THE WORLD OF MATHEMATICS

Volume 3

EDITED BY
JAMES NEWMAN

PARTS VIII-XVII

The World of MATHEMATICS

Volume 3

Edited by
James R. Newman

DOVER PUBLICATIONS, INC.
Mineola, New York

Published in Canada by General Publishing Company, Ltd., 30 Lesmill Road, Don Mills, Toronto, Ontario.

Bibliographical Note

This Dover edition, first published in 2000, is an unabridged republication of the 4-volume work originally published by Simon & Schuster, New York, in 1956.
Page iii constitutes a continuation of this copyright page.

Library of Congress Cataloging-in-Publication Data

The world of mathematics / edited by James R. Newman.
p. cm.
Originally published: New York : Simon and Schuster, 1956.
Includes bibliographical references and indexes.
ISBN 0-486-41153-2 (pbk. : v. 1)—ISBN 0-486-41150-8 (pbk. : v. 2)—ISBN 0-486-41151-6 (pbk. : v. 3)—ISBN 0-486-41152-4 (pbk. : v. 4)
1. Mathematics. I. Newman, James Roy, 1907–1966.

QA7 .W67 2000
510—dc21

00-027006

Manufactured in the United States of America
Dover Publications, Inc., 31 East 2nd Street, Mineola, N.Y. 11501

COMMENTARY ON

An Ingenious Army Captain and on a Generous and Many-sided Man

STATISTICS was founded by John Graunt of London, a "haberdasher of small-wares," in a tiny book called *Natural and Political Observations made upon the Bills of Mortality*.[1] It was the first attempt to interpret mass biological phenomena and social behavior from numerical data—in this case, fairly crude figures of births and deaths in London from 1604 to 1661. Graunt's tract appeared in 1662. Thirty years later, the Royal Society published in its "Philosophical Transactions" a paper on mortality rates written by the eminent astronomer Edmund Halley. This famous article was entitled "An Estimate of the Degrees of the Mortality of Mankind, drawn from curious Tables of the Births and Funerals at the City of Breslaw; with an Attempt to ascertain the Prices of Annuities upon Lives." It was followed by "Some further Considerations on the Breslaw Bills of Mortality." Together, the papers are the foundation for all later work on life expectancy, indispensable of course to the solvency of life-insurance companies.[2]

John Graunt was born in 1620 in Birchin Lane, London, "at the Sign of the Seven Stars," where his father kept a shop and home. He was early apprenticed to a merchant in small wares—buttons, needles and the like—and prospered in the trade. Success gave him the leisure to indulge interests somewhat broader than those of the notions counter. Aubrey describes him as "a very ingenious and studious person . . . [who] rose early in the morning to his Study before shoptime."[3] He became a friend of Sir William Petty, later the author of a well-known book on the new study of political arithmetic, and probably discussed with him the ideas to be expressed in the *Observations*. The Bills of Mortality which attracted Graunt's attention were issued weekly by the company of parish clerks and listed the number of deaths in each parish, the causes, and also an "Accompt of all the Burials and Christnings,

[1] The full title is, *Natural and Political Observations Mentioned in a following Index, and made upon the Bills of Mortality.*

[2] "He not only gave a sound analysis of this problem (the calculation of annuity prices), but he put his results in such a convenient form that this first table of mortality has remained the pattern for all subsequent tables, as to its fundamental form of expression."—Lowell J. Reed in the introduction to *Degrees of Mortality of Mankind* by Edmund Halley, a reprint of the papers noted, issued by the Johns Hopkins Press, Baltimore, 1942; p. iv. The selection by Halley is based on this reprint.

[3] *Aubrey's Brief Lives*, edited by Oliver Lawson Dick; London, 1950, p. 114.

hapning that Week." They are described fully in the material selected from Graunt's book.

Charles II was so favorably impressed by the *Observations* that he specially proposed Graunt as an original member of the newly incorporated Royal Society. To forestall any possible objections on the ground that Graunt was a shopkeeper, "his Majesty gave this particular charge to his Society, that if they found any more such Tradesmen, they should be sure to admit them all, without any more ado."[4] He was elected F. R. S. in 1662.

The merit of the *Observations* was immediately recognized and encouraged the gathering and study of vital statistics on the Continent—particularly in France—as well as in England. The book went through several editions, the fifth of which, published after Graunt's death, was enlarged by Petty. Historians have long been vexed to decide how much Petty contributed to the original work. Aubrey, who delighted in retailing small malices, says only that Graunt had his "Hint" from Petty, but he implies much more. There seems no doubt that the book was a joint production. Graunt wrote by far the greater part, including the most valuable scientific portions; Petty, it may be supposed, added what Thomas Browne would have called "elegancy" and thereby increased the popularity of the book. Sir William was a bumptious and somewhat inflated man, unable to decide whether to patronize Graunt or to claim credit for his work. There is no evidence that he even understood the importance and originality of what his friend had done.[5] The last sentence of the preface is unmistakably Graunt's: "For herein I have, like a silly Scholeboy, coming to say my Lesson to the World (that Peevish, and Tetchie Master) brought a bundle of Rods wherewith to be whipt, for every mistake I have committed."

Graunt served as a member of the city common council and in other offices, but on turning Catholic—he was raised a Puritan—"layd down trade and all other publique Employment." Aubrey tells us that he was a man generally beloved, "a faythfull friend," prudent and just. "He had an excellent working head, and was very facetious and fluent in his conversation." He was accused of having had "some hand" in the great fire of London, and the fact that he was a Catholic gave impetus to the charge. It was said that, as an officer of a water company, he had given orders stopping the water supply just before the fire started. A diligent eighteenth-century historian proved this false by showing that Graunt had

[4] Tho. Sprat, *The History of the Royal Society of London, for the improving of Natural Knowledge*; 3rd Edition, London, 1722, p. 67.

[5] For a meticulous sifting of the evidence as to Graunt vs. Petty see the introduction to a reprint of the *Observations* (Baltimore, The Johns Hopkins Press, 1939), by Walter F. Willcox. As to Petty, no inconsiderable person even if he was inflated and bumptious, see E. Strauss, *Sir William Petty, Portrait of a Genius*, Glencoe (Ill.), 1954.

had no connection with the company until a month after the fire. Graunt died of jaundice on Easter-eve 1674, and was buried "under the piewes" in St. Dunstan's church. "What pitty 'tis," wrote Aubrey, "so great an Ornament of the Citty should be buryed so obscurely!"

* * * * *

Unlike poor Graunt, whom my edition of the *Britannica* does not deign even to notice, Edmund Halley has been amply celebrated. I shall dispose of him as briefly as possible. He was born in London in 1658, the son of a wealthy "Soape-boyler," and he enjoyed every advantage, including an excellent education, that rich and indulgent parents could confer. His passion for mathematics and astronomy showed itself in his youth: when he arrived at Queen's College, Oxford, he brought with him a large assortment of astronomical instruments, including a 24-foot telescope, whose use he had already mastered. His reputation as a theoretician and observer was established by the time he was twenty. He left the college before finishing his course, to make southern hemisphere observations at St. Helena. On his return, and by the King's command, he was awarded a Master of Arts degree; a few days later he was elected a Fellow of the Royal Society. He was then twenty-two. The next few years were spent on various astronomical labors which required him to travel widely on the Continent. Becoming deeply interested in the problem of gravity, he visited Newton at Cambridge in August 1684. It was a momentous meeting, for it resulted in the *Principia*, a work which might never have appeared except for Halley's extraordinary exertions. He suggested the project in the first place; he averted suppression of the third book; he bore all the expenses of printing and binding, corrected the proofs, and laid his own work entirely aside to see Newton's masterpiece through the press. The expense was assumed at a time when Halley could ill afford it. His father had suffered serious reverses before he died and had left an encumbered and almost worthless estate.

Halley's long life was crowded with literary and scientific activity. He was a classical scholar, hydrographer, mathematician, physicist, and astronomer. His writings include, besides a vast output in his specialty, such diverse items as "An Account of the Circulation of the Watery Vapours of the Sea, and of the Cause of Springs"; "Discourse tending to prove at what Time and Place Julius Caesar made his first Descent upon Britain"; "New and General Method of finding the Roots of Equations"; a translation from the Arabic—which language he learned for this purpose—of Apollonius' treatise *De sectione rationis* and a brilliant restoration of his two lost books *De sectione spatii*; an admirable edition of the same author's *Conics*; and more than eighty miscellaneous papers published by the Royal Society, which he served as secretary. In 1698

he commanded the war-sloop Paramour Pink in an expedition to the South Atlantic to study variations of the compass and to find new lands, if possible. On this journey he "fell in with great islands of ice, of so incredible a height and magnitude that I scarce dare write my thoughts of it." He was made Savilian professor of geometry at Oxford in 1703 and astronomer royal in 1721. One of his greatest achievements was a study of the orbits of comets, of which he described no less than twenty-four. Three of these were so much alike that he was convinced that the comets of 1531, 1607, and 1682 were one body. Assuming its period to be seventy-six years, he predicted its return in 1758. On Christmas Day of that year his conjecture was verified, and Halley's comet has since appeared in 1835 and 1910.

Halley died at the age of eighty-six. He was a generous, easygoing person, "free from rancor or jealousy," who spoke and acted with an "uncommon degree of sprightliness and vivacity." He enjoyed his work, had excellent health and owned a large circle of friends, among them Peter the Great of Russia to whose table he always had access. Bishop Berkeley thought Halley an "infidel," and it is true that in 1691 he was refused the Savilian professorship of astronomy at Oxford because of his alleged "materialistic views." The evidence is that he was a sensible man who spoke his mind freely—a dangerous practice in any age.

Halley's concern with the "curious tables" of Breslaw was one of his lesser diversions. This Silesian city had, for more than a century before his entry into the problem, kept regular records of its births and deaths. Dr. Caspar Neumann, a scientist and clergyman of Breslaw had analyzed some of these data, "disproving certain current superstitions with regard to the effect of the phases of the moon and the so-called 'climacteric' years, on health." [6] His results were submitted to Leibniz who sent them to the Royal Society. It was at about this time that the Society resumed publication of the "Transactions" after a lapse of several years. Halley promised to furnish five sheets in twenty of the forthcoming issues. He was never hard up for ideas, nor for the energy and ingenuity to express them. His Breslaw papers may therefore be regarded as a kind of filler for the "Transactions," to keep his word until something better came along. Nevertheless, the analysis reflects the exceptional power of his mind.

[6] Lowell J. Reed, *op. cit.*

Natural and *Political*

OBSERVATIONS

Mentioned in a following INDEX,

and made upon the

Bills of Mortality.

By *JOHN GRAUNT*,

Citizen of

LONDON.

With reference to the *Government*, *Religion*, *Trade*, *Growth*, *Ayre*, *Diseases*, and the several Changes of the said CITY.

—— *Non, me ut miretur Turba, laboro,*
Contentus paucis Lectoribus ——

LONDON,
Printed by *Tho: Roycroft*, for *John Martin*, *James Allestry*, and *Tho: Dicas*, at the Sign of the *Bell* in St. *Paul's* Church-yard, MDCLXII.

Our days on earth are as a shadow, and there is none abiding.
—I CHRONICLES XXIX

And so from hour to hour we ripe and ripe,
And then, from hour to hour, we rot and rot;
And thereby hangs a tale. —SHAKESPEARE (*As You Like It*)

Let nature and let art do what they please,
When all is done, life's an incurable disease. —ABRAHAM COWLEY

1 Foundations of Vital Statistics

By JOHN GRAUNT

TO THE RIGHT HONOURABLE JOHN LORD ROBERTS,
BARON OF TRURO, LORD PRIVIE-SEAL,
AND ONE OF HIS MAJESTIE'S MOST HONOURABLE PRIVIE COUNCIL.

My Lord,

AS the favours I have received from your Lordship oblige me to present you with some token of my *gratitude*: so the especial Honour I have for your Lordship hath made me *sollicitous* in the choice of the *Present*. For, if I could have given your Lordship any choice *Excerptions* out of the *Greek*, or Latine Learning, I should (according to our *English Proverb*) thereby but carry *Coals to Newcastle*, and but give your Lordship *Puddle-water*, who, by your own eminent *Knowledge* in those learned *Languages*, can drink out of the very *Fountains* your self.

Moreover, to present your Lordship with tedious *Narrations*, were but to speak my own *Ignorance* of the *Value*, which his Majesty, and the Publick have of your Lordship's Time. And in brief, to offer any thing like what is already in other Books, were but to derogate from your Lordship's learning, which the World knows to be universal, and unacquainted with few useful things contained in any of them.

Now having (I know not by what accident) engaged my thoughts upon the *Bills of Mortality*, and so far succeeded therein, as to have reduced several great confused *Volumes* into a few perspicuous *Tables*, and abridged such *Observations* as naturally flowed from them, into a few succinct *Paragraphs*, without any long Series of *multiloquious Deductions*, I have presumed to sacrifice these my small, but first publish'd, *Labours* unto your Lordship, as unto whose benigne acceptance of some other of my *Papers*, even the Birth of these is due; hoping (if I may without vanity say it) they may be of as much use to Persons in your Lordship's place, as they are of little or none to me, which is no more then the fairest

Diamonds are to the *Journey-man Jeweller* that works them, or the poor *Labourer* that first dig'd them from the Earth. For with all humble submission to your Lordship, I conceive, That it doth not ill-become a *Peer of the Parliament*, or *Member of his Majestie's Council*, to consider how few starve of the many that beg: That the irreligious *Proposals* of some, to multiply People by *Polygamy*, is withall irrational, and fruitless: That the troublesome seclusions in the *Plague-time* is not a remedy to be purchased at vast inconveniences: That the greatest *Plagues* of the City are equally, and quickly repaired from the Country: That the wasting of *Males* by Wars, and Colonies do not prejudice the due proportion between them and *Females*: That the Opinions of *Plagues* accompanying the Entrance of *Kings* is false, and seditious: That *London*, the *Metropolis* of *England*, is perhaps a Head too big for the Body, and possibly too strong: That this Head grows three times as fast as the Body unto which it belongs, that is, It doubles its People in a third part of the time: That our *Parishes* are now grown madly disproportionable: That our *Temples* are not sutable to our *Religion*: That the *Trade*, and very *City of London* removes *Westward*: That the walled City is but a one fifth of the whole Pyle: That the old Streets are unfit for the present frequencie of *Coaches*: That the passage of *Ludgate* is a throat too straight for the Body: That the fighting men about *London*, are able to make three as great Armies as can be of use in this *Island*: That the number of Heads is such, as hath certainly much deceived some of our *Senatours* in their appointments of *Pole-money*, &c. Now, although your Lordship's most excellent Discourses have well informed me, That your Lordship is no stranger to all these *Positions*; yet because I knew not that your Lordship had ever deduced them from the *Bills of Mortality*; I hoped it might not be ungratefull to your Lordship, to see unto how much profit that one Talent might be improved, besides the many curiosities concerning the waxing, and waning of Diseases, the relation between *Healthfull*, and *fruitfull Seasons*, the difference between the City and Country *Air*, &c. All which, being new, to the best of my knowledge, and the whole Pamphlet, not two hours reading, I did make bold to trouble your Lordship with a perusal of it, and by this humble Dedication of it, let your Lordship and the world see the Wisdom of our City, in appointing, and keeping these Accompts, and with how much affection and success I am

My Lord,

Your Lordship's most obedient, and most faithfull Servant,

Birchin-Lane,
25 *January* 166½.

JOHN GRAUNT.

THE PREFACE

Having been born, and bred in the City of *London*, and having always observed, that most of them who constantly took in the weekly Bills of *Mortality*, made little other use of them, then to look at the foot, how the *Burials* increased, or decreased; And, among the *Casualties*, what had happened rare, and extraordinary in the week currant: so as they might take the same as a *Text* to talk upon, in the next Company; and withall, in the *Plague-time*, how the *Sickness* increased, or decreased, that so the *Rich* might judge of the necessity of their removal, and *Trades-men* might conjecture what doings they were like to have in their respective dealings:

Now, I thought that the Wisdom of our City had certainly designed the laudable practice of takeing, and distributing these Accompts, for other, and greater uses then those above-mentioned, or at least, that some other uses might be made of them: And thereupon I casting mine Eye upon so many of the General *Bills*, as next came to hand, I found encouragement from them, to look out all the *Bills* I could, and (to be short) to furnish my self with as much matter of that kind, even as the Hall of the *Parish-Clerks* could afford me; the which, when I had reduced into Tables (the Copies whereof are here inserted) so as to have a view of the whole together, in order to the more ready comparing of one *Year*, *Season*, *Parish*, or other *Division* of the City, with another, in respect of all the *Burials*, and *Christnings*, and of all the *Diseases*, and *Casualties* happening in each of them respectively; I did then begin, not onely to examine the Conceits, Opinions, and Conjectures, which upon view of a few scattered *Bills* I had taken up; but did also admit new ones, as I found reason, and occasion from my *Tables*.

Moreover, finding some *Truths*, and not commonly-believed Opinions, to arise from my Meditations upon these neglected *Papers*, I proceeded further, to consider what benefit the knowledge of the same would bring to the World; that I might not engage my self in idle, and useless Speculations, but like those Noble *Virtuosi of Gresham-Colledge* (who reduce their subtile Disquisitions upon Nature into downright Mechanical uses) present the World with some real fruit from those ayrie Blossoms.

How far I have succeeded in the Premisses, I now offer to the World's censure. Who, I hope, will not expect from me, not professing Letters, things demonstrated with the same certainty, wherewith Learned men determine in their *Scholes*; but will take it well, that I should offer at a new thing, and could forbear presuming to meddle where any of the Learned Pens have ever touched before, and that I have taken the pains, and been at the charge, of setting out those *Tables*, whereby all men

may both correct my *Positions*, and raise others of their own: For herein I have, like a silly Scholeboy, coming to say my Lesson to the World (that Peevish, and Tetchie Master) brought a bundle of Rods wherewith to be whipt, for every mistake I have committed.

OF THE BILLS OF MORTALITY, THEIR BEGINNING, AND PROGRESS

The first of the continued weekly *Bills* of *Mortality* extant at the Parish-Clerks *Hall*, begins the 29. of *December*, 1603, being the first year of *James* his Reign; since when, a weekly Accompt hath been kept there of *Burials* and *Christnings*. It is true, There were *Bills* before, *viz.* for the years 1592, -93, -94. but so interrupted since, that I could not depend upon the sufficiencie of them, rather relying upon those Accompts which have been kept since, in order, as to all the uses I shall make of them.

I believe, that the rise of keeping these Accompts, was taken from the *Plague*: for the said *Bills* (for ought appears) first began in the said year 1592. being a time of great *Mortality*; And after some disuse, were resumed again in the year 1603, after the great *Plague* then happening likewise.

These *Bills* were Printed and published, not onely every week on *Thursdays*, but also a general Accompt of the whole Year was given in, upon the *Thursday* before *Christmas Day*: which said general Accompts have been presented in the several manners following, *viz.* from the Year 1603, to the Year 1624, *inclusive* . . .

We have hitherto described the several steps, whereby the *Bills* of *Mortality* are come up to their present state; we come next to shew how they are made, and composed, which is in this manner, *viz.* When any one dies, then, either by tolling, or ringing of a Bell, or by bespeaking of a Grave of the *Sexton*, the same is known to the *Searchers*, corresponding with the said *Sexton*.

The *Searchers* hereupon (who are antient Matrons, sworn to their Office) repair to the place, where the dead Corps lies, and by view of the same, and by other enquiries, they examine by what *Disease*, or *Casualty* the Corps died. Hereupon they make their Report to the *Parish-Clerk*, and he, every *Tuesday* night, carries in an Accompt of all the *Burials*, and *Christnings*, hapning that Week, to the *Clerk* of the *Hall*. On *Wednesday* the general Accompt is made up, and Printed, and on *Thursdays* published, and dispersed to the several Families, who will pay four shillings *per Annum* for them. . . .

24

The Diseases, and Casualties this year being 1632.

Abortive, and Stilborn ..	445	Grief	11
Affrighted	1	Jaundies	43
Aged	628	Jawfaln	8
Ague	43	Impostume	74
Apoplex, and Meagrom	17	Kil'd by several accidents..	46
Bit with a mad dog.....	1	King's Evil.....	38
Bleeding	3	Lethargie	2
Bloody flux, scowring, and flux	348	Livergrown	87
Brused, Issues, sores, and ulcers, ...	28	Lunatique	5
Burnt, and Scalded.....	5	Made away themselves.....	15
Burst, and Rupture.....	9	Measles	80
Cancer, and Wolf.....	10	Murthered	7
Canker	1	Over-laid, and starved at nurse	7
Childbed	171	Palsie	25
Chrisomes, and Infants.....	2268	Piles.....	1
Cold, and Cough.....	55	Plague.....	8
Colick, Stone, and Strangury	56	Planet	13
Consumption	1797	Pleurisie, and Spleen.....	36
Convulsion	241	Purples, and spotted Feaver	38
Cut of the Stone.....	5	Quinsie	7
Dead in the street, and starved	6	Rising of the Lights.....	98
Dropsie, and Swelling.....	267	Sciatica	1
Drowned	34	Scurvey, and Itch.....	9
Executed, and prest to death	18	Suddenly	62
Falling Sickness.....	7	Surfet	86
Fever	1108	Swine Pox	6
Fistula	13	Teeth	470
Flocks, and small Pox.....	531	Thrush, and Sore mouth...	40
French Pox.....	12	Tympany	13
Gangrene	5	Tissick	34
Gout	4	Vomiting	1
		Worms	27

Christened	Males....4994 Females..4590 In all....9584	Buried	Males....4932 Females..4603 In all....9535	Whereof, of the Plague.8

Increased in the Burials in the 122 Parishes, and at the Pest-house this year..... 993

Decreased of the Plague in the 122 Parishes, and at the Pest-house this year..... 266 [10]

GENERAL OBSERVATIONS UPON THE CASUALTIES

In my Discourses upon these *Bills* I shall first speak of the *Casualties*, then give my Observations with reference to the *Places*, and *Parishes* comprehended in the *Bills*; and next of the *Years*, and *Seasons*.

1. There seems to be good reason, why the *Magistrate* should himself take notice of the numbers of *Burials*, and *Christnings*, viz. to see, whether the City increase or decrease in people; whether it increase proportionably with the rest of the Nation; whether it be grown big enough, or too big, *&c.* But why the same should be made known to the People, otherwise then to please them as with a curiosity, I see not.

2. Nor could I ever yet learn (from the many I have asked, and those not of the least *Sagacity*) to what purpose the distinction between *Males* and *Females* is inserted, or at all taken notice of; or why that of *Marriages* was not equally given in? Nor is it obvious to everybody, why the Accompt of *Casualties* (whereof we are now speaking) is made? The reason, which seems most obvious for this latter, is, That the state of health in the City may at all times appear.

3. Now it may be Objected, That the same depends most upon the Accompts of *Epidemical Diseases*, and upon the chief of them all, the *Plague*; wherefore the mention of the rest seems onely matter of curiosity.

4. But to this we answer; That the knowledg even of the numbers, which die of the *Plague*, is not sufficiently deduced from the meer Report of the *Searchers*, which onely the Bills afford; but from other Ratiocinations, and comparings of the *Plague* with some other *Casualties*.

5. For we shall make it probable, that in Years of *Plague* a quarter part more dies of that *Disease* then are set down; the same we shall also prove by the other *Casualties*. Wherefore, if it be necessary to impart to the World a good Accompt of some few *Casualties*, which since it cannot well be done without giving an Accompt of them all, then is our common practice of so doing very apt, and rational.

6. Now, to make these Corrections upon the perhaps, ignorant, and careless *Searchers* Reports, I considered first of what Authority they were in themselves, that is, whether any credit at all were to be given to their Distinguishments: and finding that many of the *Casualties* were but matter of sense, as whether a Childe were *Abortive*, or *Stilborn*; whether men were *Aged*, that is to say, above sixty years old, or thereabouts, when they died, without any curious determination, whether such *Aged* persons died purely of *Age*, as for that the *Innate heat* was quite extinct, or the *Radical moisture* quite dried up (for I have heard some Candid *Physicians* complain of the darkness, which themselves were in hereupon) I say, that

these Distinguishments being but matter of sense, I concluded the *Searchers* Report might be sufficient in the Case.

7. As for *Consumptions*, if the *Searchers* do but truly Report (as they may) whether the dead Corps were very lean, and worn away, it matters not to many of our purposes, whether the Disease were exactly the same, as *Physicians* define it in their Books. Moreover, In case a man of seventy five years old died of a *Cough* (of which had he been free, he might have possibly lived to ninety) I esteem it little errour (as to many of our purposes) if this Person be, in the Table of *Casualties*, reckoned among the *Aged*, and not placed under the Title of *Coughs*.

8. In the matter of *Infants* I would desire but to know clearly, what the *Searchers* mean by *Infants*, as whether Children that cannot speak, as the word *Infants* seems to signifie, or Children under two or three years old, although I should not be satisfied, whether the *Infant* died of *Winde*, or of *Teeth*, or of the *Convulsion*, &c. or were choak'd with *Phlegm*, or else of *Teeth*, *Convulsion*, and *Scowring*, apart or together, which, they say, do often cause one another: for, I say, it is somewhat, to know how many die usually before they can speak, or how many live past any assigned number of years.

9. I say, it is enough, if we know from the *Searchers* but the most predominant Symptomes; as that one died of the *Head-Ache*, who was sorely tormented with it, though the *Physicians* were of Opinion, that the Disease was in the *Stomach*. Again, if one died *suddenly*, the matter is not great, whether it be reported in the Bills, *Suddenly*, *Apoplexie*, or *Planet-strucken*, &c.

10. To conclude, In many of these cases the *Searchers* are able to report the Opinion of the *Physician*, who was with the Patient, as they receive the same from the Friends of the Defunct, and in very many cases, such as *Drowning*, *Scalding*, *Bleeding*, *Vomiting*, *making-away them selves*, *Lunatiques*, *Sores*, *Small-Pox*, &c. their own senses are sufficient, and the generality of the World, are able prettie well to distinguish the *Gowt*, *Stone*, *Dropsie*, *Falling-Sickness*, *Palsie*, *Agues*, *Plurisy*, *Rickets*, &c. one from another.

11. But now as for those Casualties, which are aptest to be confounded, and mistaken, I shall in the ensuing Discourse presume to touch upon them so far, as the Learning of these Bills hath enabled me.

12. Having premised these general Advertisements, our first Observation upon the *Casualties* shall be, that in twenty Years there dying of all diseases and *Casualties*, 229250. that 71124. dyed of the *Thrush*, *Convulsion*, *Rickets*, *Teeth*, and *Worms*; and as *Abortives*, *Chrysomes*, *Infants*, *Liver-grown*, and *Over-laid*; that is to say, that about ⅓. of the whole died

of those Diseases, which we guess did all light upon Children under four or five Years old.

13. There died also of the *Small-Pox*, *Swine-Pox*, and *Measles*, and of *Worms* without *Convulsions*, 12210. of which number we suppose likewise, that about ½. might be Children under six Years old. Now, if we consider that 16. of the said 229 thousand died of that extraordinary and grand *Casualty* the *Plague*, we shall finde that about thirty six *per centum* of all quick conceptions, died before six years old.

14. The second Observation is; That of the said 229250. dying of all Diseases, there died of acute Diseases (the *Plague* excepted) but about 50000. or 2/9 parts. The which proportion doth give a measure of the state, and disposition of this *Climate*, and *Air*, as to health, these *acute*, and *Epidemical* Diseases happening suddenly, and vehemently, upon the like corruptions, and alterations in the *Air*.

15. The third Observation is, that of the said 229. thousand about 70. died of *Chronical* Diseases, which shews (as I conceive) the state, and disposition of the Country (including as well it's *Food*, as *Air*) in reference to health, or rather to *longævity*: for as the proportion of *Acute* and *Epidemical* Diseases shews the aptness of the *Air* to suddain and vehement Impressions, so the *Chronical* Diseases shew the ordinary temper of the Place, so that upon the proportion of *Chronical* Diseases seems to hang the judgment of the fitness of the Country for *long Life*. For, I conceive, that in Countries subject to great *Epidemical* sweeps men may live very long, but where the proportion of the *Chronical* distempers is great, it is not likely to be so; because men being long sick, and alwayes sickly, cannot live to any great age, as we see in several sorts of *Metalmen*, who although they are less subject to acute Diseases then others,

Table of notorious Diseases.

Apoplex	1306
Cut of the Stone	0038
Falling Sickness	0074
Dead in the Streets	0243
Gowt	0134
Head-Ach	0051
Jaundice	0998
Lethargy	0067
Leprosy	0006
Lunatique	0158
Overlaid, and Starved	0529
Palsy	0423
Rupture	0201
Stone and Strangury	0863
Sciatica	0005
Sodainly	0454

Table of Casualties.

Bleeding	069
Burnt, and Scalded	125
Drowned	829
Excessive drinking	002
Frighted	022
Grief	279
Hanged themselves	222
Kil'd by several accidents	1021
Murthered	0086
Poysoned	014
Smothered	026
Shot	007
Starved	051
Vomiting	136

yet seldome live to be old, that is, not to reach unto those years, which *David* saies is the age of man.

16. The fourth Observation is; That of the said 229000. not 4000. died of outward Griefs, as of *Cancers, Fistulaes, Sores, Ulcers, broken and bruised Limbs, Impostumes, Itch, King's-evil, Leprosie, Scald-head, Swine-Pox, Wens,* &c. *viz.* not one in 60.

17. In the next place, whereas many persons live in great fear, and apprehension of some of the more formidable, and notorious diseases following; I shall onely set down how many died of each: that the respective numbers, being compared with the Total 229250, those persons may the better understand the hazard they are in.

OF PARTICULAR CASUALTIES

My first Observation is, That few are *starved*. This appears, for that of the 229250 which have died, we find not above fifty one to have been *starved*, excepting helpless *Infants* at Nurse, which being caused rather by carelessness, ignorance, and infirmity of the Milch-women, is not properly an effect, or sign of want of food in the Countrey, or of means to get it.

The Observation, which I shall add hereunto, is, That the vast numbers of *Beggars*, swarming up and down this City, do all live, and seem to be most of them healthy and strong; whereupon I make this Question, Whether, since they do all live by Begging, that is, without any kind of labour; it were not better for the State to keep them, even although they earned nothing; that so they might live regularly, and not in that Debauchery, as many Beggars do; and that they might be cured of their bodily Impotencies, or taught to work, *&c.* each according to his condition, and capacity; or by being employed in some work (not better undone) might be accustomed, and fitted for labour. . . .

My next Observation is; That but few are *Murthered*, viz. not above 86 of the 22950 [*sic*]. which have died of other diseases, and casualties; whereas in *Paris* few nights scape without their *Tragedie*.

The Reasons of this we conceive to be *Two:* One is the *Government*, and *Guard* of the City by *Citizens* themselves, and that alternately. No man settling into a Trade for that employment. And the other is, The natural, and customary abhorrence of that inhumane *Crime*, and all *Bloodshed* by most *Englishmen:* for of all that are *Executed* few are for *Murther*. Besides the great and frequent Revolutions, and Changes of *Government* since the year 1650, have been with little *bloodshed*; the *Usurpers* themselves having *Executed* few in comparison, upon the Accompt of disturbing their Innovations.

In brief, when any dead Body is found in *England*, no *Algebraist*, or *Uncipherer* of Letters, can use more subtile suppositions, and varietie of conjectures to finde out the Demonstration, or Cipher; then every common unconcerned Person doth to finde out the Murtherers, and that for ever, untill it be done.

The *Lunaticks* are also but few, *viz.* 158 in 229250. though I fear many more then are set down in our *Bills*, few being entred for such, but those who die at *Bedlam*; and there all seem to die of their *Lunacie*, who died *Lunaticks*; for there is much difference in computing the number of *Lunaticks*, that die (though of *Fevers*, and all other Diseases, unto which *Lunacie* is no *Supersedeas*) and those, that die by reason of their *Madness*.

So that, this *Casualty* being so uncertain, I shall not force my self to make any inference from the numbers, and proportions we finde in our Bills concerning it: onely I dare ensure any man at this present, well in his Wits, for one in the thousand, that he shall not die a *Lunatick* in *Bedlam*, within these seven years, because I finde not above one in about one thousand five hundred have done so.

The like use may be made of the Accompts of men, that made away themselves, who are another sort of Madmen, that think to ease themselves of pain by leaping into *Hell*; or else are yet more Mad, so as to think there is no such place; or that men may go to rest by death, though they die in *self-murther*, the greatest Sin.

We shall say nothing of the numbers of those, that have been *Drowned, Killed by falls from Scaffolds*, or by *Carts running over them*, &c. because the same depends upon the casual Trade, and Employment of men, and upon matters, which are but circumstantial to the Seasons, and Regions we live in; and affords little of that Science, and Certainty we aim at.

We finde one *Casualty* in our Bills, of which though there be daily talk, there is little effect, much like our abhorrence of *Toads*, and *Snakes*, as most poisonous Creatures, whereas few men dare say upon their own knowledge, they ever found harm by either; and this *Casualty* is the *French-Pox*, gotten, for the most part, not so much by the intemperate use of *Venery* (which rather causeth the *Gowt*) as of many common Women.

I say, the Bills of *Mortality* would take off these Bars, which keep some men within bounds, as to these extravagancies: for in the afore-mentioned 229250 we finde not above 392 to haved died of the *Pox*. Now, forasmuch as it is not good to let the World be lulled into a security, and belief of Impunity by our Bills, which we intend shall not be onely as *Death's-heads* to put men in minde of their *Mortality*, but also as *Mercurial*

Statues to point out the most dangerous ways, that lead us into it, and misery. We shall therefore shew, that the *Pox* is not as the *Toads*, and *Snakes* afore-mentioned, but of a quite contrary nature, together with the reason, why it appears otherwise.

17. Forasmuch as by the ordinary discourse of the world it seems a great part of men have, at one time, or other, had some *species* of this disease, I wondering why so few died of it, especially because I could not take that to be so harmless, whereof so many complained very fiercely; upon inquiry I found that those who died of it out of the Hospitals (especially that of *King's-Land*, and the *Lock* in *Southwark*) were returned of *Ulcers*, and *Sores*. And in brief I found, that all mentioned to die of the *French-Pox* were retured by the *Clerks* of Saint *Giles's*, and Saint *Martin*'s *in the Fields* onely; in which place I understood that most of the vilest, and most miserable houses of uncleanness were: from whence I concluded, that onely *hated* persons, and such, whose very *Noses* were eaten of, were reported by the *Searchers* to have died of this too frequent *Maladie*. . . .

OF THE DIFFERENCE BETWEEN THE NUMBERS OF MALES, AND FEMALES

The next Observation is, That there be more *Males* then *Females*.

There have been *Buried* from the year 1628, to the year 1662, *exclusive*, 209436 *Males*, and but 190474 *Females:* but it will be objected, that in *London* it may indeed be so, though otherwise elsewhere; because *London* is the great Stage and Shop of business, wherein the *Masculine Sex* bears the greatest part. But we Answer, That there have been also *Christned* within the same time, 139782 *Males*, and but 130866 *Females*, and that the Country Accompts are consonant enough to those of *London* upon this matter.

What the Causes hereof are, we shall not trouble our selves to conjecture, as in other Cases, onely we shall desire, that Travellers would enquire whether it be the same in other Countries.

We should have given an Accompt, how in every Age these proportions change here, but that we have Bills of distinction but for 32 years, so that we shall pass from hence to some inferences from this Conclusion; as first,

I. That *Christian Religion*, prohibiting *Polygamy*, is more agreeable to the *Law of Nature*, that is, the *Law of God*, then *Mahumetism*, and others, that allow it; for one man his having many women, or wives by Law, signifies nothing, unless there were many women to one man in Nature also.

II. The obvious Objection hereunto is, That one *Horse*, *Bull*, or *Ram*, having each of them many *Females*, do promote increase. To which I Answer, That although perhaps there be naturally, even of these *species*, more *Males* then *Females*, yet *artificially*, that is, by making *Geldings*, *Oxen*, and *Weathers*, there are fewer. From whence it will follow, That when by experience it is found how many *Ews* (suppose twenty) one *Ram* will serve, we may know what proportion of *male-Lambs* to castrate, or geld, *viz.* nineteen, or thereabouts: for if you emasculate fewer, *viz.* but ten, you shall by promiscuous copulation of each of those ten with two *Females*, (in such as admit the *Male* after conception) hinder the increase so far, as the admittance of two *Males* will do it: but, if you castrate none at all, it is highly probable, that every of the twenty *Males* copulating with every of the twenty *Females*, there will be little, or no conception in any of them all.

III. And this I take to be the truest Reason, why *Foxes*, *Wolves*, and other *Vermin Animals* that are not gelt, increase not faster than *Sheep*, when as so many thousands of these are daily Butchered, and very few of the other die otherwise then of themselves.

We have hitherto said there are more *Males*, then *Females;* we say next, That the one exceed the other by about a thirteenth part; so that although more men die violent deaths then women, that is, more are *slain* in Wars, *killed* by mischance, *drowned* at *Sea*, and die by the *Hand of Justice*. Moreover, more men go to *Colonies*, and travel into foreign parts, then women. And lastly, more remain unmarried, then of women, as *Fellows* of *Colleges*, and *Apprentises*, above eighteen, *&c.* yet the said thirteenth part difference bringeth the business but to such a pass, that every woman may have an Husband, without the allowance of *Polygamy*.

Moreover, although a man be *Prolifique* fourty years, and a woman but five and twenty, which makes the *Males* to be as 560 to 325 *Females*, yet the causes above named, and the later marriage of the men, reduce all to an equality. . . .

It is a Blessing to Man-kind, that by this overplus of *Males* there is this natural Bar to *Polygamy*: for in such a state Women could not live in that parity, and equality of expence with their Husbands, as now, and here they do.

The reason whereof is, not, that the Husband cannot maintain as splendidly three, as one; for he might, having three Wives, live himself upon a quarter of his Income, that is in a parity with all three, as-well as, having but one, live in the same parity at half with her alone: but rather, because that to keep them all quiet with each other, and himself, he must keep them all in greater awe, and less splendor which power he having will probably use it to keep them all as low, as he pleases, and at no more

cost then makes for his own pleasure; the poorest Subjects (such as this plurality of Wives must be) being most easily governed.

THE CONCLUSION

It may be now asked, to what purpose tends all this laborious buzzling, and groping? To know,

1. The number of the People?
2. How many *Males*, and *Females*?
3. How many Married, and single?
4. How many *Teeming* Women?
5. How many of every *Septenary*, or *Decad* of years in *age*?
6. How many *Fighting* Men?
7. How much *London* is, and by what steps it hath increased?
8. In what time the housing is replenished after a *Plague*?
9. What proportion die of each general and perticular *Casualties*?
10. What years are Fruitfull, and Mortal, and in what Spaces, and Intervals, they follow each other?
11. In what proportion Men neglect the Orders of the *Church*, and *Sects* have increased?
12. The disproportion of Parishes?
13. Why the Burials in *London* exceed the Christnings, when the contrary is visible in the Country?

To this I might answer in general by saying, that those, who cannot apprehend the reason of these Enquiries, are unfit to trouble themselves to ask them.

I might answer by asking; Why so many have spent their times, and estates about the Art of making Gold? which, if it were much known, would onely exalt Silver into the place, which Gold now possesseth; and if it were known but to some one Person, the same single *Adeptus* could not, nay, durst not enjoy it, but must be either a Prisoner to some Prince, and Slave to some Voluptuary, or else skulk obscurely up and down for his privacie, and concealment.

I might Answer; That there is much pleasure in deducing so many abstruse, and unexpected inferences out of these poor despised Bills of *Mortality*; and in building upon that ground, which hath lain waste these eighty years. And there is pleasure in doing something new, though never so little, without pestering the World with voluminous Transcriptions.

But, I Answer more seriously; by complaining, That whereas the Art of Governing, and the true *Politiques*, is how to preserve the Subject in *Peace*, and *Plenty*, that men study onely that part of it, which teacheth

how to supplant, and over-reach one another, and how, not by fair outrunning, but by tripping up each other's heels, to win the Prize.

Now, the Foundation, or Elements of this honest harmless *Policy* is to understand the Land, and the hands of the Territory to be governed, according to all their intrinsick, and accidental differences: as for example; It were good to know the *Geometrical* Content, Figure, and Scituation of all the Lands of a Kingdom, especially, according to its most natural, permanent, and conspicuous Bounds. It were good to know, how much Hay an Acre of every sort of Meadow will bear? how many Cattel the same weight of each sort of Hay will feed, and fatten? what quantity of Grain, and other Commodities the same Acre will bear in one, three, or seven years *communibus Annis?* unto what use each soil is most proper? All which particulars I call the intrinsick value: for there is also another value meerly accidental, or extrinsick, consisting of the Causes, why a parcel of Land, lying near a good Market, may be worth double to another parcel, though but of the same intrinsick goodness; which answers the Queries, why Lands in the North of *England* are worth but sixteen years purchase, and those of the West above eight and twenty. It is no less necessary to know how many People there be of each Sex, State, Age, Religion, Trade, Rank, or Degree, *&c.* by the knowledg whereof Trade, and Government may be made more certain, and Regular; for, if men knew the People as aforesaid, they might know the consumption they would make, so as Trade might not be hoped for where it is impossible. As for instance, I have heard much complaint, that Trade is not set up in some of the *South-western*, and *North-western* Parts of *Ireland*, there being so many excellent Harbours for that purpose, whereas in several of those Places I have also heard, that there are few other Inhabitants, but such as live *ex sponte creatis*, and are unfit Subjects of Trade, as neither employing others, nor working themselves.

Moreover, if all these things were clearly, and truly known (which I have but guessed at) it would appear, how small a part of the People work upon necessary Labours, and Callings, *viz.* how many Women, and Children do just nothing, onely learning to spend what others get? how many are meer Voluptuaries, and as it were meer Gamesters by Trade? how many live by puzling poor people with unintelligible Notions in Divinity, and Philosophie? how many by perswading credulous, delicate, and Litigious Persons, that their Bodies, or Estates are out of Tune, and in danger? how many by fighting as Souldiers? how many by Ministeries of Vice, and Sin? how many by Trades of meer Pleasure, or Ornaments? and how many in a way of lazie attendance, *&c.* upon others? And on the other side, how few are employed in raising, and working necessary food, and covering? and of the speculative men, how few do truly studie *Nature*,

and *Things*? The more ingenious not advancing much further then to write, and speak wittily about these matters.

I conclude, That a clear knowledge of all these particulars, and many more, whereat I have shot but at rovers, is necessary in order to good, certain, and easie Government, and even to balance Parties, and factions both in *Church* and *State*. But whether the knowledge thereof be necessary to many, or fit for others, then the Sovereign, and his chief Ministers, I leave to consideration.

PHILOSOPHICAL TRANSACTIONS:

Giving some

ACCOUNT

OF THE

Present *Undertakings, Studies* and *Labours*

OF THE

INGENIOUS;

In many

Considerable Parts of the WORLD.

VOL. XVII. For the Year 1693.

LONDON:

Printed for *S. Smith* and *B. Walford*, Printers to the *Royal Society*, at the *Prince's Arms*, in St. *Paul's* Church-yard. 1694.

Factual science may collect statistics and make charts. But its predictions are, as has been well said, but past history reversed. —JOHN DEWEY

2 First Life Insurance Tables

By EDMUND HALLEY

AN ESTIMATE OF THE DEGREES OF THE MORTALITY OF MANKIND, DRAWN FROM CURIOUS TABLES OF THE BIRTHS AND FUNERALS AT THE CITY OF BRESLAW; WITH AN ATTEMPT TO ASCERTAIN THE PRICE OF ANNUITIES UPON LIVES. BY MR. E. HALLEY, R.S.S.

THE Contemplation of the *Mortality* of *Mankind*, has besides the *Moral*, its *Physical* and *Political Uses*, both which have been some years since most judiciously considered by the curious Sir *William Petty*, in his *Natural* and *Political* Observations on the Bills of *Mortality* of *London*, owned by Captain *John Graunt*. And since in a like Treatise on the Bills of *Mortality* of *Dublin*. But the Deduction from those Bills of *Mortality* seemed even to their Authors to be defective: First, In that the *Number* of the People was wanting. Secondly, That the *Ages* of the People dying was not to be had. And Lastly, That both *London* and *Dublin* by reason of the great and casual Accession of *Strangers* who die therein, (as appeared in both, by the great Excess of the *Funerals* above the *Births*) rendred them incapable of being Standards for this purpose; which requires, if it were possible, that the People we treat of should not at all be changed, but die where they were born, without any Adventitious Increase from Abroad, or Decay by Migration elsewhere.

This *Defect* seems in a great measure to be satisfied by the late curious Tables of the Bills of *Mortality* at the City of *Breslaw*, lately communicated to this Honourable Society by Mr. *Justell*, wherein both the *Ages* and *Sexes* of all that die are monthly delivered, and compared with the number of the *Births*, for Five Years last past, *viz.* 1687, 88, 89, 90, 91, seeming to be done with all the Exactness and Sincerity possible.

This City of *Breslaw* is the Capital City of the Province of *Silesia*; or, as the *Germans* call it, *Schlesia*, and is scituated on the Western Bank of the River *Oder*, anciently called *Viadrus*; near the Confines of *Germany* and *Poland*, and very nigh the Latitude of *London*. It is very far from the Sea, and as much a *Mediterranean Place* as can be desired, whence the Confluence of Strangers is but small, and the Manufacture of Linnen employs chiefly the poor People of the place, as well as of the Country round about: whence comes that sort of Linnen we usually call your

Sclesie Linnen; which is the chief, if not the only Merchandize of the place. For these Reasons the People of this City seem most proper for a *Standard*; and the rather, for that the *Births* do, a small matter, exceed the *Funerals*. The only thing wanting is the Number of the whole People, which in some measure I have endeavoured to supply by comparison of the *Mortality* of the People of all Ages, which I shall from the said Bills trace out with all the Acuracy possible.

It appears that in the Five Years mentioned, *viz*. from 87 to 91 inclusive, there were *born* 6193 Persons, and *buried* 5869; that is, born *per Annum* 1238, and *buried* 1174; whence an *Encrease* of the People may be argued of 64 *per Annum*, or of about a 20th part, which may perhaps be ballanced by the Levies for the *Emperor*'s Service in his Wars. But this being contingent, and the Births certain, I will suppose the People of *Breslaw* to be encreased by 1238 *Births* annually. Of these it appears by the same Tables, that 348 do die *yearly* in the *first Year* of their *Age*, and that but 890 do arrive at a full *Years Age*; and likewise, that 198 do die in the *Five Years* between 1 and 6 compleat, taken at a *Medium*; so that but 692 of the Persons *born* do survive *Six* whole *Years*. From this *Age* the Infants being arrived at some degree of Firmness, grow less and less *Mortal*; and it appears that of the whole People of *Breslaw* there die *yearly*, as in the following Table, wherein the upper Line shews the *Age*, and the next under it the *Number* of Persons of that Age *dying yearly*.

7	8	9	.	14	.	18	.	21	.	27	28	.	35	36
11	11	6	5½	2	3½	5	6	4½	6½	9	8	7	7	8

	42		45		49	54	55	56	.	63		70	71	72
9½	8	9	7	7	10	11	9	9	10	12	9½	14	9	11

	77		81		84		90	91	98	99	100.
9½	6	7	3	4	2	1	1	1	0	⅕	⅗

And where no *Figure* is placed over, it is to be understood of those that die between the Ages of the preceding and consequent *Column*.

From this Table it is evident, that from the Age of 9 to about 25 there does not die above 6 *per Annum* of each *Age*, which is much about one *per Cent*. of those that are of those *Ages:* And whereas in the 14, 15, 16, 17 *Years* there appear to die much fewer, as 2 and 3½, yet that seems rather to be attributed to Chance, as are the other Irregularities in the Series of Ages, which would rectifie themselves, were the number of Years much more considerable, as 20 instead of 5. And by our own Experience in *Christ-Church Hospital*, I am informed there die of the *Young Lads*, much about one *per Cent. per Annum*, they being of the foresaid *Ages*. From 25 to 50 there seem to die from 7 to 8 and 9 *per Annum* of each Age; and after that to 70, they growing more *crasie*,

though the number be much diminished, yet the *Mortality encreases*, and there are found to die 10 or 11 of each Age *per Annum:* From thence the number of the *Living* being grown very small, they gradually decline till there be none left to *die*; as may be seen at one View in the Table.

From these Considerations I have formed the *adjoyned Table*, whose Uses are manifold, and give a more just *Idea* of the *State* and *Condition* of *Mankind*, than any thing yet extant that I know of. It exhibits the *Number* of *People* in the City of *Breslaw* of all Ages, from the *Birth* to extream *Old Age*, and thereby shews the Chances of *Mortality* at all *Ages*, and likewise how to make a certain Estimate of the value of *Annuities* for *Lives*, which hitherto has been only done by an imaginary *Valuation:* Also the *Chances* that there are that a *Person* of any *Age* proposed does live to any other *Age* given; with many more, as I shall hereafter shew. This *Table* does shew the *number* of *Persons* that are living in the *Age* current annexed thereto, as follows:

Age Curt.	Persons	Age Curt.	Persons	Age Curt.	Persons	Age Curt.	Persons	Age Curt.	Persons	Age Curt.	Persons
1	1000	8	680	15	628	22	586	29	539	36	481
2	855	9	670	16	622	23	579	30	531	37	472
3	798	10	661	17	616	24	573	31	523	38	463
4	760	11	653	18	610	25	567	32	515	39	454
5	732	12	646	19	604	26	560	33	507	40	445
6	710	13	640	20	598	27	553	34	499	41	436
7	692	14	634	21	592	28	546	35	490	42	427
Age Curt.	Persons	Age Curt.	Persons	Age Curt.	Persons	Age Curt.	Persons	Age Curt.	Persons	Age Curt.	Persons
43	417	50	346	57	272	64	202	71	131	78	58
44	407	51	335	58	262	65	192	72	120	79	49
45	397	52	324	59	252	66	182	73	109	80	41
46	387	53	313	60	242	67	172	74	98	81	34
47	377	54	302	61	232	68	162	75	88	82	28
48	367	55	292	62	222	69	152	76	78	83	23
49	357	56	282	63	212	70	142	77	68	84	20

Age	Persons
7	5547
14	4584
21	4270
28	3964
35	3604
42	3178
49	2709
56	2194
63	1694
70	1204
77	692
84	253
100	107
	34000
	Sum Total

Thus it appears, that the whole People of *Breslaw* does consist of 34000 *Souls*, being the Sum *Total* of the Persons of all Ages in the *Table*: The first use hereof is to shew the Proportion of *Men* able to bear *Arms* in any *Multitude*, which are those between 18 and 56, rather than 16 and 60; the one being generally too weak to bear the *Fatigues* of *War* and the Weight of *Arms*, and the other too crasie and infirm from *Age*, notwithstanding particular Instances to the contrary. Under 18 from the *Table*, are found in this City 11997 Persons, and 3950 above 56, which together make 15947. So that the Residue to 34000 being 18053 are Persons between those *Ages*. At least one half thereof are Males, or 9027: So that

the whole Force this City can raise of *Fencible Men*, as the *Scotch* call them, is about 9000, or 9/34, or somewhat more than a quarter of the *Number* of *Souls*, which may perhaps pass for a Rule for all other places.

The *Second Use* of this Table is to shew the differing degrees of *Mortality*, or rather *Vitality* in all *Ages*; for if the number of Persons of any *Age* remaining after one year, be divided by the difference between that and the number of the Age proposed, it shews the *odds* that there is, that a Person of that Age does not die in a *Year*. As for Instance, a Person of 25 *Years* of *Age* has the odds of 560 to 7 or 80 to 1, that he does not *die* in a *Year*: Because that of 567, living of 25 years of Age, there do die no more than 7 in a *Year*, leaving 560 of 26 Years old.

So likewise for the *odds*, that any Person does not die before he attain any proposed *Age*: Take the *number* of the remaining Persons of the Age proposed, and divide it by the difference between it and the number of those of the *Age* of the Party proposed; and that shews the *odds* there is between the Chances of the Party's living or dying. As for Instance; What is the *odds* that a Man of 40 lives 7 Years: Take the number of Persons of 47 years, which in the Table is 377, and subtract it from the number of Persons of 40 years, which is 445, and the *difference* is 68: Which shews that the *Persons dying* in that 7 years are 68, and that it is 377 to 68 or 5½ to 1, that a Man of 40 does live 7 Years. And the like for any other *number* of *Years*.

Use III. But if it be enquired at what number of *Years*, it is an even Lay that a Person of any *Age* shall die, this Table readily performs it: For if the *number* of Persons *living* of the *Age* proposed be *halfed*, it will be found by the *Table* at what Year the said *number* is reduced to half by *Mortality*; and that is the Age, to which it is an even Wager, that a Person of the *Age* proposed shall arrive before he *die*. As for Instance; A Person of 30 Years of *Age* is proposed, the number of that Age is 531, the half thereof is 265, which number I find to be between 57 and 58 Years; so that a Man of 30 may reasonably expect to live between 27 and 28 Years.

Use IV. By what has been said, the *Price* of *Insurance* upon *Lives* ought to be regulated, and the difference is discovered between the *price* of ensuring the *Life* of a *Man* of 20 and 50, for Example: it being 100 to 1 that a Man of 20 dies not in a year, and but 38 to 1 for a Man of 50 Years of Age.

Use V. On this depends the Valuation of *Annuities* upon *Lives*; for it is plain that the *Purchaser* ought to pay for only such a part of the value of the *Annuity*, as he has Chances that he is living; and this ought to be computed yearly, and the Sum of all those yearly Values being added together, will amount to the value of the *Annuity* for the *Life* of the Person proposed. Now the present value of Money payable after a term

of years, at any given rate of Interest, either may be had from Tables already computed; or almost as compendiously, by the Table of Logarithms: For the Arithmetical Complement of the Logarithm of Unity and its yearly Interest (that is, of 1, 06 for Six *per Cent.* being 9, 974694.) being multiplied by the number of years proposed, gives the present value of One Pound payable after the end of so many years. Then by the foregoing Proposition, it will be as the number of Persons living after that term of years, to the number dead; so are the Odds that any one Person is Alive or Dead. And by consequence, as the Sum of both or the number of Persons living of the Age first proposed, to the number remaining after so many years, (both given by the Table) so the present value of the yearly Sum payable after the term proposed, to the Sum which ought to be paid for the Chance the person has to enjoy such an Annuity after so many Years. And this being repeated for every year of the persons Life, the Sum of all the present Values of those Chances is the true Value of the Annuity. This will without doubt appear to be a most laborious Calculation, but it being one of the principal Uses of this Speculation, and having found some Compendia for the Work, I took the pains to compute the following Table, being the short Result of a not ordinary number of Arithmetical Operations; It shews the Value of Annuities for every Fifth Year of Age, to the Seventieth, as follows.

Age	Years Purchase	Age	Years Purchase	Age	Years Purchase
1	10,28	25	12,27	50	9,21
5	13,40	30	11,72	55	8,51
10	13,44	35	11,12	60	7,60
15	13,33	40	10,57	65	6,54
20	12,78	45	9,91	70	5,32

This shews the great Advantage of putting Money into the present *Fund* lately granted to their Majesties, giving 14 *per Cent. per Annum*, or at the rate of 7 years purchase for a Life; when young Lives, at the usual rate of Interest, are worth above 13 years Purchase. It shews likewise the Advantage of young Lives over those in Years; a Life of Ten Years being almost worth 13½ years purchase, whereas one of 36 is worth but 11.

Use VI. Two Lives are likewise valuable by the same Rule; for the number of Chances of each single Life, found in the Table, being multiplied together, become the Chances of the Two Lives. And after any certain Term of Years, the Product of the two remaining Sums is the Chances that both the Persons are living. The Product of the two Differences, being the numbers of the Dead of both Ages, are the Chances that both the Persons are dead. And the two Products of the remaining Sums of the one Age multiplied by those dead of the other, shew the

Chances that there are that each Party survives the other: Whence is derived the Rule to estimate the value of the Remainder of one Life after another. Now as the Product of the Two Numbers in the Table for the Two Ages proposed, is to the difference between that Product and the Product of the two numbers of Persons deceased in any space of time, so is the value of a Sum of Money to be paid after so much time, to the value thereof under the Contingency of Mortality. And as the aforesaid Product of the two Numbers answering to the Ages proposed, to the Product of the Deceased of one Age multiplied by those remaining alive of the other; So the Value of a Sum of Money to be paid after any time proposed, to the value of the Chances that the one Party has that he survives the other whose number of Deceased you made use of, in the second Term of the proportion. This perhaps may be better understood, by putting N for the number of the younger Age, and n for that of the Elder; Y, y the deceased of both Ages respectively, and R, r for the Remainders; and $R + Y = N$ and $r + y = n$. Then shall $N\,n$ be the whole number of Chances; $N\,n - Y\,y$ be the Chances that one of the two Persons is living, $Y\,y$ the Chances that they are both dead; $R\,y$ the Chances that the elder Person is dead and the younger living; and $r\,Y$ the Chances that the elder is living and the younger dead. Thus two Persons of 18 and 35 are proposed, and after 8 years these Chances are required. The Numbers for 18 and 35 are 610 and 490, and there are 50 of the First Age dead in 8 years, and 73 of the Elder Age. There are in all 610×490 or 298900 Chances; of these there are 50×73 or 3650 that they are both dead. And as 298900, to $298900 - 3650$, or 295250: So is the present value of a Sum of Money to be paid after 8 years, to the present value of a Sum to be paid if either of the two live. And as 560×73, so are the Chances that the Elder is dead, leaving the Younger; and as 417×50, so are the Chances that the Younger is dead, leaving the Elder. Wherefore as 610×490 to 560×73, so is the present value of a Sum to be paid at eight years end, to the Sum to be paid for the Chance of the Youngers Survivance; and as 610×490 to 417×50, so is the same present value to the Sum to be paid for the Chance of the Elders Survivance.

This possibly may be yet better explained by expounding these Products by Rectangular Parallelograms, as in *Figure 1*, wherein $A\,B$ or $C\,D$ represents the number of persons of the younger Age, and $D\,E$, $B\,H$ those remaining alive after certain term of years; whence $C\,E$ will answer the number of those dead in that time: So $A\,C$, $B\,D$ may represent the number of the Elder Age; $A\,F$, $B\,I$ the Survivors after the same term; and $C\,F$, $D\,I$, those of that Age that are dead at that time. Then shall the whole Parallelogram $A\,B\,C\,D$ be $N\,n$, or the Product of the two Numbers of persons, representing such a number of Persons of the two Ages given; and by what was said before, after the Term proposed the Rectangle $H\,D$

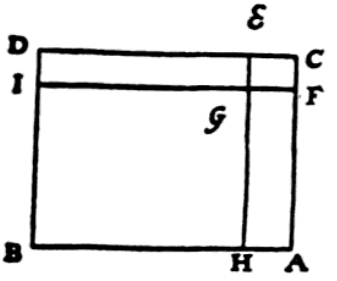

FIGURE 1

shall be as the number of Persons of the younger Age that survive, and the Rectangle *A E* as the number of those that die. So likewise the Rectangles *A I*, *F D* shall be as the Numbers, living and dead, of the other Age. Hence the Rectangle *H I* shall be as an equal number of both Ages surviving. The Rectangle *F E* being the Product of the deceased, or *Y y*, an equal number of both dead. The Rectangle *G D* or *R y*, a number living of the younger Age, and dead of the Elder: And the Rectangle *A G* or *r Y* a number living of the Elder Age, but dead of the younger. This being understood, it is obvious, that as the whole Rectangle *A D* or *N n* is to the *Gnomon F A B D E G* or $N n - Y y$, so is the whole number of Persons or Chances, to the number of Chances that one of the two Persons is living: And as *A D* or *N n* is to *F E* or *Y y*, so are all the Chances, to the Chances that both are dead; whereby may be computed the value of the Reversion after both Lives. And as *A D* to *G D* or *R y*, so the whole number of Chances, to the Chances that the younger is living and the other dead; whereby may be cast up what value ought to be paid for the Reversion of one Life after another, as in the case of providing for Clergy-mens Widows and others by such Reversions. And as *A D* to *A G* or *r Y*, so are all the Chances, to those that the Elder survives the younger. I have been the more particular, and perhaps tedious, in this matter, because it is the Key to the Case of Three Lives, which of it self would not have been so easie to comprehend.

VII. If Three Lives are proposed, to find the value of an Annuity during the continuance of any of those three Lives. The Rule is, *As the Product of the continual multiplication of the Three Numbers, in the Table, answering to the Ages proposed, is to the difference of that Product and of the Product of the Three Numbers of the deceased of those Ages, in any given term of Years; So is the present value of a Sum of Money to be paid certainly after so many Years, to the present value of the same Sum to be paid, provided one of those three Persons be living at the Expiration of that term.* Which proportion being yearly repeated, the Sum of all those present values will be the value of an Annuity granted for

three such Lives. But to explain this, together with all the Cases of Survivance in three Lives: Let N be the Number in the Table for the Younger Age, n for the Second, and ν for the Elder Age; let Y be those dead of the Younger Age in the term proposed, y those dead of the Second Age, and υ those of the Elder Age; and let R be the Remainder of the younger Age, r that of the middle Age, and $\int$ the Remainder of the Elder Age. Then shall $R + Y$ be equal to N, $r + y$ to n, and $\int + \upsilon$ to ν, and the continual Product of the three Numbers $N\ n\ \nu$ shall be equal to the continual Product of $R + Y \times r + Y \times \int + \upsilon$,[1] which being the whole number of Chances for three Lives is compounded of the eight Products following. (1) $R\ r\ \int$, which is the number of Chances that all three of the Persons are living. (2) $r\ \int\ Y$, which is the number of Chances that the two Elder Persons are living, and the younger dead. (3) $R\ \rho\ y$ the number of Chances that the middle Age is dead, and the younger and Elder living. (4) $R\ r\ \upsilon$ being the Chances that the two younger are living, and the elder dead. (5) $\int\ Y\ y$ the Chances that the two younger are dead, and the elder living. (6) $r\ Y\ \upsilon$ the Chances that the younger and elder are dead, and the middle Age living. (7) $R\ y\ \upsilon$, which are the Chances that the younger is living, and the two other dead. And Lastly and Eightly, $Y\ y\ \upsilon$, which are the Chances that all three are dead. Which latter subtracted from the whole number of Chances $N\ n\ \nu$, leaves $N\ n\ \nu - Y\ y\ \upsilon$ the Sum of all the other Seven Products; in all of which one or more of the three Persons are surviving:

To make this yet more evident, I have added *Figure 2*, wherein these Eight several Products are at one view exhibited. Let the rectangled Parallelepipedon $A\ B\ C\ D\ E\ F\ G\ H$ be constituted of the sides $A\ B$, $G\ H$, &c. proportional to N the number of the younger Age; $A\ C$, $B\ D$, &c. proportional to n; and $A\ G$, $C\ E$, &c. proportional to the number of the Elder, or ν. And the whole Parallelepipedon shall be as the Product $N\ n\ \nu$, or our whole number of Chances. Let $B\ P$ be as R, and $A\ P$ as Y: let $C\ L$ be as r, and $L\ n$ as y; and $G\ N$ as $\int$, and $N\ A$ as υ; and let the Plain $P\ R\ e\ a$ be made parallel to the plain $A\ C\ G\ E$; the plain $N\ V\ b\ Y$ parallel to $A\ B\ C\ D$; and the plain $L\ X\ T\ Q$ parallel to the plain $A\ B\ G\ H$. And our first Product $R\ r\ \int$ shall be as the Solid $S\ T\ W\ I\ F\ Z\ e\ b$. The Second, or $r\ \int\ Y$ will be as the Solid $E\ Y\ Z\ e\ Q\ S\ M\ I$. The Third, $R\ \int\ y$, as the Solid $R\ H\ O\ V\ W\ I\ S\ T$. And the Fourth, $R\ r\ \upsilon$, as the Solid $Z\ a\ b\ D\ W\ X\ I\ K$. Fifthly, $\int\ Y\ y$, as the Solid $G\ Q\ R\ S\ I\ M\ N\ O$. Sixthly, $r\ Y\ \upsilon$, as $I\ K\ L\ M\ G\ Y\ Z\ A$. Seventhly, $R\ y\ \upsilon$, as the Solid $I\ K\ P\ O\ B\ X\ V\ W$. And Lastly, $A\ I\ K\ L\ M\ N\ O\ P$ will be as the Product of the 3 numbers of persons dead, or $Y\ y\ \upsilon$. I shall not apply this in all the cases thereof for brevity sake; only to shew in one how all the rest may be performed, let

[1] EDITOR'S NOTE: The manuscript for the 9th line of this page contains a misprint. It should read $R + Y \times r + y$, etc. Likewise, in line 13, $R\ \rho\ y$ should read $R\ \int\ y$.

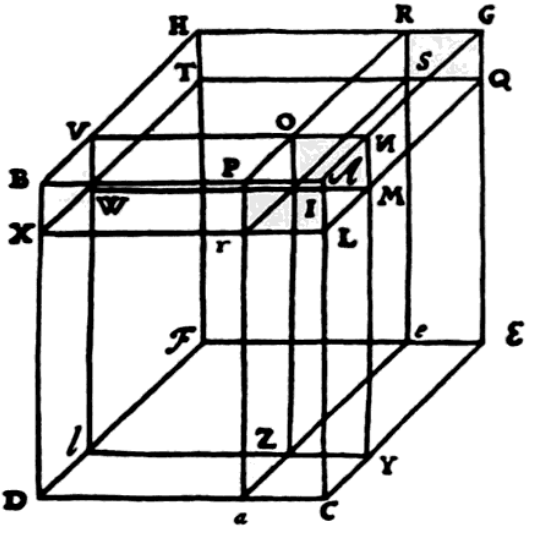

FIGURE 2

it be demanded what is the value of the Reversion of the younger Life after the two elder proposed. The proportion is as the whole number of Chances, or $N\ n\ \nu$ to the Product $R\ y\ \nu$, so is the certain present value of the Sum payable after any term proposed, to the value due to such Chances as the younger person has to bury both the elder, by the term proposed; which therefore he is to pay for. Here it is to be noted, that

Years	Present value of 1 *l.*	Years	Present value of 1 *l.*	Years	Present value of 1 *l.*
1	0,9434	19	0,3305	37	0,1158
2	0,8900	20	0,3118	38	0,1092
3	0,8396	21	0,2941	39	0,1031
4	0,7921	22	0,2775	40	0,0972
5	0,7473	23	0,2618	45	0,0726
6	0,7050	24	0,2470	50	0,0543
7	0,6650	25	0,2330	55	0,0406
8	0,6274	26	0,2198	60	0,0303
9	0,5919	27	0,2074	65	0,0227
10	0,5584	28	0,1956	70	0,0169
11	0,5268	29	0,1845	75	0,0126
12	0,4970	30	0,1741	80	0,0094
13	0,4688	31	0,1643	85	0,0071
14	0,4423	32	0,1550	90	0,0053
15	0,4173	33	0,1462	95	0,0039
16	0,3936	34	0,1379	100	0,0029
17	0,3714	35	0,1301		
18	0,3503	36	0,1227		

the first term of all these Proportions is the same throughout, viz. $N\ n\ v$. The Second changing yearly according to the Decrease of R, r, $\int$, and Encrease of Y, y, v. And the third are successively the present values of Money payable after one, two, three, *&c.* years, according to the rate of Interest agreed on. These numbers, which are in all cases of Annuities of necessary use, I have put into the following Table, they being the Decimal values of One Pound payable after the number of years in the Margent, at the rate of 6 *per Cent.* . . .

I. SOME FURTHER CONSIDERATIONS ON THE BRESLAW BILLS OF MORTALITY. BY THE SAME HAND, &C.

S I R,

What I gave you in my former Discourse on these Bills, was chiefly designed for the Computation of the Values of Annuities on Lives, wherein I believe I have performed what the short Period of my Observations would permit, in relation to exactness, but at the same time do earnestly desire, that their Learned Author Dr. *Newman* of *Breslaw* would please to continue them after the same manner for yet some years further, that so the casual Irregularities and apparent Discordance in the Table, [p. 1438] may by a certain number of Chances be rectified and ascertain'd.

Were this *Calculus* founded on the Experience of a very great number of Years, it would be very well worth the while to think of Methods for facilitating the Computation of the Value of two, three, or more Lives; which as proposed in my former, seems (as I am inform'd) a Work of too much Difficulty for the ordinary Arithmetician to undertake. I have sought, if it were possible, to find a Theorem that might be more concise than the Rules there laid down, but in vain; for all that can be done to expedite it, is by Tables of Logarithms ready computed, to exhibit the *Rationes* of N to Y in each single Life, for every third, fourth or fifth Year of Age, as occasion shall require; and these Logarithms being added to the Logarithms of the present Value of Money payable after so many Years, will give a Series of Numbers, the Sum of which will shew the Value of the Annuity sought. However for each Number of this Series two Logarithms for a single Life, three for two Lives, and four for three Lives, must necessarily be added together. If you think the matter, under the uncertainties I have mentioned, to deserve it, I shall shortly give you such a Table of Logarithms as I speak of, and an Example or two of the use thereof: But by Vulgar Arithmetick the labour of these Numbers were immense; and nothing will more recommend the useful Invention of Logarithms to all Lovers of Numbers, than the advantage of Dispatch in this and such like Computations.

Besides the uses mentioned in my former, it may perhaps not be an unacceptable thing to infer from the same Tables, how unjustly we repine at the shortness of our Lives, and think our selves wronged if we attain not Old Age; whereas it appears hereby, that the one half of those that are born are dead in Seventeen years time, 1238 being in that time reduced to 616. So that instead of murmuring at what we call an untimely Death, we ought with Patience and unconcern to submit to that Dissolution which is the necessary Condition of our perishable Materials, and of our nice and frail Structure and Composition: And to account it as a Blessing that we have survived, perhaps by many Years, that Period of Life, whereat the one half of the whole Race of Mankind does not arrive.

A second Observation I make upon the said Table, is that the Growth and Encrease of Mankind is not so much stinted by any thing in the Nature of the *Species*, as it is from the cautious difficulty most People make to adventure on the state of *Marriage*, from the prospect of the Trouble and Charge of providing for a Family. Nor are the poorer sort of People herein to be blamed, since their difficulty of subsisting is occasion'd by the unequal Distribution of Possessions, all being necessarily fed from the Earth, of which yet so few are Masters. So that besides themselves and Families, they are yet to work for those who own the Ground that feeds them: And of such does by very much the greater part of Mankind consist; otherwise it is plain, that there might well be four times as many Births as we now find. For by computation from the Table, I find that there are nearly 15000 Persons above 16 and under 45, of which at least 7000 are Women capable to bear Children. Of these notwithstanding there are but 1238 born yearly, which is but little more than a sixth part: So that about one in six of these Women do breed yearly; whereas were they all married, it would not appear strange or unlikely, that four of six should bring a Child every year. The Political Consequences hereof I shall not insist on, only the Strength and Glory of a King being in the multitude of his Subjects, I shall only hint, that above all things, Celibacy ought to be discouraged, as, by extraordinary Taxing and Military Service: And those who have numerous Families of Children to be countenanced and encouraged by such Laws as the *Jus trium Liberorum* among the *Romans*. But especially, by an effectual Care to provide for the Subsistence of the Poor, by finding them Employments, whereby they may earn their Bread, without being chargeable to the Publick.

COMMENTARY ON

The Law of Large Numbers

IN Jacob Bernoulli's famous book, *Ars Conjectandi*, appears a theorem of cardinal significance to the theory of probability. Usually called Bernoulli's Theorem, it is also known as the Law of Large Numbers, a name given to it by the French mathematician, Siméon Poisson (1781–1840). This theorem was the first attempt to deduce statistical measures from individual probabilities and Bernoulli claimed that it took him twenty years to perfect it. The time was not ill spent considering the central importance of the result, but mathematicians, scientists and philosophers have since then devoted many more than twenty years to examining and debating the exact meaning of the theorem and the proper range of its application in statistics.

The theorem is quite simple to state. Indeed, on first seeing it you may wonder what Bernoulli could have been stewing over for twenty years, and why it should have stirred so much controversy later on. The fact is, it is a nest of subtleties and traps; the harder one thinks about it, the more one grows uneasy. Bernoulli, of course, had his hands full making the machinery; he was entirely unaware of the logical and philosophical snares he was setting.

Here, then, is a fairly simple statement of Bernoulli's Theorem: If the probability of an event's occurrence on a single trial is p, and if a number of trials are made, independently and under the same conditions, the most probable *proportion* of the event's occurrences to the total number of trials is also p; further, the probability that the proportion in question will differ from p by less than a given amount, however small, increases as the number of trials increases.[1] By "throwing mathematical discretion to the winds," a well-known student of the subject arrives at this rough everyday definition: "If the probability of an event is p, and if an infinity of trials are made, the proportion of successes is sure to be p." [2] Nothing bad will happen to you if you adopt this as a reference point in my discussion and in reading the selection from Bernoulli. Nevertheless, I should like to make one more definition available. This one is a skillful compromise between a mathematical and a nonmathematical formulation: " 'In a sufficiently large set of α things it is almost certain that the relative

[1] For two excellent definitions, from which mine has been synthesized, see John Maynard Keynes, *A Treatise on Probability*, London, 1921, pp. 337–338; and Thornton C. Fry, *Probability and Its Engineering Uses*, New York, 1928, p. 100.

[2] Fry, *op. cit.*, p. 100.

frequency of β things will approximate to the probability of an α thing's being β within any degree of approximation which may be desired.' Here the phrase 'almost certain' is to be understood as a convenient way of saying that there is a probability as near as we like to 1." [3]

Each of these definitions raises several difficult questions. I cannot begin to state, let alone answer, them all adequately. But a few points should be set forth briefly to give the reader an inkling of the importance of Bernoulli's result and of the problems surrounding it.

1. The first thing to notice is that the theorem has to do with *proportions* or *frequencies* and not with the absolute number of times an event occurs. Increase the number of trials and it becomes increasingly likely (i.e., the probability approaches 1) that the *ratio* of successes to total trials will differ from p (the probability of the single event) by less than a fixed amount, *no matter how small*; but as the number of trials increases, it becomes increasingly unlikely (i.e., the probability approaches zero) that the *number* of successes will differ from p by less than a fixed amount, *no matter how large*.[4] This point is easily illustrated. In tossing a coin, one may assume that the probability of getting a head is ½. On this assumption the most probable number of heads in a series of trials is half the total number of tosses. It is not surprising, however, to find a deviation from this ratio: six heads in ten tosses, forty-eight in a hundred, and so on. The question is, what is the probability that in n successive throws the proportion of heads will differ from ½ by, for example, not more than $\frac{1}{10}$? According to Bernoulli's theorem, the probability approaches 1 as the number n (of successive throws) increases. In other words the probability of a *relative deviation* of at least $\frac{1}{10}$ "sinks rapidly." Six heads in ten exceeds our limit, but it is much less likely that the relative frequency of heads will differ from ½ by more than $\frac{1}{10}$ in 100 throws, and even more unlikely in 1,000 throws. At the same time, it grows more probable as n increases that the absolute deviation of the number of heads from half the total number of throws will exceed any given number. For example, there is a much greater chance of a deviation of 1 in 100 tosses (i.e., 51 or 49 heads instead of 50), than of 1 in 10 tosses (i.e., 6 or 4 heads instead of 5); of 10 in 1,000 tosses than 10 in 100; and so on. That is the essence of the theorem.[5]
2. Bernoulli's theorem is the subject of gross misconceptions, some of

[3] William Kneale, *Probability and Induction*, Oxford, 1949, p. 139.

[4] See Fry, *op. cit.*, p. 101, for an instructive table comparing facts about number and proportion of successes.

[5] See Kneale, *op. cit.*, pp. 139–140; also Ernest Nagel, *Principles of the Theory of Probability*, in *International Encyclopaedia of Unified Science*, Vol. I, no. 6, p. 35. Each of these studies presents an admirable analysis not only of Bernoulli's theorem but of current problems of probability.

them due to the picturesque name coined by Poisson. It is often supposed, for example, that the theorem is "a mysterious law of nature which guarantees that in a sufficiently large number of trials a probability will be 'realized as a frequency.' " [6] This odd opinion derives apparently from the conviction that in the long run nature is bound to imitate man. The theorem is a part of the mathematical calculus of probability. The propositions of this calculus are not statements of fact or experience. They are formal arithmetical propositions, valid in their own domain, and neither capable of validating "facts" nor of being invalidated by them.[7] It is no more to be expected that the theorem can be proved by experiment than that the multiplication table or binomial theorem can be so proved.

Another fallacious inference from the theorem is the so-called "law of averages." This is an article of faith widely and fervently adhered to. It is the basis for the belief that when a player has had a bad run at cards, his luck is bound to turn; that after red has come up five times in a row at the roulette table, it is prudent to bet on black; that if a coin falls heads three or four times in succession, the chances are that tails will come up more frequently in the next three or four throws to "even things up." The only safe thing to say about these beliefs is that the player who acts on them consistently is more likely to be ruined than the player whose guide of action is erratic impulse. Bernoulli's Theorem is itself the sole ground for expecting a particular proportion of heads in the coin-tossing game, and it is an essential condition of the theorem that the trials be independent, i.e., without influence on each other. It is patently foolish, then, to invoke the theorem that sets out from the premise that the probability of a head at every toss is ½, to prove that the probability is less than ½ after a consecutive run of heads. Yet this is the muddleheaded idea underlying all gambling systems. "In a genuine game of chance there can be no system for improving one's chances of winning. That is part of what we mean by calling it a game of chance." [8]

3. While the theorem of Bernoulli (and certain generalizations of it by Poisson, Tchebycheff, Markoff and others) has proved exceptionally

[6] Kneale, *op. cit.*, p. 139.

[7] "The calculus of probability has the same general function as a demonstrative geometry or a demonstrative arithmetic: given certain initial probabilities, the calculus of probability makes it possible to calculate the probabilities of certain properties which are related to the initial ones in various ways. Thus, arithmetic cannot tell us how many people live in either China or Japan; but if the population of China and the population of Japan are given, we can compute the combined populations of these countries. The calculus of probability functions in the same way." Nagel, *op. cit.*, p. 27.

[8] The example is from Kneale, *op. cit.*, p. 140. See also Julian Lowell Coolidge, "The Gambler's Ruin," *Annals of Mathematics*, Vol. X, Series 2, 1909; and the same author's *An Introduction to Mathematical Probability*, Oxford, 1925, pp. 52–59.

important in practical affairs—insurance, for example,[9] and in scientific research—e.g., the kinetic theory of gases, its limitations and weaknesses have repeatedly been stressed. Among the points raised most frequently are these: that the theorem applies only in special cases (e.g., where the events are truly independent) and under conditions which are the exception rather than the rule; "that a knowledge of what has occurred at some of the trials would not affect the probability of what may occur at any of the others";[10] that the law of large numbers is useless as a tool of prediction concerning "sequences of observations" unless the concept of probability itself is defined in terms of relative frequency as against the formulation employed in classical probability theory.[11] None of these limitations, it should be observed, affect the *mathematical* validity of the theorem, nor do they depreciate either its contribution to the growth of statistical theory or its continuing indispensability as an instrument of inquiry. A sound grasp of the theorem is essential to an appreciation of the mathematics of probability and statistics—which, in turn, carries an increasingly responsible part in almost every branch of modern science, in industry, commerce, government and other activities. That is why I have devoted so much space to introducing the selection which follows.

[9] "As an illustration of the importance of the law of large numbers in practical affairs it will be sufficient to mention the business of insurance. Let us suppose that the probability that a man of a certain age and constitution will die within a year is 1/10. If such an individual considers insuring his life, this is the fraction which he should bear in mind and use in making his decision. But the insurance company which offers to cover the risk of his dying within the year considers another probability derived from this probability. If there are a great many people of the same characteristics insuring their lives with that company, there is a very high probability that the company will not have to pay claims on more than about one tenth of the policies. If, therefore, the company charges in each case a premium of rather more than one tenth of the amount of the policy, it is very likely that it will have enough over after all claims are paid to meet its administrative expenses and distribute a dividend to its shareholders. The greater the number of persons insuring with the company, the greater the probability that the company's finances will remain sound, provided always that its premiums are calculated in the way described. This is the all-important consideration which distinguishes the business of an insurance company from gambling." Kneale, *op. cit.*, p. 141.

[10] Keynes, *op. cit.*, p. 342; also, generally, pp. 341–345.

[11] A strong protagonist of this view argues his case persuasively in a most interesting book written for nonmathematicians: Richard von Mises, *Probability, Statistics and Truth*, 1939. See especially Lecture Four, "The Laws of Large Numbers," pp. 156–193.

Defendit numerus: There is safety in numbers. —AUTHOR UNKNOWN

The number is certainly the cause. The apparent disorder augments the grandeur. —EDMUND BURKE (*On the Sublime and the Beautiful*)

It doesn't depend on size, or a cow would catch a rabbit.
—PENNSYLVANIA GERMAN PROVERB

3 The Law of Large Numbers[1]

By JACOB BERNOULLI

WE have now reached the point where it seems that, to make a correct conjecture about any event whatever, it is necessary only to calculate exactly the number of possible cases,[2] and then to determine how much more likely it is that one case will occur than another. But here at once our main difficulty arises, for this procedure is applicable to only a very few phenomena, indeed almost exclusively to those connected with games of chance. The original inventors of these games designed them so that all the players would have equal prospects of winning, fixing the number of cases that would result in gain or loss and letting them be known beforehand, and also arranging matters so that each case would be equally likely. But this is by no means the situation as regards the great majority of the other phenomena that are governed by the laws of nature or the will of man. In the game of dice, for instance, the number of possible cases [or throws] is known, since there are as many throws for each individual die as it has faces; moreover all these cases are equally likely when each face of the die has the same form and the weight of the die is uniformly distributed. (There is no reason why one face should come up more readily than any other, as would happen if the faces were of different shapes or part of the die were made of heavier material than the rest.) Similarly, the number of possible cases is known in drawing a white or a black ball from an urn, and one can assert that any ball is equally likely to be drawn: for it is known how many balls of each kind are in the jar, and there is no reason why this or that ball should be drawn more readily than any other. But what mortal, I ask, could ascertain the number of diseases, counting all possible cases, that afflict the human body in every one of its many parts and at every age, and say how much more likely

[1] Translated from "*Klassische Stücke der Mathematik*, ausgewählt von A. Speiser" (Zürich, 1925), pp. 90–95. The selection is from the German translation of the *Ars Conjectandi* by R. Haussner in Ostwald's *Klassiker der exakten Wissenschaften*, Leipzig, 1899, nr. 108.

[2] For "case," the correct translation of the German, one may read *result* or *outcome*. ED.

one disease is to be fatal than another—plague than dropsy, for instance, or dropsy than fever—and on that basis make a prediction about the relationship between life and death in future generations? Or who could enumerate the countless changes that the atmosphere undergoes every day, and from that predict today what the weather will be a month or even a year from now? Or again, who can pretend to have penetrated so deeply into the nature of the human mind or the wonderful structure of the body that in games which depend wholly or partly on the mental acuteness or the physical agility of the players he would venture to predict when this or that player would win or lose? These and similar forecasts depend on factors that are completely obscure, and which constantly deceive our senses by the endless complexity of their interrelationships, so that it would be quite pointless to attempt to proceed along this road.

There is, however, another way that will lead us to what we are looking for and enable us at least to ascertain *a posteriori* what we cannot determine *a priori*, that is, to ascertain it from the results observed in numerous similar instances. It must be assumed in this connection that, under similar conditions, the occurrence (or nonoccurrence) of an event in the future will follow the same pattern as was observed for like events in the past. For example, if we have observed that out of 300 persons of the same age and with the same constitution as a certain *Titius*, 200 died within ten years while the rest survived, we can with reasonable certainty conclude that there are twice as many chances that Titius also will have to pay his debt to nature within the ensuing decade as there are chances that he will live beyond that time. Similarly, if anyone has observed the weather over a period of years and has noted how often it was fair and how often rainy, or has repeatedly watched two players and seen how often one or the other was the winner, then on the basis of those observations alone he can determine in what ratio the same result will or will not occur in the future, assuming the same conditions as in the past.

This empirical process of determining the number of cases by observation is neither new nor unusual; in chapter 12 and following of *L'art de penser* [3] the author, a clever and talented man, describes a procedure that is similar, and in our daily lives we can all see the same principle at work. It is also obvious to everyone that it is not sufficient to take any single observation as a basis for prediction about some [future] event, but that a large number of observations are required. There have even been instances where a person with no education and without any previous instruction has by some natural instinct discovered—quite remarkably—that the larger the number of pertinent observations available, the smaller

[3] *La logique, ou L'art de penser*, by Antoine Arnauld and Pierre Nicole. 1662. (Makes use of Pascal, Fragment no. 14.) There are in fact *two* authors but Bernoulli makes it appear there is only one.

the risk of falling into error. But though we all recognize this to be the case from the very nature of the matter, the scientific proof of this principle is not at all simple, and it is therefore incumbent on me to present it here. To be sure I would feel that I were doing too little if I were to limit myself to proving this one point with which everyone is familiar. Instead there is something more that must be taken into consideration—something that has perhaps not yet occurred to anyone. *What is still to be investigated is whether by increasing the number of observations we thereby also keep increasing the probability that the recorded proportion of favorable to unfavorable instances will approach the true ratio, so that this probability will finally exceed any desired degree of certainty,* or whether the problem has, as it were, an asymptote. This would imply that there exists a particular degree of certainty that the true ratio has been found which can never be exceeded by any increase in the number of observations: thus, for example, we could never be more than one-half, two-thirds, or three-fourths certain that we had determined the true ratio of the cases. The following illustration will make clear what I mean: We have a jar containing 3000 small white pebbles and 2000 black ones, and we wish to determine empirically the ratio of white pebbles to the black—something we do not know—by drawing one pebble after another out of the jar, and recording how often a white pebble is drawn and often a black. (I remind you that an important requirement of this process is that you put back each pebble, after noting its color, before drawing the next one, so that the number of pebbles in the urn remains constant.) Now we ask, is it possible by indefinitely extending the trials to make it 10, 100, 1000, etc., times more probable (and ultimately "morally certain") that the ratio of the number of drawings of a white pebble to the number of drawings of a black pebble will take on the same value (3:2) as the actual ratio of white to black pebbles in the urn, than that the ratio of the drawings will take on a different value? If the answer is no, then I admit that we are likely to fail in the attempt to ascertain the number of instances of each case [i.e., the number of white and of black pebbles] by observation. But if it is true that we can finally attain moral certainty by this method[4] . . . then we can determine the number of instances *a posteriori* with almost as great accuracy as if they were known to us *a priori*. Axiom 9 [presented in an earlier chapter] shows that in our everyday lives, where moral certainty is regarded as absolute certainty, this consideration enables us to make a prediction about any event involving chance that will be no less scientific than the predictions made in games of chance. If, instead of the jar, for instance, we take the atmosphere or the human body, which conceal within themselves a multitude of the most varied processes or diseases, just as the jar conceals the pebbles, then for

[4] Bernoulli demonstrates that this is true in his next chapter.

these also we shall be able to determine by observation how much more frequently one event will occur than another.

Lest this matter be imperfectly understood, it should be noted that the ratio reflecting the actual relationship between the numbers of the cases—the ratio we are seeking to determine through observation—can never be obtained with absolute accuracy; for if this were possible, the ruling principle would be opposite to what I have asserted: that is, the more observations were made, the *smaller* the probability that we had found the correct ratio. The ratio we arrive at is only approximate: it must be defined by two limits, but these limits can be made to approach each other as closely as we wish. In the example of the jar and the pebbles, if we take two ratios, 301/200 and 299/200, 3001/2000 and 2999/2000, or any two similar ratios of which one is slightly less than 1½ and the other slightly more, it is evident that we can attain any desired degree of probability that the ratio found by our many repeated observations will lie between these limits of the ratio 1½, rather than outside them.

It is this problem that I decided to publish here, after having meditated on it for twenty years. . . .

. . . If all events from now through eternity were continually observed (whereby probability would ultimately become certainty), it would be found that everything in the world occurs for definite reasons and in definite conformity with law, and that hence we are constrained, even for things that may seem quite accidental, to assume a certain necessity and, as it were, fatefulness. For all I know that is what Plato had in mind when, in the doctrine of the universal cycle, he maintained that after the passage of countless centuries everything would return to its original state.

COMMENTARY ON

Statistics and the Lady with a Fine Palate

STATISTICS has shot up like Jack's beanstalk in the present century. And as fast as the theory has developed, politics, economics, social affairs, business and science have taken it over for their special purposes. The anthologist must survey a literature so new and so vast that even the expert can scarcely comprehend it. It is impossible to know where to begin, let alone how to make a representative choice. I have tried to subdue the problem by a gross expedient. The selections which follow do no more than discuss a few fundamental concepts; the history of statistics, the great bulk of theoretical considerations and the applications of statistical method are scarcely mentioned. This procedure is justified on grounds other than expediency. Statistics, as Tippett observes, resembles arithmetic in its impact on science and human affairs.[1] But while arithmetic "is so woven into our thinking that we use it almost subconsciously," statistics is invariably regarded as a separate branch of study, and even its basic ideas are grasped by only a small proportion of educated persons. The material presented below may clarify a few essential principles for readers who, though they recognize the importance of the subject, would not dream of working through an entire book on statistics.

The first two selections are from two very good primers, L. H. C. Tippett's *Statistics*, and M. J. Moroney's *Facts from Figures*. Tippett, a leading British statistician, explains what is meant by sampling, theory of random errors, the nature of statistical laws. Two chapters are excerpted from Moroney's fat little volume in the Pelican series. They treat the concept of averages, scatter, mean and standard deviations. Both Tippett and Moroney—the latter is a British statistician, industrial consultant and lecturer in mathematics at Leicester College of Technology and Commerce—write for the nonmathematician and express themselves with commendable clarity.

The selection by R. A. Fisher is more advanced. It has to do with the design of experiments, a branch of scientific inquiry dealing with the nature of scientific inquiry itself. This study of method represents one of the most fruitful advances of scientific thought in the past two or three decades.

The theory of experimental design does not of course represent the

[1] L. H. C. Tippett, *Statistics*, Oxford University Press, 1944, p. 178.

first effort of scientists to examine and systematize their methods of observation and inference. It has long been known, for example, that scientific measurements, however careful and precise (such as those of astronomy), never yield the same results in successive determinations. These irregularities or "errors" are regarded as unavoidable accompaniments of experiment due to small, undetectable causes. Mathematicians have lumped together the various mysterious disturbances under the name of chance and have invented brilliant theories to control the mischief and confusion caused by chance. Laplace and Gauss, shortly after the opening of the nineteenth century, laid the foundation for the theory of errors of observation, a mathematical achievement of the first order. These matters are considered in the selections from Tippett and Moroney and I need dwell on them no further.

The design of experiments is also concerned with observational error, but it moves along a somewhat broader front than the classical mathematical attack on the problem. It may help to an appreciation of the difference in approach if I say that the theory of experimental design is less "complacent" than the theory of random errors. The latter thrives on errors, merely fixing limits within which experimental results are acceptable despite variations. The former worries the experiment itself to make certain that its structure is logical, that it is broad enough to serve as a foundation for inference, that the objects studied are fully and fairly exploited, that every recognizable and avoidable source of error, however small, has in fact been eliminated. Concretely, the designer must consider such factors as the variables of the system to be examined (e.g., do they form the simplest set compatible with the objective of telling the experimenter what he really wants to know?), the adequacy and representativeness of the sample, the sources of psychological bias in subjects, instruments or experimenter, the selection of suitable controls to serve as standards of comparison for the significant variables in the main experiment, the appropriate "level of significance" for the given test (i.e., what is the minimum probability the experimenter would require "before he would be willing to admit that his observations have demonstrated a positive result"?), the value of enlarging and repeating the experiment to increase its "sensitiveness." These and other points like them apply generally to all experiments, whether laid out in the physics, chemistry, biology or psychology laboratory, the hospital, the agricultural station, the schoolroom, or the lagoon at Bikini. One feels that the earlier experimental geniuses—Galileo, Faraday, Boyle, Galton, Pasteur, to name a few—avoided almost instinctively many of the pitfalls now carefully fenced off by the modern theories of experimental procedure. At the same time, one must recognize that they had less to fret about when they performed their experiments. The great increase in the complexity

of experiments multiplies the opportunities for going astray as much as it enlarges the effectiveness and scope of scientific research.

The Fisher excerpt has to do with a specific example of testing design. Sir Ronald Fisher, one of the foremost statisticians of the century, is the pioneer of the theory of design of experiments. He is professor of genetics at Cambridge University, a Fellow of the Royal Society and a Foreign Associate of the U. S. National Academy of Sciences. His name is associated prominently with the development of elaborate experimental techniques in agriculture at the famous Rothamsted Experimental Station, with important advances in genetics and the mathematical theory of natural selection, with searching reforms and innovations of statistical method such as factorial design, the "confounding" procedure, the use of Latin Squares, the exploitation of small samples and substantial refinements of randomization. (For enlightenment as to these terms I refer the reader to the selection from Fisher and also to a recently published book by E. Bright Wilson: *Introduction to Scientific Research*, New York, 1952.) Fisher's writings include the standard textbook *Statistical Methods for Research Workers* [2] and *The Design of Experiments*, regarded as a classic work of statistics and scientific method. The second chapter of the latter book is entitled "The Principles of Experimentation, Illustrated by a Psycho-Physical Experiment." It concerns a lady who says that when a cup of tea is made with milk she is able to tell whether the tea or milk was first added to the cup. The surpassing nicety of taste displayed by this hypothetical lady provides Sir Ronald with the excuse for a most remarkable series of experiments. Fisher is not an easy writer; the presentation of this case, however, is a model of lucidity and requires no mathematics other than elementary arithmetic. It demands of the reader the ability to follow a closely reasoned argument, but it will repay the effort by giving a vivid understanding of the richness, complexity and subtlety of modern experimental method.

[2] Eleventh Revised Edition, 1951, Oliver and Boyd Ltd., Edinburgh.

It is then, but an exceeding little Way, and in but a very few Respects, that we can trace up the natural Course of things before us, to general Laws. And it is only from Analogy, that we conclude, the Whole of it to be capable of being reduced into them; only from our seeing, that Part is so. It is from our finding, that the Course of Nature, in some Respects and so far, goes on by general Laws, that we conclude this of the Whole.

—BISHOP BUTLER

4 Sampling and Standard Error

By L. C. TIPPETT

SAMPLING

THE practice of taking a small part of a large bulk to represent the whole is fairly generally understood and widely used. The housewife will 'sample' a piece of cheese at the shop before making a purchase; and a cotton spinner will buy a bale of cotton, having seen only a small sample of it. The sample is also a very important tool of the statistician.

There are two general reasons for working with samples instead of the bulk. (1) Some appraisals of the thing in question involve destructive tests, and there is no point in appraising it if the whole is destroyed in the process; the housewife cannot eat her cheese and have it. (2) It is very much more economical to investigate a sample than the whole bulk. In social and economic work, for example, it is usually prohibitively expensive to investigate the whole field of inquiry in any detail. Even the population census, which has behind it the financial and coercive resources of the state, is made only at infrequent intervals, and the questions asked are few and comparatively simple. If a sample inquiry is made, on the other hand, it is feasible to employ experienced field workers who can collect information that is comparatively detailed and elaborate, and can ensure that the records are reasonably accurate.

Unfortunately, however, the method of inquiry by sample is somewhat mistrusted, sometimes honestly and sometimes, I suspect, because a sample has in some instance given a result the sceptic does not like. Yet a sample *may* give reliable results. For example, an earthquake disaster in 1923 interrupted the tabulation of the results of the Japanese census of 1920 and interim figures were given based on a sample containing one family in every thousand. These results agreed well with those given later when the regular tabulations were completed. Nevertheless it must be agreed that samples do not represent the bulk exactly, and that they may sometimes be much in error.

In the present discussion [of sampling methods] I shall follow the usual practice of statisticians of referring to the bulk that is being sampled as the *population*. The population in this chapter is to be thought of specially as contrasting with the sample. I shall refer only to populations consisting of recognizably discrete individuals, e.g., men or electric lamps.

The ideal sample is the simple random one in which chance alone decides which of the individuals in the population are chosen. Suppose we wish to obtain a random sample of the people of England and Wales in order to make an estimate of their average height. To do this we may, in principle, take forty-odd million exactly similar cards, one for each person, and write each person's national registration number on the appropriate card. These cards may then be put in a large churn, thoroughly mixed, and (say) one thousand cards be drawn, somewhat in the way the names are drawn for the Irish sweepstake. The thousand people whose numbers are on the cards are a random sample, and we can measure their heights, find the average, and so obtain a figure which is an estimate of the average height for the population.

To investigate the error in the average so estimated we could, again in principle, subsequently measure the heights of all individuals in the population and so obtain the true average. An easier thing to do is to draw a number of samples, each of one thousand, and calculate the several averages. These will vary above and below the true, or population value, and the extent to which they vary gives some idea of the error with which any one sample estimates the true average.

To do such an experiment in fact requires far greater resources than I can command, but there are other experiments that are similar in principle and are easier to do. What we really want to know is how chance works in deciding the choice of the sample, and chance also operates in games of the table, with such things as cards, dice and roulette wheels. In these games, a population does not exist in the sense that the population of England and Wales does, but we may use the concept of a *hypothetical* population. Suppose, for example, we threw a perfectly balanced six-sided die millions of times. We should expect one-sixth of the throws to score aces, one-sixth to score twos, and so on, and the average score would be $\frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) = 3{\cdot}5$. These millions of throws are a population, and any thousand of them including the first thousand is a random sample. But the millions of throws need not, in fact, be made; they need only be imagined as a hypothetical population, of which any number of actual throws form a sample.

To illustrate the way in which random sampling errors arise I have made an experiment which I need not describe in exact detail. The experiment is equivalent to that described here, which is not quite so easy to perform but easier to imagine. The imagined apparatus consists of ten

packs, each of ten cards, the cards in each pack being numbered respectively 1, 2, 3 . . . 10. The packs are shuffled separately, one card is drawn from each, and the ten numbers on the cards are added to give a score. For example, the numbers might be 2, 4, 2, 10, 2, 5, 9, 2, 9, 8 and the score would then be 53. Then the cards are put back in their packs, the packs are reshuffled, and again ten cards are drawn to give another score. This is repeated, so that a large number of scores results, which are individuals from a hypothetical population consisting of the very large number of scores that could conceivably be obtained. The lowest conceivable score is 10, resulting from ten aces; the highest is 100, resulting from ten tens; and the true average score is 55. Now let us consider the results of the experiment.

It would take too much space to give in full the results of a really extensive experiment, but enough are given in Table 1 to show the kind of thing that happens. The top part of the table gives the first thirty individual scores. Chance has not given a score as high as 100 or as low as 10, as it might have done, and presumably would have done had I continued long enough with the experiment. The first thirty scores vary between 36 and 72, the range being 36. Now, in order to see what happens when we take samples and find the averages, I took 30 samples, each of ten scores. Such samples are far too small for most statistical inquiries (although statisticians sometimes have to be content with small

TABLE 1

Individual Scores and Average Scores in Samples of Ten and Forty

Individual Scores									
52	46	72	53	36	55	42	56	61	53
56	65	48	54	62	65	48	65	61	60
58	42	58	46	63	61	68	53	54	43
Averages of Samples of Ten									
52·6	58·4	54·6	52·6	48·6	54·0	52·8	50·8	46·0	55·8
53·4	59·4	55·0	56·2	61·6	53·6	54·2	56·8	52·3	54·0
56·7	55·2	56·3	52·3	53·8	57·8	55·9	61·8	58·6	49·2
Averages of Samples of Forty									
54·6	51·6	53·6	56·6	54·3	55·1	57·3	54·4	56·0	55·4
55·3	54·1	55·8	55·4	56·0	53·2	55·1	54·3	54·8	54·2
54·3	57·2	53·2	56·0	54·5	51·5	53·7	56·0	54·8	55·4

samples) but they illustrate the errors of random sampling. The average scores are in the middle section of Table 1. The first average of 52·6 is obtained from the ten individual scores in the top row of the table. The thirty averages vary between 46·0 and 61·8, the range being 15·8, and

no average differs from the population value of 55 by more than 9·0. In so far as these thirty samples show the variations we are likely to get in the averages of the millions of samples we could draw, we may say that the biggest error with which the average of any one sample of ten scores estimates the population average is 9·0. When I took larger samples, each of forty scores, I obtained results given in the lowest section of Table 1. They vary between 51·5 and 57·3 with a range of 5·8, and the biggest error with which any one sample of forty scores estimates the population average is $55 - 51{\cdot}5 = 3{\cdot}5$. Thus we see that the averages estimated from random samples vary among themselves and differ from the average for the population, but that the biggest error decreases as the size of the sample is increased from ten to forty; and you may take on trust that this tendency would have continued had I extended the experiment to deal with still larger samples. For example, by calculating the average of the thirty averages of samples of forty, we have the average of a single sample of 1,200 scores, which comes to 54·8—very close to the population value of 55.

These results are shown in the frequency distributions of Figure 1 where, instead of a frequency for each sub-range, there are dots, each dot representing an individual score or the average of a sample. Notice how the averages tend to be clustered more closely round the population value as the size of the sample is increased. A frequency distribution of sample averages for any given size of sample is called the *sampling distribution* of the average.

The errors of random sampling, which in an experiment like that just described show themselves as variations between sample means arise from the variation between the individuals in the original population. Other things being equal, such sampling errors are proportional to the amount of variation in the population. As an extreme example, it is easy to see that had there been no variation between the individual scores and they had all been 55, the means of all samples of all sizes would have been 55 and there would have been no sampling errors.

When the statistician thinks of the random error of the average of a sample he thinks of a whole collection of possible values of error, any one of which the given sample may have: of the sampling distribution of errors. The actual error of the given sample probably exceeds the smallest of these values; it may easily exceed the intermediate values; and it is unlikely to exceed the very largest values. There is a whole list of probabilities with which the various values of error are likely to be exceeded, and these can be calculated from a quantity called the *standard error.* The standard error is a measure of the variation in the sampling distribution analogous to the standard deviation [1] and for the statistician it sums

[1] [See selection by M. J. Moroney, p. 1487. ED.]

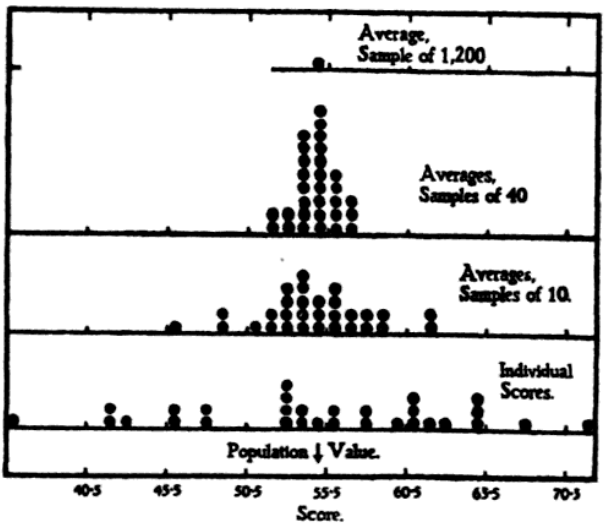

FIGURE 1—Frequency distributions giving results of sampling experiments.

up the whole distribution of errors. If the standard error of a sample is large, the errors to which that sample is liable are, as a whole, large; if the standard error is small, the likely errors are small. This quantity, carrying with it the idea of errors occurring with various probabilities, should replace the cruder 'biggest error' I introduced in describing the results of the experiment.

It is not usually necessary to do an actual experiment to measure sampling errors, as the mathematical theory of probability enables statisticians to deduce sampling distributions and standard errors theoretically. This method is better because it is less laborious and more exact, giving results as accurate as an experiment involving millions of samples. The results of the theoretical calculations are of the same kind as those given by the experiment, and in some instances they have been checked by very large-scale experiments.

I have considered only the sampling errors of the average, but the same principles apply to other statistical quantities such as ratios, and the measures of variation and correlation. The theoretical deduction of sampling distributions of the many statistical quantities in use is a very highly developed branch of mathematical statistics; and sometimes the problems have proved so difficult to solve that statisticians have had to fall back upon actual sampling experiments.

With the ability to calculate errors of sampling, statisticians can make allowances for them when making deductions from sample results. It is a standard procedure to examine the results of a sample to see how far they

can be explained by random errors. This is called *testing the significance* of the results, and only such results as cannot reasonably be attributed to errors of random sampling are held to be *statistically significant*.

Before going on to the more practical problems of sampling, I will summarize the ground covered so far. When many samples of the same size are taken from a population of variable individuals, the sample averages show variation which may be described by a sampling distribution and measured by the standard error. A given sample of that size may have any one of the averages in the distribution, and the probability that its error will exceed any stated value can be calculated from the standard error. The standard error of the average is a measure of the errors to which a sample average is liable. For a sample of given size, this standard error increases as the variation between the individuals in the original population increases; for a given population, the standard error becomes smaller as the size of the sample is increased. (For the sake of the mathematically minded it may be stated that the standard error is inversely proportional to the square root of the number in the sample.) Consequently the random errors can be made as small as we please by making the sample large enough, and for a given population it is possible to calculate the size of sample necessary to reduce these errors to any desired value. Similar remarks apply to quantities other than the average.

The tendency for large samples from some population to have averages that vary little amongst themselves and differ but little from the population value is the reality behind the popular conception of the Law of Averages. This law does not operate, as some people think, so that an abnormally high individual score or run of scores is followed by an abnormally low score or run, correcting the average by compensation. In a random series, the scores following an abnormal score or run are quite unaffected by what has gone before; they tend to be nearer the general average than the abnormal scores are, i.e. to be more normal, so that when included in the average they reduce the effect of the abnormal scores. Averaging has more of a swamping than a compensating effect. Thus, if we may regard the days of weather as individuals from a population, the average weather for the population being the general type experienced at a given time of the year and place, the law of averages does not require that a very wet spell shall be followed by a very dry spell. For all I know, there may be a law to that effect, but if so, it is not the law of averages.

If the individuals in a statistical population are well mixed up, no known method of investigation can give more accurate results for a given cost than the method of purely random sampling just described, unless something is known about the individuals to enable some sort of selection to be made. Sometimes, however, a more complex form of random sample called the representative sample gives greater accuracy. Suppose, for

example, that in a housing survey we wish to find the average number of rooms per family in some town. Some families at one end of the scale of wealth will live in one room each and at the other end there may be families that have say twelve rooms each; and this variation over a range of eleven rooms per family will give rise to a certain standard error in a simple random sample. Suppose further that we can divide the town into three districts—'poor,' 'middle-class,' and 'wealthy'—in each of which the total number of families is known, and that the range of variation of rooms per family is from one to seven in the poor, from four to ten in the middle class, and from six to twelve in the wealthy district. Then if we take a random sample from any one district, the district average is estimated with a smaller standard error than that just mentioned, resulting from a range of variation of six rooms per family (i.e. 7 minus 1, 10 minus 4, or 12 minus 6). Further, it can be proved that if a *representative* sample of the same size is taken, in which the proportion of families from each district is the same as in the whole town, the standard error of the average of that sample will be the same as the smaller error resulting from a range of variation of six rooms per family. This is because the proportion of families from each district is left to chance in the simple random sample; in the representative sample it is not, and that source of error is removed.

Random sampling is the basis of the representative sample, however, which is nothing more than a weighted combination of random subsamples.

Representative sampling is used in the Gallup polls of public opinion, where care is taken to see that the opinions of various classes of people are represented in appropriate proportions instead of leaving it to chance to determine what these proportions shall be.

If it is granted that the ideal random sample can be a reliable instrument of investigation, the questions remain: Can the ideal be attained? Are the actual samples that are used as reliable as random samples? As a random sample is increased in size, it gives a result that progressively comes closer to the population value, whereas samples taken in some of the ways that are used give results that progressively come closer to some value other than the population value, results that may for some kinds of sample be too high, or for others too low. A sample of this kind is said to be *biased*, and the difference between the value given by a very large sample and the corresponding population value is called an *error of bias*. A biased die, for example, is one for which the fraction of throws showing an ace, say, tends to a value other than one-sixth (the value for the hypothetical population), and the greater the number of throws, the clearer is it that the fraction of actual aces is not one-sixth. Errors of

bias are added to the random errors, and since they follow no laws from which they can be calculated, they must be eliminated entirely or reduced so that they become unimportant. This may be difficult to do, and it is often necessary to use very elaborate sampling methods to avoid errors of bias.

It is nearly impossible for anyone to select individuals at random without some randomizing apparatus. If a teacher tries to select a few children from a class, he will tend to choose too many clever ones, or dull ones, or average ones; or if he tries to be random he may select too many clever and dull children and too few intermediate ones. In selecting a sample of houses 'at random,' the investigator will be very unlikely to select anything like the right proportions of large and small ones, shabby and smart ones, new and old ones, and so on. Bias almost inevitably will creep in. This is illustrated by the results of large experiments conducted on several thousands of school children in Lanarkshire in 1930 to measure the effect of feeding them with milk, on their growth during the period of the experiment—about six months. At each school the children were divided into two comparable groups; one group received the milk and the other did not, and the effect of the milk was to be measured by comparing the growth rates of the two groups. The results for a number of schools were combined. In an experiment of this kind, the accuracy depends very much on the two groups or samples of children being similar on the average before the feeding with milk begins, i.e. on one being unbiased with respect to the other. To secure this, the children were selected for the two groups either by ballot or by a system based on the alphabetical order of the names. Usually, these are both good ways of making unbiased random samples of the two groups, but the whole thing was spoilt by giving the teachers discretionary powers, where either method gave an undue proportion of well-fed or ill-nourished children, 'to substitute others to obtain a more level selection.' Presumably the substitution was not done on the basis of the actual weights of the children, but was left to the personal judgement of the teachers. The result was that at the start of the experiment, the children in the group that were later fed with milk were smaller than those in the other group, the average difference being an amount that represented three months of growth. It has been suggested that teachers tended, perhaps subconsciously, to allow their natural sympathies to cause them to put into the 'milk' group more of the children who looked as though they needed nourishment. This bias did not ruin the experiment, but unfortunately the interpretation of some of the results was left somewhat a matter of conjecture instead of relative certainty, and there was later a certain amount of controversy about some of the interpretations. The substitutions of

the children could have been done without introducing bias had the actual weights been made the basis, and there would have been an improvement on the purely random sampling; but by unwittingly introducing the bias, it seems that the teachers actually made matters worse.

A sampling method that is very liable to give biased results, particularly when testing opinion on controversial matters, is that of accepting voluntary returns. An undue proportion of people with strong views one way or the other are likely to make the returns, and people with moderate views are not so likely to take the trouble to represent them. For this reason, the post-bags of newspapers and Members of Parliament do not give random samples of public opinion.

A spectacular example of a biased sample is provided by the attempt of the American magazine, the *Literary Digest*, to forecast the results of the Presidential election of 1936 by means of a 'straw vote.' Some ten millions of ballot post cards were sent to people whose names were in telephone directories and lists of motor-car owners, and several million cards were returned each recording a vote for one of the candidates. Of those votes, only 40·9 per cent were in favour of President Roosevelt, whereas a few weeks later in the actual election he actually polled 60·7 per cent of the votes. Those from among telephone users and motor-car owners who returned voting cards did not provide a random sample of American public opinion on this question.

Bias does not result only from obviously bad sampling methods; it may arise in more subtle ways when a perfectly satisfactory method is modified slightly, perhaps because practical conditions make this necessary. In some Ministry of Labour samples of the unemployed, a 1 per cent sample was made by marking every hundredth name in the register of claims, which was in alphabetical order. Bias was introduced by not confining the inquiry to the marked names; instead, the first claimant appearing at the Exchange whose name was marked or was among the five names on either side of the marked one, was interviewed to provide the necessary data. Claimants who are in receipt of benefit attend at the exchange several days in a week, whereas those whose claims are disallowed but who are maintaining registration only attend once a week. The effect of this and of the latitude allowed in the choice of persons for interview was that too many claimants in receipt of benefit were included in the sample. It was only when the existence of this bias was realized that some of the results that were apparently inconsistent with other known facts made sense. A similar kind of effect can arise in surveys of households if no one is at home when the investigator first calls at some house chosen to be one of the sample. Such houses are likely to contain small families with few or no young children, since in large households

someone is almost certain to be at home to answer the door; and unless the houses with no one at home are re-visited, the sample will be biased in respect of size and character of household.

Although there is no general theory of errors of bias by which the amount of such errors can be calculated in any particular instance, as can be done for random errors, statisticians do not work entirely in the dark. Sometimes the sample gives, as part of its results, information that is also known accurately from a full census, and the sample is usually regarded as free from bias in all respects if in this one respect it agrees with the census. The soundness of the results of a sample inquiry may sometimes be checked by comparing them with data obtained in other ways, perhaps by other investigators. Where none of these checks are available, it may be necessary to rely on the statistician's general experience of sampling methods in deciding whether the sample in question is a good one. I have given enough examples to show that a good deal is known of the ways in which errors of bias arise, and what must be done to avoid them.

It is implicit in my definition of errors of bias that they cannot be 'drowned' by taking very large samples, in the way that random errors can; a fact that the experience of the *Literary Digest's* straw-vote on the American Presidential election of 1936 amply confirms. From this point of view, a good sample can be arrived at only by employing a good sampling method. I have already mentioned some methods incidentally, and it is only necessary here to give it as a warning that when a statistician advises adherence to an elaborate method with a closeness that may seem to the layman to be 'fussy,' that advice had better be followed; failure to do so has been known to lead to biased results.

Altogether, the method of inquiry by sample is difficult and full of pitfalls. But statisticians could not get on without it and experience of its use is both wide and deep, so that in competent hands the method is capable of giving results that are reasonably accurate. Moreover, the inevitable errors in the results can be estimated, and allowance can be made for them in arriving at conclusions.

TAKING ACCOUNT OF CHANCE

Chance operates in many fields besides that of random sampling, and many of its effects can be calculated by applying the same general methods as are used to calculate the errors of random sampling. Some of the further applications of those methods will be described in the present chapter.

The effects of chance can be calculated only because they follow certain laws, but these differ in kind from the *exact* laws of subjects like physics.

Events that follow exact laws can be described or predicted precisely; but we can only specify probabilities that chance events will occur, or specify limits within which chance variations will probable lie. Newton's laws of motion, for example, are exact because they describe exactly the relations between the motions of bodies and the forces acting upon them; the errors of random sampling follow chance laws because we cannot predict exactly what average a random sample will have; we can only state, as I have suggested on p. 1462, the probability that it will lie within certain limits.

I cannot embark upon a full discussion of what we mean by chance, but as a preliminary I shall indicate a few ideas associated with the word. Statisticians attribute to chance, phenomena (events or variations) that are not exactly determined, or do not follow patterns described by known exact laws, or are not the effects of known causes. That is to say, the domain of chance varies with our state of knowledge—or rather of ignorance. Such ignorance may be fundamental because the relevant exact laws or causes are unknowable; it may be non-essential or temporary, and exist because the exact laws do not happen to have been discovered; or the ignorance may be deliberately assumed because the known exact laws and causes are not of such a character that they can profitably be used in the particular inquiry in hand.

An example of ignorance that, according to present-day ideas, is fundamental, is in the Principle of Indeterminacy of modern physics; we do not and cannot know the precise motion of an electron. We do not know what determines the position of a shot on a target, but that ignorance is non-essential and in some degree temporary. The variations in the positions of the shots depend on a host of factors such as variations in the primary aim of the marksman, the steadiness of his hand, the weight, size, and shape of the bullets, the propelling charges, the force of the wind, and so on; but presumably these factors can be investigated and laws be discovered. Indeed, this has happened; and the history of gunnery shows the temporary character of the ignorance. Gunnery is much more of a science and more exact than it was in the days of the Battle of Waterloo, or even during the 1914–18 War; and as knowledge has increased, unpredictable variations in placing of shots have been reduced; but at each stage these variations are regarded as due to chance. Ignorance of causes is assumed by an insurance company in using its past experience of accident claims to establish future premiums for motor-car insurance. The company has considerable knowledge of the circumstances surrounding every accident on which a claim is made, but is unable to make more than limited use of that knowledge, and so treats accidents largely as chance events, except for a few special allowances such as 'no claims bonuses' or extra premiums charged to people with bad accident records.

Usually, events regarded as coming within the domain of chance are those governed by a complicated system of many causes, each of which produces only a small variation; and one frequent characteristic of such events is that small changes in the circumstances surrounding them make a big difference to the results.

Chance as I have described it operates in a very wide field, covering the whole of the unknown; but mathematical calculations can be made and chance laws be propounded only for comparatively simple systems covering a portion of this field. Nevertheless, such calculations have a wide range of usefulness, which the following examples will illustrate.

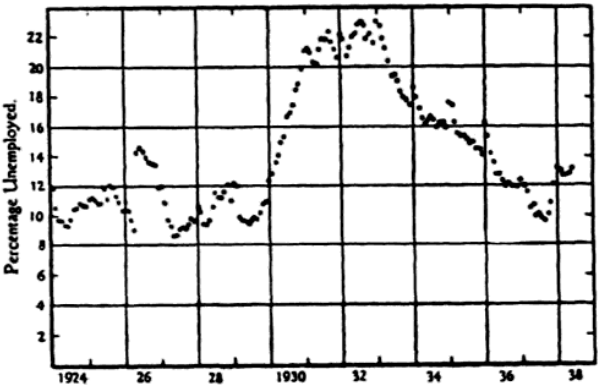

FIGURE 2—Percentage of insured workers unemployed, Great Britain 1924–38.

One use of chance calculations is for deciding which of the fluctuations in a time series are random and which are trends having some significance. As an example of a time series, consider the unemployment data represented in Figure 2. Readers will have no difficulty in recognizing the broad changes, viz. the minor waves in 1925 and towards the end of 1928, the large upward sweep in 1930, the improvement from the end of 1932 to the end of 1937, and the upward movement again in 1938. For the time being we shall omit 1926, the year of the General Strike, as being exceptional. These changes are reasonably attributed to fundamental causes that operate fairly slowly and may be represented by a smooth curve drawn through the actual points of the graph. There are a number of mathematically determined curves that have the property of changing in level in such a slow, regular way, and are of the nature of exact laws or descriptions.

Let us imagine such a curve to be drawn through the points of Figure 2. Then the actual points will be seen to deviate from this curve. They may in some degree follow a seasonal pattern (another exact law), but in Figure 2 that pattern is not very evident, and most of the deviations are

irregular. Presumably many of them can be explained in terms of a minor strike in some industry, a political change in some country affecting another industry, an exceptionally hard winter, and so on; but we cannot bring such knowledge into a system, and so we assume ignorance. Hence, having tried all the known kinds of exact laws that are relevant, we consider the deviations to be due to a complex system of chance causes that operate we know not how; and we apply to them the same laws as describe the results of games of chance and sampling experiments. This is the argument for applying the theory of errors of random sampling to testing the statistical significance of fluctuations in time series. For example, there was a sudden and temporary rise in unemployment in the beginning of 1936; is it significant? It actually occurred, and therefore is real, but when we ask the above question we in effect ask: Can the rise be reasonably considered as a random fluctuation arising from that system of causes we have labelled *chance*, or has some unusual event happened? And so we apply the theory of random errors. If this theory had been applied to testing the statistical significance of the sudden rise in unemployment in 1926 it would have shown the operation of something unusual—we know that to be the General Strike.

This kind of application of the theory enables us, in retrospect, to decide whether any particular events with which we try to associate fluctuations have had important effects compared with the system of random fluctuations. When following changes week by week or month by month as they occur it is useful, too, to be able to decide whether the last increase or decrease is large enough to call for action, or whether it is random. At one period, for example, a local newspaper used to publish weekly figures of deaths due to road accidents in a certain town, and the number used to fluctuate about an average of four or five per week. Should we worry if between two particular weeks the number rises from three to six, or rejoice if it falls from five to two? No! Such changes are no greater than any that can be attributed to chance, and do not indicate a real change in conditions. Sometimes the chance coincidence of random fluctuations may give rise to several consecutive small increases or decreases, giving a spurious appearance of a trend. Sampling theory can show when such is the case.

To arrive at results of these kinds, it is necessary to analyse the time series so as to separate the random fluctuations from the secular movements; and additional complications occur if the system of random fluctuations changes. Some would say, for example, that in trade the random fluctuations during a slump and a boom are different. The whole analysis is only approximate, but it is based on ideas that are sufficiently close to reality to give useful results.

The theory of random errors was used for measuring the accuracy of

astronomical measurements long before it was applied to statistical samples, and it is used somewhat in measuring the errors of experimental observations in general. When the astronomer measures, say, the position of a star, he finds that in spite of the precision of his apparatus, and the care with which he adjusts it and makes his observations, he does not get the same answer from successive determinations. He repudiates the idea that the position is varying and attributes the variations in his results to unavoidable errors of observation. The question arises: What is the true position? And if it cannot be measured exactly, how accurately can it be estimated? A similar situation arises in the other so-called exact sciences: e.g., in physics and chemistry. Several determinations have been made of the velocity of light, but they do not agree exactly; and a chemist would be very surprised if he got exactly the same result every time he measured an atomic weight.

This interpretation of experimental results as being due to an invariable quantity plus observational or experimental errors is purely a mental conception. The only reality is the set of observations, the characteristics of which can, if desired, be expressed by any statistical constants such as the average, or a measure of variation, or by a frequency distribution. For most experiments, however, it is useful and (within limits) valid to adopt the more common conception.

The errors do not follow any known exact laws, and so the laws of chance are sometimes used to describe them. In applying these laws, the results are regarded as a random sample from a hypothetical population of results, the average of this population being the true value. Then, the average of the sample is an estimate of the true value, and the error in that estimate can be calculated as for any statistical sample. Is this idea valid? On the face of things, it seems as reasonable to imagine the millions of results that would have been obtained had the experiment been repeated millions of times under the same conditions as to imagine the results of millions of tosses of a die. But it is not so certain that the variations between experimental results are entirely of the same kind as those we get when we toss dice.

On this question there are differences of opinion among experimentalists. Some refuse to admit any similarities between experimental and random errors. Others, faced with otherwise intractable results, use the theory of random errors as the only way out. Experimental errors are not, in general, random. There are 'personal' factors, and any one person shows a bias that changes from time to time. I prefer to regard a set of experimental results as a biased sample from a population, the extent of the bias varying from one kind of experiment and method of observation to another, from one experimenter to another, and, for any one experi-

menter, from time to time. If this view is accepted, experimental errors can be regarded as forming a chance system, but the system is not as simple as that assumed in calculating the errors of random sampling.

In general the bias cannot be estimated and the theory of random errors is therefore not enough. Sometimes, however, one can say that the bias is likely to be small compared with the random errors, and then the theory may give useful, if approximate, results. For example, if, say, five separate chemists were to determine the atomic weight of an element independently, in different times and places, and possibly by different methods, the results would vary because of the effects of random errors and bias. But the separate biases for the five chemists would differ and so would appear as the random errors between the results, the group as a whole would probably exhibit but little bias, and the theory of errors would provide a reasonably close measure of the precision with which the average of the five results estimates the true atomic weight. This might not be so, on the other hand, for the average of, say, twenty consecutive determinations made by one chemist in one laboratory.

Errors of bias are often relatively unimportant when the observed quantity is the difference between two similar quantities. In measuring the distance between two lines in a spectrum, for example, the main error is often due to the uncertainty of setting the cross-hairs of the measuring microscope on the centres of the lines. If there is a bias in doing this, it is likely to be similar for the two lines (provided they are not too dissimilar in width and appearance), and the difference in the two settings will probably be practically unbiased. The theory of errors gave a result that was at least qualitatively right, when applied to Lord Rayleigh's measurements of the density of nitrogen. He made a number of determinations on 'atmospheric' and 'chemical' nitrogen and found a difference in the two averages. Subsequent treatment by the theory of errors has shown that the difference is greater than can be attributed to random variations, and this result is in accordance with a real difference we now know to exist, owing to the presence of the rarer inert gases in 'atmospheric' nitrogen.

Where the bias is completely unknown, I doubt if it is possible to do more than hope that the true value lies somewhere between the highest and lowest of the actual values, and regard the average as an estimate of the true value, that is as good as, but no better than, any other single estimate that could be made from the data. It is, of course, the experimenter's job to reduce bias and random errors to a minimum.

To sum up, the theory of random errors may be usefully applied to some experimental observations, particularly of differences in values, but great caution must be observed on account of bias. Certainly such an application is no substitute for careful experimental control.

Much experimental work, particularly in biological subjects, is now done under conditions, many of which can be well controlled, and the observations can be made accurately; but the material is inherently variable and the results have to be treated statistically. The Lanarkshire experiment made to measure the effect of milk on the growth of children, already mentioned on p. 1466, is of this kind. The amount of milk fed can be controlled, children fed and not fed with milk can be kept in the same environment, and the changes in weight can be measured accurately; but it would not do to base conclusions on an experiment on, say, two children. Children vary, and it is necessary to observe a large number and take averages.

The problem of interpreting the results of such experiments is essentially statistical, and it has fallen to the lot of statisticians to study the general questions of arranging experiments with variable material, of drawing conclusions from the results, and of testing them. Under the leadership of Professor R. A. Fisher, who started this work at the Rothamsted Experimental Station (for agriculture), an elaborate technique for doing this has been developed and is very widely used. I propose to give some description of this subject.

There are three main principles to be observed in designing such an experiment; they are replication, randomization, and economy in arrangement.

The necessity for replication has already been stated. The problem first arose chiefly in agricultural field trials made to measure such things as the effects of various fertilizers on wheat yield. It was early seen that different plots treated in the same way gave different yields. Hence, it was not sufficient to have two plots, say, to treat one with a fertilizer, to grow the crops and measure the yields, and to regard the difference as measuring the effect of the fertilizer. The experiment had to be replicated by treating several plots in each way and measuring the difference between the average yields.

Even differences in such averages can be affected by variations between plots, as we can see from the results of the sampling experiment described in the last chapter; and it is desirable to estimate the accuracy of the observed difference. The only known way of doing this is by the theory of random errors. It was found, however, that variations in plot fertility were not random. There was usually a fertility pattern, e.g. a gradient in fertility across the field. In order that the theory of sampling could be applied, an element of randomization was introduced artificially by using some such device as a ballot to decide which plots should receive the various experiment treatments. This is a 'trick of the trade' for making fertility variations into a comparatively simple chance system. A statis-

tician might apply this principle to the above-mentioned experiment of feeding milk to school children by tossing a coin once for each child, giving that child milk if the result is 'heads,' say, and no milk if the result is 'tails.'

The pattern in fertility differences between plots in a field was used to increase the accuracy of experimental comparisons. Adjacent plots tend to be more alike than those in different parts of the field, and by comparing the treatments on adjacent plots the random variations affecting the comparison were reduced, with an increase in accuracy. The other way of increasing accuracy is to increase the number of plots, and hence the expense of the experiments; the arrangement using adjacent plots is therefore more economical. In the same way, had it been possible in the Lanarkshire milk experiment to use identical twins, giving milk to one of each pair, far fewer children would have given the same accuracy as thousands chosen at random. This kind of arrangement can be made to satisfy the condition of randomness sufficiently for the application of the theory of random errors in an appropriate form.

The above are the elementary principles of the modern approach to the design of what I shall term *statistical-experimental* investigations. The whole subject has, however, become very complicated as several treatments of one kind have been included, and treatments of several kinds. Thus, experiments may be done with various quantities and combinations of several kinds of fertilizer on several varieties of wheat. Further complication arises when experiments are done on different farms and in different years, and it is necessary to consider to what extent results obtained on one group of farms in one year apply to other farms and other years.

In spite of the fact that sound methods are available, experimenters continue to work with variable material on non-statistical lines, and they get discordant results which they cannot fit into a system. Different workers sometimes get different results in the same subject, and controversies arise. When, in such circumstances, the experimenters turn to sound methods of statistical analysis, involving proper experimental arrangements, difficulties of these kinds tend to disappear. Then, experiments which were previously done on an inadequate scale are increased in size, often they are designed more economically than before, and the advancement of knowledge is made more orderly and certain.

Statistical methods are often regarded as applying only to very large numbers of observations, but that is no longer true. It would be far too costly to replicate some experiments hundreds and thousands of times, and statisticians have had to make do with small numbers. They have, however, developed the theory of errors to apply to small samples as well as to large ones.

There are many chance events that occur in life, to which the general theory of random errors may in some degree be applied.

For example, many telephone subscribers have access to one trunk line, and a multitude of causes determine how many will want to use it at any given instant, i.e. it is to some extent a question of chance whether more than one subscriber will want to use it at once and thus cause delay. In so far as this is true, the extent of delays of this kind can be calculated from the theory of probability which is the basis of the theory of errors. This is typical of a number of congestion problems that arise in telephony, in road and rail traffic, and so on; and although many of them are difficult mathematically, the theory is being applied.

Accidents have a large element of chance in their causation—the circumstances preceding a 'near shave' often differ by only a hairbreadth from those preceding a catastrophically fatal accident, and the theory of probability has been useful for studying accident problems in calculating the effects of chance and showing the importance of other factors. The following is an example.

Records were kept of the numbers of accidents that happened during the course of one year to 247 men workers engaged in moulding chocolate in a factory. Some of the men had no accident, some had one, some two, and so on, a few having as many as twenty-one accidents. The data are arranged in a frequency distribution in the first two columns of Table 2. Now we ask: Were all the variations between the men in the numbers of accidents they suffered due to chance, or were there differences between the men in their tendency to have accidents? Were the 42 men who had no accidents exceptionally skilful or just lucky; and were the 22 men who had ten accidents or more clumsy or unlucky? The average number of accidents per man is 3·94, and even if all the men were equally skilful in avoiding accidents, chance would give rise to some variation. It has been calculated from the extended theory of random sampling that this variation would result in the frequency distribution of the last column of figures of Table 2. This is very different from the actual distribution. We may say, roughly, that 5 of the 42 men with no accidents were lucky and the remaining 37 skilful; that one of the 22 men with ten or more accidents was unlucky and the remainder clumsy. Comparisons of this kind between actual and calculated chance distributions have led to investigations that have shown how people differ in 'accident proneness,' i.e. in their tendency under given circumstances to suffer accidents. The chance distribution given in Table 2 is calculated by assuming a very simple system of chance variations; more complicated systems taking into account variations in accident proneness have been used in the more advanced investigations on the subject.

TABLE 2

FREQUENCY DISTRIBUTION OF MEN WHO HAD VARIOUS NUMBERS OF ACCIDENTS. COMPARISON BETWEEN ACTUAL AND CHANCE DISTRIBUTIONS

(Data by E. M. Newbold. Report No. 34, *Industrial Fatigue (Now Health) Research Board*)

Number of Accidents	*Frequency of Men*	
	Actual	*Chance*
0	42	5
1	44	21
2	30	40
3	30	50
4	25	48
5	11	37
6	12	23
7	15	13
8	8	6
9	8	3
10–15	19	1
15–21	3	–
Total	247	247

STATISTICAL LAWS

The central problem of statistics is dealing with groups variously described as collections, crowds, aggregates, masses, or populations, rather than with individual or discrete entities; with events that happen on the average or in the long run rather than with those that happen on particular occasions; with the general rather than with the particular. A fuller consideration of this aspect of statistics is the subject of the present chapter.

Again I shall use the language common in statistical writings and refer to *populations* of *individuals*. The population is regarded (in the discussion of sampling) as something from which samples are taken, but here as an aggregate of individuals, which will in most instances be represented by a sample, i.e. I shall not distinguish between the population and the sample.

The population has characteristics and properties of its own, which are essentially derived from and are an aggregate of those of the individuals, although the two sets of properties may be different in kind. In the population, the individuals merge and their individuality is dissolved, but from the dissolution rises a new entity like a phoenix from the flames. The population is at the same time less and more than the totality of the individuals.

This conception is not peculiar to statistics. Rousseau, for example, distinguishes in *The Social Contract* between the General Will and the wills of all the people:

'In fact, each individual, as a man, may have a particular will contrary or dissimilar to the general will which he has as a citizen. His particular interest may speak to him quite differently from the common interest.'

'There is often a great deal of difference between the will of all and the general will; the latter considers only the common interest, while the former takes private interest into account, and is no more than a sum of particular wills: but take away from these same wills the pluses and minuses that cancel one another, and the general will remains as the sum of the differences.'

The general idea is expressed in another way in the following passage from *Old Junk* by Mr. H. M. Tomlinson:

'His shop had its native smell. It was of coffee, spices, rock-oil, cheese, bundles of wood, biscuits and jute bags, and yet was none of these things, for their separate essences were so blended by old association that they made one indivisible smell, peculiar, but not unpleasant, when you were used to it.'

The loss of individuality results from the method of the statistician in confining his attention to only a few characteristics of the individuals and grouping them into classes. Consider a married couple, say Mr. and Mrs. Tom Jones. As a couple their individuality consists of a unique combination of a multitude of characteristics. Mr. Jones is tall and thin, is aged 52 years, has brown hair turning grey, and is a farmer. Mrs. Jones is called Mary and at 38 years is still handsome; she is blonde and is really a little too 'flighty' for a farmer's wife. The couple have been married for 16 years and have three children: two boys aged 14½ and 11 years, and a girl aged 2. In addition to these and similar attributes the couple have a number of moral and spiritual qualities that we may or may not be able to put down on paper. It is by all these, and a host of other qualities that their relatives and neighbours know Mr. and Mrs. Jones; the uniqueness of the combination of qualities is the individuality of the couple.

The statistician who is investigating, say, the ages of husbands and wives in England and Wales is interested only in the ages, and does not wish to describe even these accurately. So he puts our couple in that class for which the age of the husband is 45–55 years and that of the wife is 35–45 years. Mr. and Mrs. Jones are now merely one of a group of some 320,000 other couples, and are indistinguishable from the others in their group.

Statistical investigations are not always confined to one or two characters of the individuals, and elaborate methods have been developed for

dealing with many attributes, e.g. the ages of married couples at marriage, income, number of children, fertility of the grandparents of the children, and so on, but however many attributes are included, they are very few compared with the number that make up the individuality of each couple.

A population of individuals is the most characteristic and simplest chance system the statistician has to deal with. We do not know, or do not take any account of, the causes of the differences between the individuals, and so we dismiss them as being due to chance, and fasten our attention on the population.

Statistics is essentially totalitarian because it is not concerned with individual values of even the few characters measured, but only with classes. However much we analyse the data to show the variation between the parts, we still deal with sub-groups and sub-averages; we never get back to the individuals. In studying the death rate of a country, for example, we may decompose the general average into sub-averages for the two sexes, for the separate age-groups, for different localities, industries, and social classes; but the death rate of an individual has no meaning. When we think of variation, we think of a mass of variable individuals rather than of one or two being very different from the remainder.

This part of statistical technique in selecting only a few characteristics for investigation, and in classifying the data, is not only necessary because of the limited power of the human brain to apprehend detail, but is a part of the general scientific method. It is an essential step in the development of general scientific laws. However much we know of Mr. and Mrs. Jones in particular, if we know nothing more we have no basis for drawing conclusions about married couples in general. It is only by paying attention to such features as individuals have in common with others that we can generalize. Individuals are important, as such, to themselves, to their neighbours and relations, and to professional consultants—the parson, the doctor, and the lawyer; they have no importance for the statistician, nor indeed for any scientist, except that they, with a host of other individuals, provide data.

Our first and, for most of us, our only reactions to our environment are individualistic. We *are* individuals, our experience is mostly with individuals, and even when considering a group we are conscious mostly of our personal relationship to it. The concept of the population as an entity does not come easily, and our ordinary education does little to correct this defect. The mental effort required to realize this concept is perhaps something like that necessary to appreciate a fugue with its contrapuntal pattern, as compared with the ease of following a tune with simple harmonies.

The characteristics of the population are described by frequencies and

by the statistical constants and averages already described, but it is apparently so difficult to think of the reality behind these constants—the mass of individuals—that we personify the population and speak in such terms as 'the average man.' This is only possible because of a similarity between some of the measures of a population and those of an individual; the average height of a group of men is expressed in feet and inches, just as the height of one man is; but the similarity is only superficial.

We have already seen the inadequacy of the average as a description of variable material,[2] but the average individual sometimes is also a rather absurd figure. In 1938, for example, he was among those comparatively rare individuals who died at the age of 58 years. His age in England and Wales in 1921 was 29·9, and in 1938 it was 33·6 years; i.e., in 17 years the average man aged by only 3·7 years! The average family can have fractions of a person. Books on the upbringing of babies usually contain a curve showing the growth in weight of an average baby; but few actual curves are like that. The curve for a real baby may be above or below that for the average and it may have a different slope in various parts. It will also usually have 'kinks' due to teething troubles and minor illnesses, whereas the curve for the average baby is fairly smooth; this paragon among children has no troubles!

Variation is, of course, an important characteristic of populations that individuals cannot have. I have already been at pains to describe this,[3] and to point out how, for example, the deviations from any relationship shown by a contingency or correlation table are as characteristic of the data as the relationship itself. Indeed, without variation, a collection of individuals is scarcely a population in the statistical sense. A thousand exactly similar steel bearing balls (if such were possible) would be no more than one ball multiplied one thousand times. It is the quality of variation that makes it difficult at first to carry in mind a population in its complexity.

All the special properties of populations I have considered arise in aggregates of independent individuals, but there are additional characteristics due to interactions between individuals. The behaviour of men in the mass is often different from their behaviour as individuals. Some men affect (or 'infect') others and such phenomena as mass enthusiasms and panics arise. We speak of mass-psychology. Similarly the effect of an infectious disease on a community of people in close contact is different from its effect on a number of more or less isolated individuals. Statistical description can take account of interactions between individuals, but it is seldom necessary to do so.

[2] Discussed in an earlier chapter of Tippett's book. Ed.
[3] Discussed in an earlier chapter of Tippett's book. Ed.

Although the individuals in a population vary, the characteristics of the population itself are very stable. Sir Arthur Eddington has well said: 'Human life is proverbially uncertain; few things are more certain than the solvency of a life-insurance company.' This means that we do not know when any individual will die, but an insurance company can estimate the incidence of death in its population of policy-holders with great accuracy.

This contrast between individualistic variability and statistical stability, and the fact that the latter emerges from the former, this apparent paradox of order coming out of chaos, has from time to time given rise to metaphysical speculations. People in the eighteenth century, accustomed to considering the variations between individuals, seem to have been struck by the statistical regularities and saw evidences of a Divine order. Sir Arthur Eddington, on the other hand, presumably taking for granted the regularity of the laws of physics, is more struck by the compatibility with these laws of the unpredictable variation in the behaviour of individual electrons, and offers comfort to those who want to believe in free will and scientific law at the same time. The practical statistician may accept it as a fact requiring no special metaphysical explanation, that mass regularities can often be discerned where the individuals apparently follow no regular laws.

Galton writes of the regularity of form of the frequency distribution in the following terms:

'I know of scarcely anything so apt to impress the imagination as the wonderful form of cosmic order expressed by the "Law of Frequency of Error." The law would have been personified by the Greeks and deified, if they had known of it. It reigns with serenity and in complete self-effacement, amidst the wildest confusion. The huger the mob, and the greater the apparent anarchy, the more perfect is its sway. It is the supreme law of Unreason. Whenever a large sample of chaotic elements are taken in hand and marshalled in the order of their magnitude, an unsuspected and most beautiful form of regularity proves to have been latent all along.'

Let us re-examine the data from the sampling experiment described on pp. 1460–1462 and see if we can repeat Galton's experience and recapture something of his mood.

I have extended the experiment to obtain 4,000 scores altogether. The first thirty are given in the top part of Table 1 (p. 1461) in the order in which they occurred, and these together with the 3,970 other scores are the 'large sample of chaotic elements'—and chaotic they undoubtedly

appear. I then proceeded to marshal the scores in the order of their magnitude by forming a frequency distribution, and stage by stage stopped to look at the result as the distribution began to grow. The results for 50, 200, 1,000, and 4,000 scores are in Figure 3. Since the scores are whole numbers, I have not grouped them into sub-ranges; the scales of the distributions in the vertical direction have been reduced as the numbers of scores have increased. At 50 scores, there is no sign of any regularity or form in the distribution, but at 200 scores, a vague suggestion of a form seems to be emerging; the scores show a slight tendency to pile up in

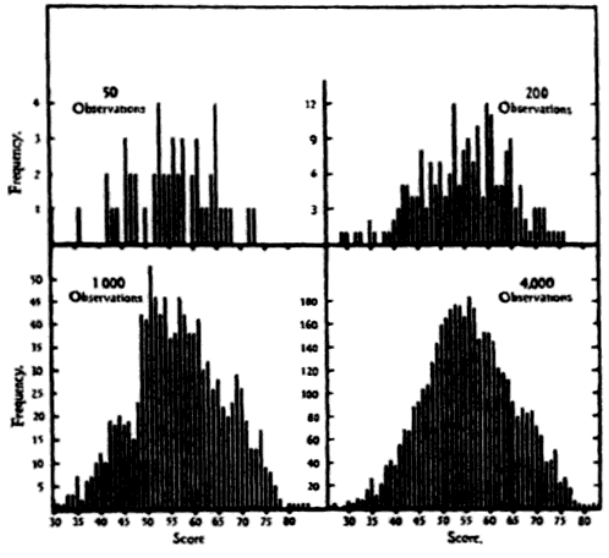

FIGURE 3—Frequency distribution of various numbers of observations from the same population.

the middle of the range. At 1,000 scores, the form is clearly apparent, although irregularities are still pronounced; but at 4,000 scores, the 'most beautiful form of regularity' is there, almost in perfection. It is not difficult to imagine the regularity that would be apparent were the sample so large as to be indistinguishable from the population.

The formulae and laws that describe populations and their behaviour as opposed to individuals are termed *statistical laws*. The various statistical constants (e.g. standard deviation, mean deviation, measures of association and correlation) are elementary statistical laws. Other laws of a higher order of complexity describe how populations change with time or place, or other circumstances. Laws of heredity, for example, are a way of describing how some characters in populations of plants or animals change from generation to generation.

Some statistical laws are discovered by simple observation of the population as a whole. For example, the change in the death rate for the

country may be recorded from year to year, without any consideration being given to the changes in the chances of death from various causes, to which the individual is exposed. A public lighting authority could compare two batches of electric lamps by counting how many of each are burnt out after having been in use for, say, 500 hours. Or a colony of the banana fly may be kept in a bottle under standard conditions, and the growth in numbers observed. However, there is nothing necessarily statistical in the technique applied in such experiments, although investigations of this character are often classed as statistical in the widest sense of the word. The introduction of the concept of pieces of matter as populations of electrons or atoms does not necessarily turn an ordinary physical investigation of the macroscopic properties of matter into a statistical one.

Statistical methods and calculations are involved, however, when the laws for the population are deduced from those for individuals. The calculation of statistical constants is a case in point, and the estimation of some quality of a batch of electric lamps from calculations made on the full frequency distribution of lives is another. Estimates, made by demographers, of the size and age composition of the future population from a consideration of the characteristics of the present population and the various birth and death rates, are an important example of the statistical deduction of statistical laws. Such calculations may involve complicated mathematics.

It is implicit in all I have written that statistical laws have nothing to do with individuals. It is no exception to the statistical law that old men have old wives, on the average, if one old man of one's acquaintance has a young wife. A failure to recognize the distinction between the two types of laws sometimes leads to attempts to apply statistical laws to individuals, with paradoxical results.

We now return to the starting-point of this chapter—a consideration of individuals. They in the aggregate are the population, and from their characteristics we can calculate those of the population. We cannot perform the reverse process. Individuality is lost, as far as the statistician is concerned, for good and all. Does this mean we know absolutely nothing of the individual when we know the population? Not quite.

Consider a single electric lamp taken at random from the batch represented by the distribution of Table 3. Even if we do not know its life, we know that it will be an exceptional lamp if its life is greater than, say, 2,800 hours—it will be one of $\frac{1}{150}$ths of the batch. Indeed, it is more likely to be one of the $\frac{89}{150}$ths of the lamps with lives between, say, 1,000 and 2,000 hours.

We are used, in ordinary life, to dealing with data of this kind by

TABLE 3

LENGTH OF LIFE OF ELECTRIC LAMPS

(Data by E. S. Pearson, *Journal of the Royal Statistical Society*, **96**, 1933, p. 21)

Life (hours)	*Frequency of Lamps*
0– 200	1
200– 400	3
400– 600	2
600– 800	10
800–1,000	19
1,000–1,200	21
1,200–1,400	23
1,400–1,600	18
1,600–1,800	17
1,800–2,000	10
2,000–2,200	8
2,200–2,400	5
2,400–2,600	5
2,600–2,800	4
2,800–3,000	2
3,000–3,200	1
3,200–3,400	1
Total	150

introducing the concept of probability. In the example quoted we would say that the probability of any one lamp having a life greater than 2,800 hours is $\frac{4}{150} = 0{\cdot}027$, and that the probability of the life being between 1,000 and 2,000 hours is $\frac{89}{150} = 0{\cdot}593$.

This is an application of what is commonly regarded as the statistician's definition of probability as a ratio of frequencies. Corresponding to any frequency distribution there can be calculated a whole series of probabilities of a random individual lying within various stated limits, and statistical probability is a device (a verbal trick!) for attaching to the random individual the characteristics of the whole distribution. In this way, a population is epitomized in an individual much more satisfactorily than in the concept of 'the average man.' But statistical probability does more than this. It corresponds closely to the more popular idea of probability as a measure of the strength of belief in a thing. Most people if asked what is the probability of a tossed penny falling heads uppermost would reflect that heads was as likely as tails and would reply: one-half. The statistician, if in a pedantic mood, would reply: in the hypothetical population of tosses, one-half of the total give heads, therefore the probability of a head is one-half. An alternative method of expression is to state that the chances of a head are even, or for the lamps, that they are 593 to 407 in favour of a life of between 1,000 and 2,000 hours.

Probability is, in ordinary life, also applied to events that do not occur as frequencies. We speak of the probability of, or the chances in favour of, a particular horse winning a race. Even in such instances, however, I think that people carry at the backs of their minds the idea of frequencies; they in effect imagine a lot of races, in a given proportion of which the particular horse wins. The idea is described in the following quotation from a lecture given by Karl Pearson in 1892:

'A friend is leaving us, say in Chancery Lane at 4 o'clock in the afternoon, and we tell him that he will find a Hansom cab at the Fleet Street corner. There is no hesitation in our assertion. We speak with knowledge, because an invariable experience has shown us Hansom cabs at 4 o'clock in Fleet Street. But given the like conditions within reach of a suburban cab-stand, and our statement becomes less definite. We hesitate to say absolutely that there will be a cab: "You are sure to find a cab," "I believe there will be a cab on the stand," "There is likely to be a cab on the stand," "There will possibly be a cab on the stand," "There might *perhaps* be a cab," "I don't expect there will be a cab," "It's very improbable," "You are sure not to find a cab," etc., etc. In each and every case we go through some rough kind of statistics, *once* we remember to have seen the stand without a cab; on occasions few and far between, "perhaps on an average once a month," "perhaps once a week," "every other day," "more often than not there has been no cab there." Certainty in the case of Fleet Street passes through every phase of belief to disbelief in the case of the suburban cab-stand. If once a month is the very maximum of times I have seen an empty cab-stand, my belief that my friend will find a cab there to-day is far stronger than if I have seen it vacant once a week. A measure of my belief in the occurrence of some event in the future is thus based upon my statistical experience of its occurrence or failure in the past.'

Thus probability in its most general use is a measure of our degree of confidence that a thing will happen. If the probability is 1·0, we know the thing will certainly happen, and if the probability is high, say 0·9, we feel that the event is likely to happen. A probability of 0·5 denotes that the event is as likely to happen as not, and one of zero means that it certainly will not. This interpretation, applied to statistical probabilities calculated from frequencies, is the only way of expressing what we know of the individual from our knowledge of the population.

Statistical laws, which describe the characters and behaviour of populations in one way or another, may be transformed into probabilities—i.e. from them the probabilities and frequencies in the population may be calculated. Thus, statistical laws are the chance laws referred to earlier.

It may have been noticed that probabilities have been calculated from the frequencies of a distribution, either known as for the lamps, or as-

sumed as for the penny. In general, it is necessary to have some data on which to calculate probabilities. I am often asked what is the probability of some queer or interesting event, without being given any data. Statisticians do not evolve probabilities out of their inner consciousness, they merely calculate them.

"Let us sit on this log at the roadside, says I, and forget the inhumanity and the ribaldry of the poets. It is in the glorious columns of ascertained facts and legalized measures that beauty is to be found. In this very log that we sit upon, Mrs. Sampson, says I, is statistics more wonderful than any poem. The ring shows it was sixty years old. At the depth of two thousand feet it would become coal in three thousand years. The deepest coal mine in the world is at Killingworth, near Newcastle. A box four feet long, three feet wide, and two feet eight inches deep will hold one ton of coal. If an artery is cut, compress it above the wound. A man's leg contains thirty bones. The Tower of London was burned in 1841."

"Go on, Mr. Pratt," says Mrs. Sampson. "Them ideas is so original and soothing. I think statistics are just as lovely as they can be."

—O. HENRY (*The Handbook of Hymen*)

When Tennyson wrote The Vision of Sin, Babbage read it. After doing so, it is said he wrote the following extraordinary letter to the poet:

"In your otherwise beautiful poem, there is a verse which reads:

'Every moment dies a man,
Every moment one is born.'

"It must be manifest that, were this true, the population of the world would be at a standstill. In truth the rate of birth is slightly in excess of that of death. I would suggest that in the next edition of your poem you have it read:

'Every moment dies a man,
Every moment 1 1/16 is born.'

"Strictly speaking this is not correct. The actual figure is a decimal so long that I cannot get it in the line, but I believe 1 1/16 will be sufficiently accurate for poetry. I am etc." —MATHEMATICAL GAZETTE

5 On the Average and Scatter

By M. J. MORONEY

ON THE AVERAGE

'The figure of 2·2 children per adult female was felt to be in some respects absurd, and a Royal Commission suggested that the middle classes be paid money to increase the average to a rounder and more convenient number.' (*Punch*)

IN former times, when the hazards of sea voyages were much more serious than they are today, when ships buffeted by storms threw a portion of their cargo overboard, it was recognized that those whose goods were sacrificed had a claim in equity to indemnification at the expense of those whose goods were safely delivered. The value of the lost goods was paid for by agreement between all those whose merchandise had been in the same ship. This sea damage to cargo in transit was known as 'havaria' and the word came naturally to be applied to the compensation money which each individual was called upon to pay. From this Latin word

derives our modern word *average*. Thus the idea of an average has its roots in primitive insurance. Quite naturally, with the growth of shipping, insurance was put on a firmer footing whereby the risk was shared, not simply by those whose goods were at risk on a particular voyage, but by large groups of traders. Eventually the carrying of such risks developed into a separate skilled and profit-making profession. This entailed the payment to the underwriter of a sum of money which bore a recognizable relation to the risk involved.

The idea of an average is common property. However scanty our knowledge of arithmetic, we are all at home with the idea of goal averages, batting and bowling averages, and the like. We realize that the purpose of an average is *to represent a group of individual values* in a simple and concise manner so that the mind can get a quick understanding of the general size of the individuals in the group, undistracted by fortuitous and irrelevant variations. It is of the utmost importance to appreciate this fact that the average is to act as a *representative*. It follows that it is the acme of nonsense to go through all the rigmarole of the arithmetic to calculate the average of a set of figures which do not in some real sense constitute a single family. Suppose a prosperous medical man earning £3,000 a year had a wife and two children none of whom were gainfully employed and that the doctor had in his household a maid to whom he paid £150 a year and that there was a jobbing gardener who received £40 a year. We can go through all the processes of calculating the average income for this little group. Six people between them earn £3,190 in the year. Dividing the total earnings by the number of people we may determine the average earnings of the group to be £531 13s. 4d. But this figure is no more than an impostor in the robes of an average. It represents not a single person in the group. It gives the reader a totally meaningless figure, because he cannot make one single reliable deduction from it. This is an extreme example, but mock averages are calculated with great abandon. Few people ask themselves: What conclusions will be drawn from this average that I am about to calculate? Will it create a false impression?

The idea of an average is so handy that it is not surprising that several kinds of average have been invented so that as wide a field as possible may be covered with the minimum of misrepresentation. We have a choice of averages; and we pick out the one which is appropriate both to our data and our purpose. We should not let ourselves fall into the error that because the idea of an average is easy to grasp there is no more to be said on the subject. Averages can be very misleading.

The simplest average is that which will be well known to every reader. This common or garden average is also called the *mean*, a word meaning 'centre.' (All averages are known to statisticians as 'measures of central

tendency,' for they tell us the point about which the several different values cluster.) The *arithmetic mean* or average of a set of numbers is calculated by totalling the items in the set and dividing the total by the number of individuals in the set. No more need be said on this point, save that the items to be averaged must be of the same genus. We cannot, for example, average the wages of a polygamist with the number of his wives.

A second kind of average is the *harmonic mean*, which is the reciprocal [1] of the arithmetic mean of the reciprocals of the values we wish to average. The harmonic mean is the appropriate average to use when we are dealing with rates and prices. Consider the well-known academic example of the aeroplane which flies round a square whose side is 100 miles long, taking the first side at 100 m.p.h., the second side at 200 m.p.h., the third side at 300 m.p.h., and the fourth side at 400 m.p.h. What is the average speed of the plane in its flight around the square? If we average the speeds using the arithmetic average in the ordinary way, we get:

$$\text{Average speed} = \frac{100 + 200 + 300 + 400}{4} = 250 \text{ m.p.h.}$$

But this is not the correct result as may easily be seen as follows.

Time to travel along the first side	= 1 hour
Time to travel along the second side	= 30 minutes
Time to travel along the third side	= 20 minutes
Time to travel along the fourth side	= 15 minutes
Hence total time to travel 400 miles	= 2 hours 5 minutes
	= $2\frac{5}{12}$ hours

From this it appears that the average velocity is $400 \div 2\frac{5}{12} = 192$ m.p.h.

The ordinary arithmetic average, then, gives us the wrong result. A clue as to the reason for this will be found in the fact that the different speeds are not all maintained for the same time—only for the same distance. The correct average to employ in such a case is the harmonic mean.

In order to give the formula for this we shall here introduce a little more mathematical notation which will be of great benefit to us later in this book. In calculating averages we have to *add up* a string of items which make up the set whose average is required. The mathematician uses a shorthand sign to tell us when to add up. He calls adding up 'summing' and uses the Greek letter *S* which is written Σ and called 'sigma' to indicate when terms are to be added. (This is actually the capital sigma. Later we shall have a lot to say about the small letter sigma which is written σ.) Each of the numbers which have to be taken into account in our calcula-

[1] The reciprocal of a number is found by dividing that number into unity, e.g., the reciprocal of $4 = \frac{1}{4} = 0{\cdot}25$.

tion is denoted by the letter x. If we wish to differentiate between the various quantities we can number them thus: x_1, x_2 x_3, x_4, etc., the labelling numbers being written as subscripts so that they will not be confused with actual numbers entering into the calculation. (This may sound as confusing to the novice as it will be boring to the learned. Let the learned turn over the pages till they find something interesting, while we explain this simple and useful shorthand to the novice.) Let us take as an example the calculation of the arithmetic average of the five numbers 5, 6, 8, 7, 6. We could, if there were any reason for keeping track of these, label them as follows:

$$x_1 = 5 \qquad x_2 = 6 \qquad x_3 = 8 \qquad x_4 = 7 \qquad x_5 = 6$$

Now the advantage of using algebraic notation (i.e., letters to stand for any numbers we care to substitute for them according to the problem in hand) is that we can write down in a very compact way the rules for performing the calculation which will give us the correct answer to the type of problem we are dealing with. In fact, a formula is nothing else than the *answer to every problem* of the type to which it applies. We solve the problem once and for all when we work out a formula. The formula *is* the answer. All we have to do is to substitute for the letters the actual quantities they stand for in the given problem. Suppose, now, we denote the number of quantities which are to be averaged in our problem by the letter n (in our case here, $n = 5$). To calculate the arithmetic average we have to add up all the five quantities thus: $5 + 6 + 8 + 7 + 6 = 32$. This adding part of the calculation would appear in algebraic form as $x_1 + x_2 + x_3 + x_4 + x_5$. The next step would be to divide the total by the number of items to be averaged, viz. 5, giving the result 6·4 for the average. In algebraic notation this would appear as

$$\text{Average} = \frac{x_1 + x_2 + x_3 + x_4 + x_5}{n}$$

This method of writing the formula would be very inconvenient if there were a large number of items to be averaged; moreover, there is no need to keep the individual items labelled, for in an average the identity of the individuals is deliberately thrown away as irrelevant. So we introduce the summation sign, Σ, and write our formula in the very compact form:

$$\text{Average} = \frac{\Sigma x}{n}$$

The formula thus tells us that to get the average we 'add up all the x values and divide their total by the number of items, n.'

In similar fashion, now, the *harmonic mean*, which we have said is the average to be used in averaging speeds and so on and which is defined

as the reciprocal (the reciprocal of a number x is equal to $\frac{1}{x}$) of the arithmetic mean of the reciprocals of the values, x, which we wish to average, has the formula:

$$\text{Harmonic mean} = \frac{n}{\sum\left(\frac{1}{x}\right)}$$

To illustrate the use of this formula let us use it on our aeroplane problem. The four speeds, which were each maintained over the same distance, were 100, 200, 300, and 400 m.p.h. These are our x values. Since there are four of them the value of n in our formula is 4, and we get:

$$\text{Harmonic mean} = \frac{n}{\sum\left(\frac{1}{x}\right)} = \frac{4}{(1/100 + 1/200 + 1/300 + 1/400)} = \frac{4}{(25/1200)}$$

$$= \frac{4 \times 1200}{25} = 192 \text{ m.p.h.}$$

which we know to be the correct answer.

The reader should note carefully that the harmonic mean is here appropriate because the times were variable, with the distances constant. Had it been that times were constant and distances variable the ordinary arithmatic average would have been the correct one to use. The type of average which is appropriate always depends on the terms of the problem in hand. Formulae should never be applied indiscriminately.

Yet a third type of average is the *geometric mean*. This is the appropriate average to use when we wish to average quantities which are drawn from a situation in which they follow what W. W. Sawyer in *Mathematician's Delight* calls the 'gangster law of growth,' i.e., a geometric progression or the exponential law. Many quantities follow this type of law. For example, the population of a city, given a stable birth-rate and death-rate with no migration, will increase at a rate proportional to the number of people in the city. Suppose that in the year 1940 a certain city had a population of 250,000 and that in the year 1950 its population were 490,000. If we wished to estimate the population in the year 1945 (estimating populations at various times between successive censuses is an important matter in public health statistics) then we might, as a rough approximation, take the average of the populations at the two known dates, thus:

$$\text{Population at } 1945 = \frac{250{,}000 + 490{,}000}{2} = 370{,}000$$

This would only be a sensible method if we were able to assume that the population increased by the same number every year. This is not likely, however, for, as the city grows in size, so the number of citizens is likely to grow at an ever increasing rate (see Figure 1). A better estimate is likely to be obtained, in normal circumstances, by calculating the *geometric mean* of the population at the two known dates. To calculate the geometric mean, we multiply together all the quantities which it is desired to average. Then, if there are n such quantities, we find the nth root of the product. Denoting our n quantities by $x_1, x_2, x_3, \ldots x_n$, we may write the formula for the geometric mean as follows:

$$\text{Geometric mean} = \sqrt[n]{x_1 \times x_2 \times x_3 \times \ldots x_n}$$

Applying this to the problem given above where we wish to estimate the population of a city in 1945, given that in 1940 the population was 250,000 and in 1950 was 490,000, we have $n = 2$ items to average, and we find:

$$\text{Geometric mean} = \sqrt[2]{250{,}000 \times 490{,}000} = 350{,}000$$

as our estimate for the population at 1945. This result, it will be noted, is appreciably lower than we obtained using the arithmetic average (370,000). If the reader considers Figure 1 he will see that it is the more likely estimate.

Collecting together, at this point, our three different averages, we have:

Arithmetic Mean (usually denoted as $\bar{x}$ and called x-bar)

$$\bar{x} = \frac{\Sigma x}{n}$$

Harmonic Mean (usually denoted by H)

$$H = \frac{n}{\sum\left(\frac{1}{x}\right)}$$

Geometric Mean (usually denoted by G)

$$G = \sqrt[n]{x_1 \times x_2 \times x_3 \times \ldots x_n}$$

Each of these measures of central tendency has its own special applications. All of them are obtained by simple arithmetical processes which take into account the magnitude of every individual item.

We emphasized the important idea of any average or measure of central tendency as the *representative* of a homogeneous group in which the

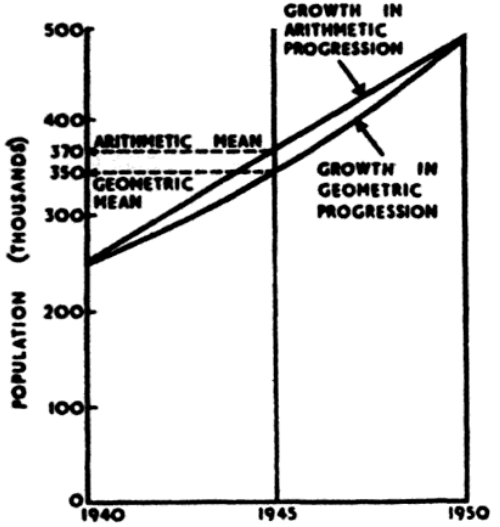

FIGURE 1—Comparison of Interpolation by Arithmetic Mean and Geometric Mean. The population of a city often grows according to the exponential law. This would certainly be true with stable birth-rate and death-rate and in absence of migration. Under these conditions, the geometric average would be more appropriate than the arithmetic average to interpolate the population at a given date between two dates at which the population was known.

members are recognizably similar. Now many distributions, while being undoubtedly homogeneous in the sense that there is continuity between the various members of the group, nevertheless are such that very great differences exist between the largest and smallest members, and, moreover, exhibit a marked lack of symmetry, the family tending to cluster much nearer to one extreme than the other. Figure 2 is a typical example. It shows the way in which annual income is distributed. There is certainly continuity, but small incomes are the norm. The reader will appreciate at once that to calculate averages for distributions of this type using the arithmetic mean would be very misleading. The relatively few people with extremely high incomes would pull up the average appreciably, so that it could not be taken as truly representative of the population in general. Figure 3, which shows the relative frequency of different sizes of family, presents the same difficulty. Some families are *very* well off for children and the calculation of an arithmetic average might well be misleading—particularly if our purpose is purely descriptive.

It is evident that what we need in such cases is a measure of central tendency which is unaffected by the relatively few extreme values in the 'tail' of the distribution. Two ideas suggest themselves. The first is that if

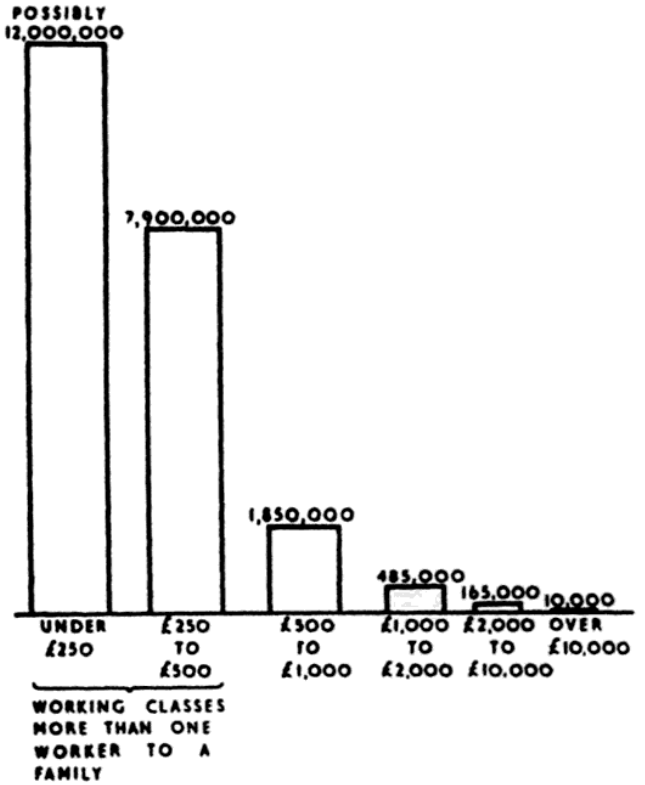

FIGURE 2—Numbers of people in different income ranges forming a positively skew distribution.

we were to take all our families and set them down in a long column starting with the smallest and working up to the largest, we could then use the size of that family which came halfway down the column as our measure of central tendency. This measure is called the *median* (meaning 'middle item'). Half of all families would have a size not less than that of the median family and half not more than that of the median family. Notice that in this way we do not take account at all of the actual numbers of children except for ranking purposes. It is evident that the number of children in the largest family could be increased to 50,000 without in any way disturbing our measure of central tendency, which would still be the middle item.

A second method of getting a measure of central tendency which is not upset by extreme values in the distribution is to use the *most commonly occurring value*. This is the fashionable value, the value *à la mode*, so to say. It is called the *mode* or *modal value*. For example, in Figure 3 the modal value for the size of family is seen to be two children. This is really

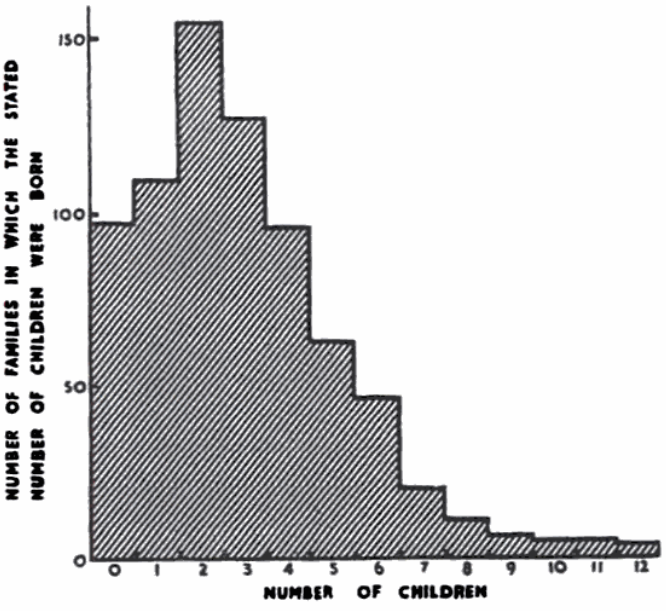

FIGURE 3—Distribution of number of children per family is also positively skewed.

a typical value and seems real to us compared with the arithmetic average which in this case works out to 2·96. It is difficult to imagine 2·96 children. Notice that the arithmetic mean is markedly affected by the relatively few very large families. Which is correct? Neither and both. Both averages serve a purpose. The mode would form a very poor basis for any further calculations of an arithmetical nature, for it has deliberately excluded arithmetical precision in the interests of presenting a typical result. The arithmetic average, on the other hand, excellent as it is for numerical purposes, has sacrificed its desire to be typical in favour of numerical accuracy. In such a case it is often desirable to quote *both* measures of central tendency. Better still, go further and present a histogram of the distribution as in Figure 3.

A problem which not infrequently arises is to make an estimate of the median value of a distribution when we do not have the actual values of each individual item given, but only the numbers of items in specified ranges.

We shall now say a few words about frequency distributions. If we have a large group of items each of which has connected with it some numerical value indicative of its magnitude, which varies as between one member of the group and another (as, for example, when we consider the heights of men or the amount of income tax paid by them), and if

we draw up a table or graph showing the relative frequency with which members of the group have the various possible values of the variable quantity (e.g., proportion of men at each different height, or proportions

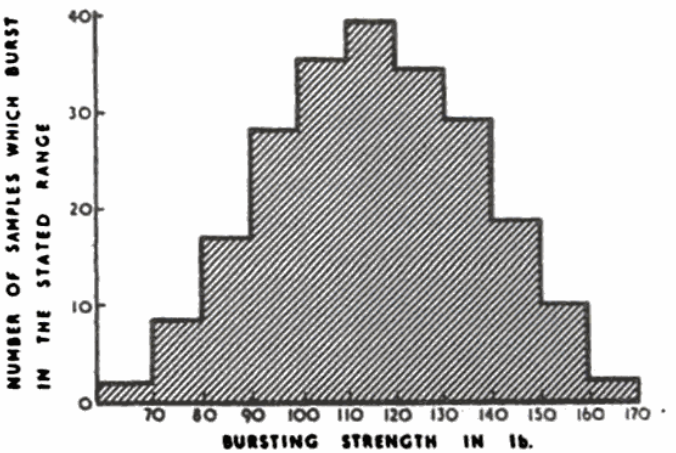

FIGURE 4—Distribution for bursting strength of samples of vinyl coated nylon exhibiting symmetry.

of the population, falling into various income tax groups), then we have what is called a *frequency distribution* for the variable quantity in question. This is usually called simply the *distribution*. Thus we have distributions for height, weight, chest size, income, living rooms per person, and so on. Similarly we have distributions for the number of deaths according to age for different diseases, number of local government areas with specified birthrates and deathrates and so on. The quantity which varies

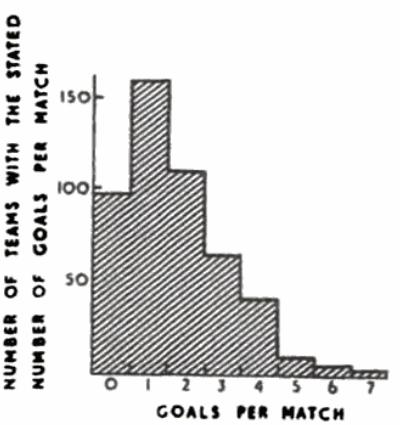

FIGURE 5—The number of goals scored per team per match gives a positively skewed distribution of a discontinuous variable.

(height, birthrate, income, and so on) is called the *variate*. Some variates are *continuous*, i.e., they can assume *any* value at all within a certain range. Income, height, birth-rate, and similar variates are continuous.

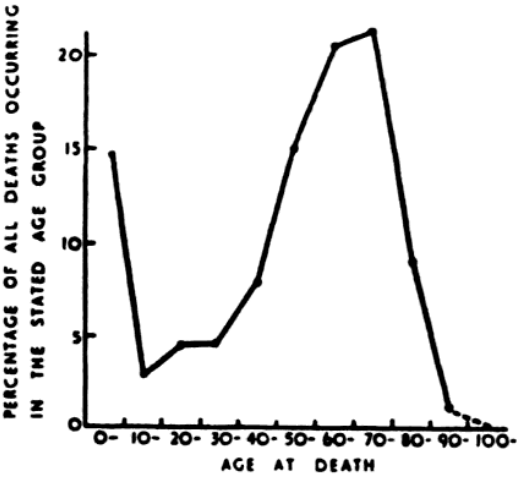

FIGURE 6—Example of a Bimodal (double peaked) Distribution. The peak in the first years of life reflects the serious loss of potential life due to the infantile mortality rate. (From the "Registrar General's Report, Years 1930–32," quoted by M. G. Kendall in "Advanced Statistics.")

Other variates are said to be *discontinuous*, because they can only assume isolated values. For example, the number of children in a family can only be a whole number, fractions being impossible. Families grow in distinct jumps. An addition to the family is an event. Goals scored in football matches, articles lost in buses, the number of petals on a flower—all such variable quantities are discontinuous.

When we collect together information for the purposes of statistical analysis it is rare that we have information about all the individuals in a group. Census data are perhaps the nearest to perfection in this sense; but even in this case the information is already getting out of date as it is collected. We may say that the census count in a certain country taken on a certain day came to 43,574,205, but it would be nothing short of silly to keep quoting the last little figure 5 for the next ten years—or even the next ten minutes. Such accuracy would be spurious. In general it is not possible to investigate the whole of a population. We have to be content with a *sample*. We take a sample with the idea of making inferences from it about the population from which it was drawn, believing, for example, that the average of a good sample is closely related to the average of the whole population. The word *population* is used in statistics to refer not simply to groups of people, but, by a natural extension, to groups of

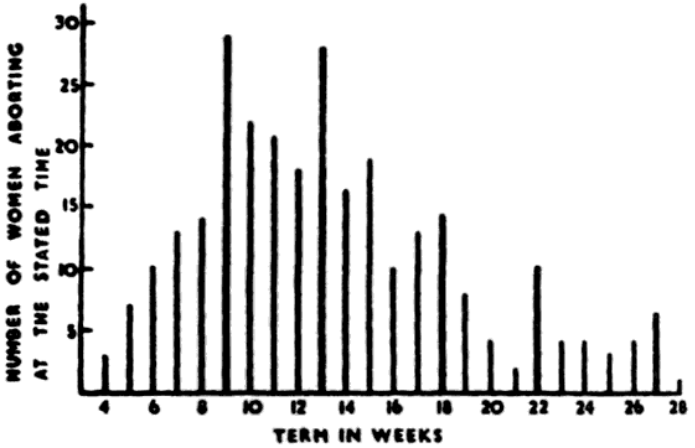

FIGURE 7—Abortion in women. Data given by T. V. Pearce (1930) and quoted by M. G. Kendall, "Advanced Statistics." The reader may care to speculate about possible periodicity in these data. Is there reasonable suggestion of a cycle whose duration is roughly one month? What other conclusion can you draw?

measurements associated with any collection of inanimate objects. By drawing a sufficiently large sample of measurements, we may arrive at a frequency distribution for any population. Figures 4–8 give examples of various types of distribution.

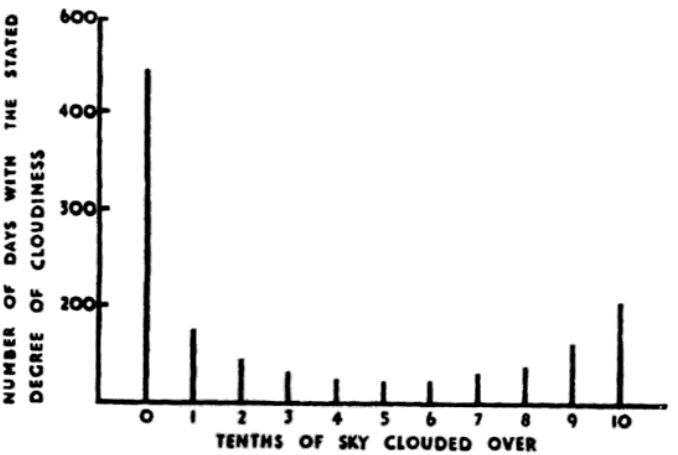

FIGURE 8—Peculiar distribution of cloudiness at Greenwich. Based on data given by Gertrude Pearse (1928) for month of July 1890–1904 (excluding 1901) and quoted by M. G. Kendall, "Advanced Statistics," Vol. I. Note tendency for sky to be either very clear or very cloudy.

Some distributions, as will be seen from the diagrams, are *symmetrical* about their central value. Other distributions have marked asymmetry and are said to be *skew*. Skew distributions are divided into two types. If the 'tail' of the distribution reaches out into the larger values of the variate, the distribution is said to show *positive skewness*; if the tail extends towards the smaller values of the variate, the distribution is called *negatively skew*. In the next chapter we shall take up the question of the concentration of the members of the distribution about their central value, for it is clearly a matter of the greatest importance to be able to measure the degree to which the various members of a population may differ from each other.

Figure 9 illustrates an interesting relationship which is found to hold approximately between the median, mode, and mean of moderately skew distributions. Figures 10 and 11 illustrate geometrical interpretations of the three measures of central tendency.

We shall close this chapter with an elementary account of Index Numbers, which are really nothing more than a special kind of average. The

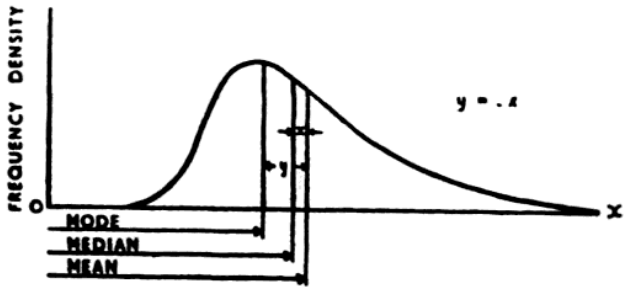

FIGURE 9—Mean, Median and Mode in moderately skew cases. For moderately skew distributions we have the simple approximate relation: Mean − Mode = 3 (Mean − Median). For a perfectly symmetrical distribution they all coincide.

best known index number is the Cost of Living Index, which, as readers will know, is a rough measure of the average price of the basic necessities of life. In many industries, the Cost of Living Index is a strong chain which keeps a man's reward tied strictly to his necessity rather than to his ambition. But index numbers are a widespread disease of modern life, or, we might better say, a symptom of the modern disease of constantly trying to keep a close check on everything. We have index numbers for exports, for imports, for wage changes and for consumption. We have others for wholesale and retail prices. The Board of Trade has an index. The Ministry of Labour has an index. *The Economist* has another. It is scarcely possible to be respectable nowadays unless one owns at least one

index number. It is a corporate way of 'keeping up with the Joneses'—the private individual having been forced by taxation to give up this inspiring aim long ago.

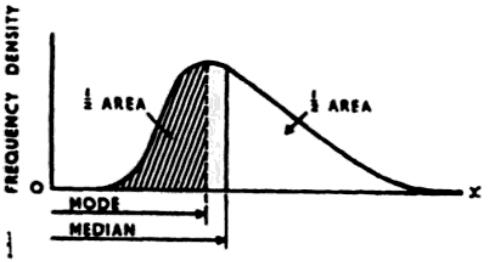

FIGURE 10—Geometrical interpretation of Mode and Median. The vertical line at the median value divides the area under the frequency curve into halves (area is proportional to frequency). The vertical line at the modal value passes through the peak of the curve, i.e., it is the value at which the frequency density is a maximum.

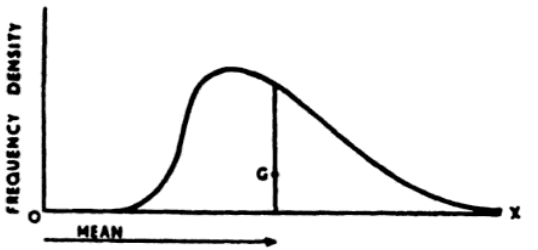

FIGURE 11—Geometrical interpretation of the Mean The vertical line at the mean will pass through the centre of gravity of a sheet of uniform thickness and density cut to the shape of the distribution. The mean is the abscissa of the centre of gravity "G."

It is really questionable—though bordering on heresy to put the question—whether we would be any the worse off if the whole bag of tricks were scrapped. So many of these index numbers are so ancient and so out of date, so out of touch with reality, so completely devoid of practical value when they have been computed, that their regular calculation must be regarded as a widespread compulsion neurosis. Only lunatics and public servants with no other choice go on doing silly things and liking it. Yet, since we become more and more the servants of our servants, and since they persist in tying us down to this lugubrious system whereby the housewife, the business man, and the most excellent groups of the citizenry have all their difficulties compressed into the brevity of an index number, we reluctantly ask the reader to bear with us while we explain, briefly, this academic tomfoolery of telling us in cryptic form

what we knew already from hard experience: namely, that the cost of living has risen in the last couple of months, sufficiently for us to be able to submit a humble claim for higher wages to offset part of our increased burden.

Consider the question of the changes which take place in retail prices. As every housewife knows, the price we are asked to pay bears only the faintest resemblance in many cases to the worth of the article. She knows, too, that for many commodities it is more accurate to speak of prices rather than price. Tomatoes in one shop may be 6d. per pound; the same tomatoes in another shop may be 10d. or 1s. Some people are well enough off to be able to shop by price. They like lots of service and servility and are willing to pay for it. Yet, even if these sections of the community are excluded, there still remains a fair variation between one district and another for the same article, things like fish and fruit being notorious in this respect. In addition to this variation in the price of the articles, we have to recognize that different families have different spending patterns. If cheese were made as dear as gold it would not matter one iota to the family that hates cheese like poison. Conscientious vegetarians would probably regard it as an excellent thing if the price of meat rose to prohibitive levels. Total abstainers positively loathe the idea of beer and spirits being cheap. Non-smokers love to see the Chancellor raise the money by piling the tax on 'non-essentials' like tobacco. It is evident that we shall get nowhere if all this individuality is to run riot. It is far too inconvenient for the statistician.

We get over the difficulty by shutting our eyes to it. All we have to do is to invent a 'standard family.' [2] We might, for example, choose the standard urban working-class family. We then do a sample survey, to find out what quantities of the various articles we are considering they consume in a week under normal conditions, and draw up a table as follows:

EXPENDITURE OF THE STANDARD WORKING-CLASS FAMILY (1949)

	Quantity	Price	Expenditure	Weight
Bread and Flour	39 lb.	4d./lb.	156d.	31·2
Meat	7 lb.	24d./lb.	168d.	33·6
Potatoes	35 lb.	2d./lb.	70d.	14·0
Tea	1 lb.	36d./lb.	36d.	7·2
Sugar	2 lb.	5d./lb.	10d.	2·0
Butter	1 lb.	18d./lb.	18d.	3·6
Margarine	1 lb.	12d./lb.	12d.	2·4
Eggs	1 doz.	30d./doz.	30d.	6·0
		Total	500d.	100·0

[2] Composed of one underpaid male, one overworked female, and 2·2 underfed children.

Now, it is a relatively simple matter to keep track of the changes in prices as time goes on. It would be very much more troublesome to keep a check on whether the spending pattern, as indicated by the amounts of the various items bought by the standard family, was tending to change. One line of approach would be to assume that our standard family will not change its demands from year to year. Suppose for the year 1950 the prices were as in the following table.

EXPENDITURE OF THE STANDARD WORKING-CLASS FAMILY (1950)

	Quantity	Price	Expenditure	Weight
Bread and Flour	39 lb.	5d./lb.	194d.	30·1
Meat	7 lb.	30d./lb.	210d.	32·6
Potatoes	35 lb.	3d./lb.	105d.	16·3
Tea	1 lb.	36d./lb.	36d.	5·6
Sugar	2 lb.	6d./lb.	12d.	1·9
Butter	1 lb.	27d./lb.	27d.	4·2
Margarine	1 lb.	15d./lb.	15d.	2·3
Eggs	1 doz.	45d./doz.	45d.	7·0
		Total	644d.	100·0

The reader should ignore, for the moment, the last column, headed 'Weight,' in each table. The obvious thing, at once, is that to buy the same quantities of the same articles, and therefore to get the same 'satisfaction,' as the economists have it, cost the standard family 644d. in 1950 as against 500d. in 1949, i.e., the cost in 1950 as compared with 1949 was $^{644}/_{500} \times 100 = 128{\cdot}8\%$. We could then say that the index of retail prices, as represented by this group of items, stood at 129 in 1950 (1949 = 100).

We could get a similar indication of the rise in retail prices as follows. Consider, first, the amount of money our standard family spent on the various items in our 'base year, 1949.' These can be reduced to percentages of the total expenditure (on the group of items considered in the index). For instance, out of a total expenditure of 500d., bread and flour claimed 156d. or 31·2%. Similarly, meat took 33·6% of the total expenditure, potatoes 14·0%, and so on. These figures are entered in the column headed 'Weight' since they tell us the relative importance of the different items in the household budget. Meat is a very heavy item, sugar a relatively small one. These weights give us a pattern of expenditure as it actually appeared to the standard housewife in the base year. They take account of both quantity and price. The first thing that is obvious from this pattern of weights is that, while a 50% increase in the cost of sugar is not a matter of great importance to the housewife, even a 10% increase in the price of meat would be a serious extra burden to carry in the

standard family where income is usually closely matched to expenditure. We must remember that our standard family is a standardized family. Its wants are not supposed to change. It is supposed to be devoid of ambition. It only gets a rise in salary when such a rise is absolutely necessary.

Now while it is true (in the absence of subsidies and purchase tax or price fixing by combines) that all commodities tend to rise in price together, nevertheless, superimposed on this general tendency, there will be a certain irregularity. Comparing the price of bread and flour in our two years we find that the 'price relative,' as it is called, of this item is $\frac{5}{4} \times 100 = 125\%$ in 1950 as compared with the base year, 1949.

The following table shows the 'prices relative' for the several items, together with the weights corresponding to the base year. The weights have been quoted to the first decimal place, further places being condemned as coming under the heading 'delusions of accuracy.'

	Price relative	Base year weight	Price-rel. × weight
Bread and Flour	125	31·2	3,900
Meat	125	33·6	4,200
Potatoes	150	14·0	2,100
Tea	100	7·2	720
Sugar	120	2·0	240
Butter	150	3·6	540
Margarine	125	2·4	300
Eggs	150	6·0	900
	Total	100·0	12,900

If, now, we divide the total of the 'prices relative × weight' by the total of the weights, we get the average prices of the commodities in 1950, as compared with the base year, 1949, equals 129·00, which we certainly quote no more accurately than 129. This would now be our index of retail prices. For every hundred pennies spent in 1949 we need to spend 129 in 1950 to get the same amount of 'satisfaction.' Evidently, every succeeding year—or month, for that matter—can be compared with our base year.

The economists, of course, have great fun—and show remarkable skill—in inventing more refined index numbers. Sometimes they use geometric averages instead of arithmetic averages (the advantage here being that the geometric average is less upset by extreme oscillations in individual items), sometimes they use the harmonic average. But these are all refinements of the basic idea of the index number which we have indicated in this chapter. Most business men seem to thrive without understanding this simple matter. Perhaps they half realize that it doesn't mean a lot, except in regard to wage negotiations between themselves and Trade Unions—and in such cases experts on both sides of the fence do all the statistics

required. The employer and employee don't much mind how much of this arithmetic goes on, so long as the final agreement is reasonably fair to both sides.

The snags in this index number game will be apparent to the reader. First of all, if he will inspect the pattern of weights in the tables for 1949 and 1950, he will see that they are not identical. Over a reasonable period of years the pattern can change appreciably. Then, again, if we try to measure the cost of living of our standard family by including heating, lighting, rent, beer, cigarettes, football pools, and the rest, we soon get into deep water. For example, if we find that in the base year the standard family spends one-tenth of its income on football pools, are we to argue that since this is a heavy item of expenditure it shall be supported somehow in the cost of living calculations? Until very recently the cost of living index in this country took account of the cost of paraffin and candles for lighting purposes, and assumed that no working-class family had heard of electricity. Then there is the difficulty that the standard family tends to become a standardized family in so far as its wages are tied to an index which is slow to recognize the right of its standard family to be anything but standard in its requirements from year to year. The reader should consider carefully the full implication of 'subsidies on essentials' (included in cost of living index) and 'purchase tax on non-essentials' (not included in the index or only modestly represented). The pernicious nature of tying wages to cost of living indexes while this jiggery-pokery is official policy will be apparent. The whole scheme is positively Machiavellian in its acceptance of deception as a necessity in politics. And does it really work so well, after all? The truth is that it is too inefficient even to keep the worker standardized. As new items are available from manufacturers, the public has to be given the power to purchase them, whether they are included in the cost of living index or not. Shall we ask the economists: What good do your indexes do—really?

SCATTER

'The words figure and fictitious both derive from the same Latin root, fingere. Beware!'—M. J. M.

We have discussed various ways of measuring the central tendency of distributions and have seen that such measures are characteristic of the distribution of any quantity, so that different populations are distinguished from each other by different values of these measures. For example, the average value for the height of women differs from the average height for men. Numerical characteristics of populations are called *parameters*. Having dealt with parameters of central tendency, we now turn to the no less

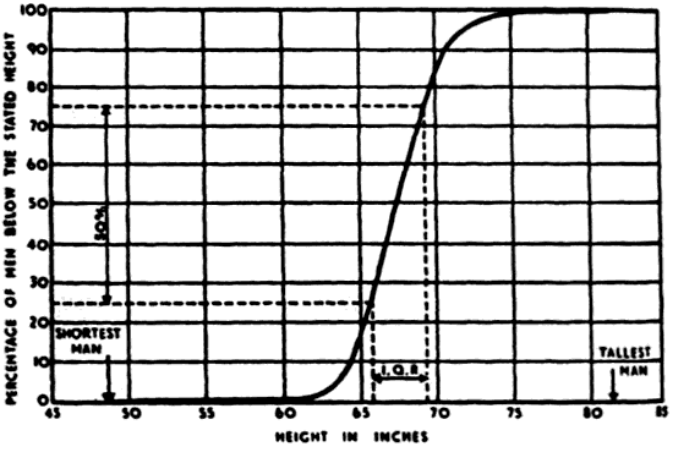

FIGURE 12—Ogive for heights for young men (I.Q.R. = interquartile range). Based on W. T. Martin, "Physique of the Young Male," by permission of H.M.S.O.)

important matter of parameters of dispersion. According to Memorandum No. 20 issued by the Medical Research Council (W. J. Martin: *The Physique of Young Males*) the height of young males, aged between 20 and 21 years, has an average value of 5 feet 7½ inches. This is information. But we should like to know more,[3] for it is evident that not all the young men were exactly of this height. The adjoining ogive (Figure 12) shows the percentages of men less than stated heights in a total of 91,163 who were measured. Figure 13 shows the data displayed in histogram form. It is evident that very considerable variability exists, so that, whilst the great majority of men differ relatively little from the average height, very noticeable departures from it are not at all infrequent. How are we to get a measure of the variability about the mean value?

The easiest way is to state the height of the tallest man seen and the shortest, thus. Tallest: 6 feet 9 inches. Average: 5 feet 7½ inches. Shortest: 4 feet 0 inches. Alternatively, we might state the *range*, i.e., the difference between the tallest and the shortest, viz. 6 feet 9 inches *minus* 4 feet 0 inches = 2 feet 9 inches. This is not a very good way. A moment's thought will make it clear that we might very easily not have met these two extreme heights. It might well have been that we should have found the shortest man to be 4 feet 4 inches and the tallest 6 feet 6 inches. This would give us a range of 6 feet 6 inches *minus* 4 feet 4 inches = 2 feet 2 inches—a result which is appreciably different from the previous one. Again, it might have happened that among those examined in this group for military service were the giant and the dwarf from some circus.

[3] The author does not disappoint us in this desire.

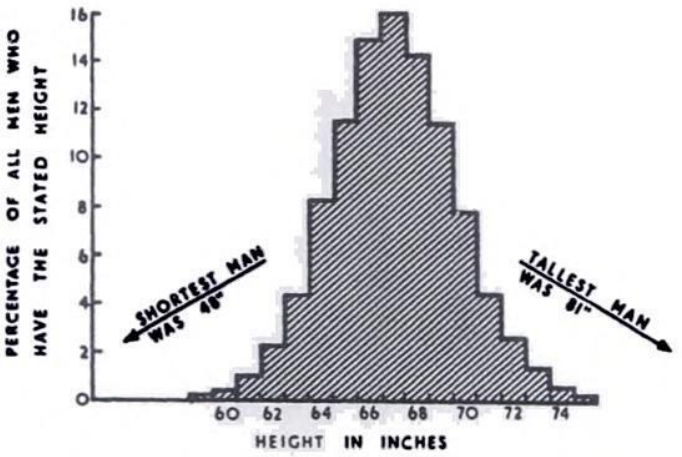

FIGURE 13—Histogram corresponding to the ogive of Figure 12.

Supposing the giant to be 9 feet 7 inches and the dwarf 3 feet 2 inches, we should have obtained for our range the value 6 feet 5 inches. It is obviously undesirable to have a measure which will depend entirely on the value of any freaks that may occur. It is impossible for a measure based on freaks to speak as the *representative* of the ordinary population. The range, then, although it is used in certain circumstances, is not ideal as a measure of dispersion.[4] It would be better to have a parameter less likely to be upset by extreme values.

We may tackle this problem by devising a measure for dispersion along the same line that we took for the median when we were discussing measures of central tendency. The median was the value above which 50% of the population fell and below which the other 50% fell. Suppose, now, we divide the population, after it has been set out in order of size, into *four equal groups.* The value above which only 25% of the population falls we call the *upper quartile,* and the value below which only 25% of the population falls we call the *lower quartile.* Evidently, 50% of the population falls between the upper and lower quartile values. The reader may care to check for himself that the upper and lower quartiles, for the table of heights we are using as an example, are roughly 5 feet 9 inches and 5 feet 6 inches respectively. Thus, we may see at once that roughly 50% of the population differ in height by amounts not exceeding three inches, despite the fact that the tallest man observed was no less than 2 feet 9 inches taller than the shortest man. This, of course, is a consequence of the way in which the large majority of heights cluster closely to the average. This is a very common effect. Intelligence Quotients be-

[4] The range is very efficient when the samples contain very few items.

have in the same sort of way. Most people are little removed from average intelligence, but geniuses and morons tend to occur in splendid isolation. (We may recall here that the modal ('fashionable') value tends to coincide with the arithmetic mean when the distribution is fairly symmetrical.) Thus the *interquartile range*, i.e., the difference between the upper and lower quartile values, makes a good measure of dispersion. It is immune from the disturbances occasioned by the incidence of extreme values. It is easy to calculate. It has a simple and meaningful significance in that it tells us the range of variability which is sufficient to contain 50% of the population. The interquartile range is frequently used in economic and commercial statistics for another reason. Often, data are collected in such a way that there are indeterminate ranges at one or both

FIGURE 14—Showing numbers of firms with the stated number of employees in the food, drink, and tobacco trades of Great Britain. (Based on Census of Production 1930, quoted by M. G. Kendall, "Advanced Statistics," Vol. I.)

ends of the table. An example is shown in Figure 14. The largest group is labelled '400 and over.' This is vague, and it would obviously be impossible to do a precise calculation for any measure depending on arithmetical processes involving the actual values in the unbounded upper class. (We shall show in the next chapter how the limited vagueness in the other *bounded* classes is dealt with.) The median and the interquartile range provide us with measures of central tendency and scatter respectively in such cases.

Median and quartiles are simply special cases of a quite general scheme for dividing up a distribution by *quantiles*. Thus, we may arrange our distribution in order of size and split it up into *ten* groups containing equal numbers of the items. The values of the variable at which the divisions occur are known then as the first, second, third, and so on, *deciles*. This idea is used by educational psychologists to divide pupils into 'top 10%, second 10%, third 10%,' and so on, with regard to inherent intelligence in so far as that characteristic may be measured by tests.

Yet another measure of dispersion, which depends on all the measurements, is the *mean deviation*. In order to calculate this parameter, we first of all find the arithmetic mean of the quantities in the distribution. We then find the difference between each of the items and this average, calling all the differences positive. We then add up all the differences thus obtained and find the average difference by dividing by the number of differences. Thus the mean deviation is the average difference of the several items from their arithmetic mean. In mathematical form we have

$$\text{Mean Deviation} = \frac{\Sigma|x - \bar{x}|}{n}$$

where as before the symbol $\bar{x}$ stands for the arithmetic mean of the various values of x. The sign $|x - \bar{x}|$ indicates that we are to find the difference between x and the average of the x values, ignoring sign. The sign Σ means 'add up all the terms like.'

Example. Find the arithmetic mean and mean deviation for the set of numbers: 11, 8, 6, 7, 8.

Here we have $n = 5$ items to be averaged. As previously shown, the average of the items is

$$\bar{x} = \frac{\Sigma x}{n} = \frac{11 + 8 + 6 + 7 + 8}{5} = \frac{40}{5} = 8$$

In order to get the mean difference, we calculate the several differences of the items from their average value of 8 and sum them, thus:

$$\begin{aligned} & |11 - 8| + |8 - 8| + |6 - 8| + |7 - 8| + |8 - 8| \\ = \quad & 3 \quad + \quad 0 \quad + \quad 2 \quad + \quad 1 \quad + \quad 0 \quad = 6 \end{aligned}$$

We then calculate the mean deviation by dividing this total of the deviations by $n = 5$, and so find the mean deviation as $\frac{6}{5} = 1 \cdot 2$.

The mean deviation is frequently met with in economic statistics.

The measures so far suggested are often used in elementary work on account of their being easy to calculate and easy to understand. They are, however, of no use in more advanced work because they are extremely difficult to deal with in sampling theory, on which so much of advanced work depends. The most important measure of dispersion is the *standard deviation*, which is a little more difficult to calculate and whose significance is less obvious at first sight. Calculation and interpretation, however, soon become easy with a little practice, and then the standard deviation is the most illuminating of all the parameters of dispersion. The standard deviation will be familiar to electrical engineers and mathematicians as the *root-mean-square deviation*.[5] The general reader will do well to remember this phrase as it will help him to remember exactly how the

[5] It is strictly analogous to radius of gyration in the theory of moments of inertia.

standard deviation is calculated. We shall detail the steps for the calculation of the standard deviation of a set of values thus:

Step 1. Calculate the arithmetic average of the set of values.

Step 2. Calculate the differences of the several values from their arithmetic average.

Step 3. Calculate the squares of these differences (the square of a number is found by multiplying it by itself. Thus the square of 4 is written 4^2 and has the value $4 \times 4 = 16$).

Step 4. Calculate the sum of the squares of the differences to get a quantity known as the *sample sum of squares.*

Step 5. Divide this 'sample sum of squares' by the number of items, n, in the set of values. This gives a quantity known as the *sample variance.*

Step 6. Take the square root of the variance and so obtain the standard deviation. (The square root of any number, x, is a number such that when it is multiplied by itself it gives the number x. Thus, if the square root of x is equal to a number y then we shall have $y^2 = y \times y = x$.)

This sounds much more complicated than it really is. Let us work out an example, step by step.

Example. Find the standard deviation of the set of values 11, 8, 6, 7, 8.

Step 1. We calculated the arithmetic average previously as $\bar{x} = 8$.

Step 2. The differences of the items from this average (sign may be ignored) are: 3, 0, 2, 1, 0.

Step 3. The squares of these differences are:

$3 \times 3 = 9 \qquad 0 \times 0 = 0 \qquad 2 \times 2 = 4 \qquad 1 \times 1 = 1 \qquad 0 \times 0 = 0$

Step 4. The sample sum of squares is: $9 + 0 + 4 + 1 + 0 = 14$.

Step 5. Dividing the sample sum of squares by the number of items, $n = 5$, we get the sample variance as $s^2 = 14/5 = 2{\cdot}8$ (s^2 is the accepted symbol for sample variance).

Step 6. The standard deviation is found as the square root of the sample variance thus: $s = \sqrt{2{\cdot}8} = 1{\cdot}673$.

The formula for the standard deviation is:

$$s = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n}}$$

We have seen how to calculate the standard deviation. What use is it to us in interpretation? Actually it is very easy to visualize. If we are given *any* distribution which is reasonably symmetrical about its average and which is *unimodal* (i.e., has one single hump in the centre, as in the histogram shown in Figure 13) then we find that we make very little error in assuming that two-thirds of the distribution lies less than one standard deviation away from the mean, that 95% of the distribution lies less than two standard deviations away from the mean, and that less than

1% of the distribution lies more than three standard deviations away from the mean. This is a rough rule, of course, but it is one which is found to work very well in practice. Let us suppose, for example, that we were told no more than that the distribution of intelligence, as measured by Intelligence Quotients (a person's I.Q. is defined as $\frac{\text{Mental Age}}{\text{Chronological Age}} \times 100$) has an average value $\bar{x} = 100$, with standard deviation $s = 13$. Then we might easily picture the distribution as something like the rough sketch shown in Figure 15.

The reader may care to compare the rough picture thus formed from a simple knowledge of the two measures $\bar{x}$ and s with the histogram shown in Figure 16 which is based on results obtained by L. M. Terman and

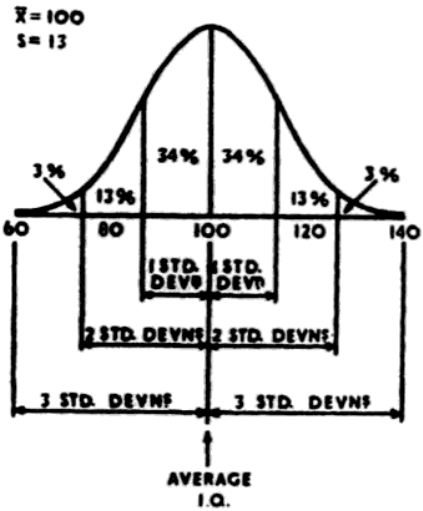

FIGURE 15—Knowing only that we have a fairly symmetrical, unimodal distribution whose mean value is I.Q. 100 units and whose standard deviation is I.Q. 13 units, we can at once picture in our minds that the distribution looks something as shown.

quoted by J. F. Kenney from his book *The Measurement of Intelligence*. This is typical of the use of measures of central tendency and dispersion in helping us to carry the broad picture of a whole distribution (provided it be reasonably symmetrical and unimodal) in the two values $\bar{x}$ and s. Such measures properly may be said to *represent* the distribution for which they were calculated.

The measures of dispersion which we have so far dealt with are all expressed in terms of the units in which the variable quantity is measured. It sometimes happens that we wish to ask ourselves whether one distribu-

tion is relatively more variable than another. Let us suppose, for example, that for the heights of men in the British Isles we find a mean value 57 inches with standard deviation 2·5 inches, and that for Spaniards the mean height is 54 inches with standard deviation 2·4 inches. It is evident that British men are taller than Spaniards and also slightly more variable in height. How are we to compare the *relative* variability bearing in mind

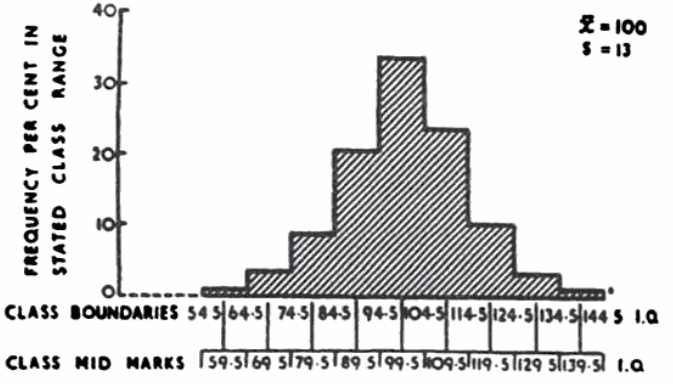

FIGURE 16—Distribution of Intelligence Quotient (compare with Figure 15). Distribution of I.Q. with $\bar{x} = 100$, $s = 13$. Based on data by L. M. Terman and quoted by J. F. Kenney, "Mathematics of Statistics," Vol. I).

that the Spaniards are shorter in height than the British? Karl Pearson's *coefficient of variation* is the most commonly used measure in practice for such a case.

It is defined as: $v = \dfrac{100s}{\bar{x}}$

If we calculate the coefficient of variation for our two cases, we get:

British $v = \dfrac{100 \times 2{\cdot}5}{67} = 37{\cdot}3\%$

Spaniards $v = \dfrac{100 \times 2{\cdot}4}{64} = 37{\cdot}5\%$

We conclude that, though the British are more variable in an absolute sense, the variability of the Spaniards, expressed as a percentage of the mean height, is just slightly greater.

For her own breakfast she'll project a scheme
Nor take her tea without a stratagem. —EDWARD YOUNG (1683–1765)

"Come little girl, you seem
To want my cup of tea
And will you take a little cream?
Now tell the truth to me"
She had a rustic woodland grin
Her cheek was soft as silk,
And she replied, "Sir, please put in
A little drop of milk." —BARRY PAIN (*The Poets at Tea*)

6 Mathematics of a Lady Tasting Tea

By SIR RONALD A. FISHER

STATEMENT OF EXPERIMENT

A LADY declares that by tasting a cup of tea made with milk she can discriminate whether the milk or the tea infusion was first added to the cup. We will consider the problem of designing an experiment by means of which this assertion can be tested. For this purpose let us first lay down a simple form of experiment with a view to studying its limitations and its characteristics, both those which appear to be essential to the experimental method, when well developed, and those which are not essential but auxiliary.

Our experiment consists in mixing eight cups of tea, four in one way and four in the other, and presenting them to the subject for judgment in a random order. The subject has been told in advance of what the test will consist, namely that she will be asked to taste eight cups, that these shall be four of each kind, and that they shall be presented to her in a random order, that is in an order not determined arbitrarily by human choice, but by the actual manipulation of the physical apparatus used in games of chance, cards, dice, roulettes, etc., or, more expeditiously, from a published collection of random sampling numbers purporting to give the actual results of such manipulation. Her task is to divide the 8 cups into two sets of 4, agreeing, if possible, with the treatments received.

INTERPRETATION AND ITS REASONED BASIS

In considering the appropriateness of any proposed experimental design, it is always needful to forecast all possible results of the experiment, and

to have decided without ambiguity what interpretation shall be placed upon each one of them. Further, we must know by what argument this interpretation is to be sustained. In the present instance we may argue as follows. There are 70 ways of choosing a group of 4 objects out of 8. This may be demonstrated by an argument familiar to students of "permutations and combinations," namely, that if we were to choose the 4 objects in succession we should have successively 8, 7, 6, 5 objects to choose from, and could make our succession of choices in $8 \times 7 \times 6 \times 5$, or 1680 ways. But in doing this we have not only chosen every possible set of 4, but every possible set in every possible order; and since 4 objects can be arranged in order in $4 \times 3 \times 2 \times 1$, or 24 ways, we may find the number of possible choices by dividing 1680 by 24. The result, 70, is essential to our interpretation of the experiment. At best the subject can judge rightly with every cup and, knowing that 4 are of each kind, this amounts to choosing, out of the 70 sets of 4 which might be chosen, that particular one which is correct. A subject without any faculty of discrimination would in fact divide the 8 cups correctly into two sets of 4 in one trial out of 70, or, more properly, with a frequency which would approach 1 in 70 more and more nearly the more often the test were repeated. Evidently this frequency, with which unfailing success would be achieved by a person lacking altogether the faculty under test, is calculable from the number of cups used. The odds could be made much higher by enlarging the experiment, while, if the experiment were much smaller even the greatest possible success would give odds so low that the result might, with considerable probability, be ascribed to chance.

THE TEST OF SIGNIFICANCE

It is open to the experimenter to be more or less exacting in respect of the smallness of the probability he would require before he would be willing to admit that his observations have demonstrated a positive result. It is obvious that an experiment would be useless of which no possible result would satisfy him. Thus, if he wishes to ignore results having probabilities as high as 1 in 20—the probabilities being of course reckoned from the hypothesis that the phenomenon to be demonstrated is in fact absent—then it would be useless for him to experiment with only 3 cups of tea of each kind. For 3 objects can be chosen out of 6 in only 20 ways, and therefore complete success in the test would be achieved without sensory discrimination, *i.e.*, by "pure chance," in an average of 5 trials out of 100. It is usual and convenient for experimenters to take 5 per cent. as a standard level of significance, in the sense that they are prepared to ignore all results which fail to reach this standard, and, by this means, to eliminate from further discussion the greater part of the fluctu-

ations which chance causes have introduced into their experimental results. No such selection can eliminate the whole of the possible effects of chance coincidence, and if we accept this convenient convention, and agree that an event which would occur by chance only once in 70 trials is decidedly "significant," in the statistical sense, we thereby admit that no isolated experiment, however significant in itself, can suffice for the experimental demonstration of any natural phenomenon; for the "one chance in a million" will undoubtedly occur, with no less and no more than its appropriate frequency, however surprised we may be that it should occur to *us*. In order to assert that a natural phenomenon is experimentally demonstrable we need, not an isolated record, but a reliable method of procedure. In relation to the test of significance, we may say that a phenomenon is experimentally demonstrable when we know how to conduct an experiment which will rarely fail to give us a statistically significant result.

Returning to the possible results of the psycho-physical experiment, having decided that if every cup were rightly classified a significant positive result would be recorded, or, in other words, that we should admit that the lady had made good her claim, what should be our conclusion if, for each kind of cup, her judgments are 3 right and 1 wrong? We may take it, in the present discussion, that any error in one set of judgments will be compensated by an error in the other, since it is known to the subject that there are 4 cups of each kind. In enumerating the number of ways of choosing 4 things out of 8, such that 3 are right and 1 wrong, we may note that the 3 right may be chosen, out of the 4 available, in 4 ways and, independently of this choice, that the 1 wrong may be chosen, out of the 4 available, also in 4 ways. So that in all we could make a selection of the kind supposed in 16 different ways. A similar argument shows that, in each kind of judgment, 2 may be right and 2 wrong in 36 ways, 1 right and 3 wrong in 16 ways and none right and 4 wrong in 1 way only. It should be noted that the frequencies of these five possible results of the experiment make up together, as it is obvious they should, the 70 cases out of 70.

It is obvious, too, that 3 successes to 1 failure, although showing a bias, or deviation, in the right direction, could not be judged as statistically significant evidence of a real sensory discrimination. For its frequency of chance occurrence is 16 in 70, or more than 20 per cent. Moreover, it is not the best possible result, and in judging of its significance we must take account not only of its own frequency, but also of the frequency for any better result. In the present instance "3 right and 1 wrong" occurs 16 times, and "4 right" occurs once in 70 trials, making 17 cases out of 70 as good as or better than that observed. The reason for including cases better than that observed becomes obvious on considering what our con-

clusions would have been had the case of 3 right and 1 wrong only 1 chance, and the case of 4 right 16 chances of occurrence out of 70. The rare case of 3 right and 1 wrong could not be judged significant merely because it was rare, seeing that a higher degree of success would frequently have been scored by mere chance.

THE NULL HYPOTHESIS

Our examination of the possible results of the experiment has therefore led us to a statistical test of significance, by which these results are divided into two classes with opposed interpretations. Tests of significance are of many different kinds, which need not be considered here. Here we are only concerned with the fact that the easy calculation in permutations which we encountered, and which gave us our test of significance, stands for something present in every possible experimental arrangement; or, at least, for something required in its interpretation. The two classes of results which are distinguished by our test of significance are, on the one hand, those which show a significant discrepancy from a certain hypothesis; namely, in this case, the hypothesis that the judgments given are in no way influenced by the order in which the ingredients have been added; and on the other hand, results which show no significant discrepancy from this hypothesis. This hypothesis, which may or may not be impugned by the result of an experiment, is again characteristic of all experimentation. Much confusion would often be avoided if it were explicitly formulated when the experiment is designed. In relation to any experiment we may speak of this hypothesis as the "null hypothesis," and it should be noted that the null hypothesis is never proved or established, but is possibly disproved, in the course of experimentation. Every experiment may be said to exist only in order to give the facts a chance of disproving the null hypothesis.

It might be argued that if an experiment can disprove the hypothesis that the subject possesses no sensory discrimination between two different sorts of object, it must therefore be able to prove the opposite hypothesis, that she can make some such discrimination. But this last hypothesis, however reasonable or true it may be, is ineligible, as a null hypothesis to be tested by experiment, because it is inexact. If it were asserted that the subject would never be wrong in her judgments we should again have an exact hypothesis, and it is easy to see that this hypothesis could be disproved by a single failure, but could never be proved by any finite amount of experimentation. It is evident that the null hypothesis must be exact, that is free from vagueness and ambiguity, because it must supply the basis of the "problem of distribution," of which the test of significance is the solution. A null hypothesis may, indeed, contain arbitrary elements,

and in more complicated cases often does so: as, for example, if it should assert that the death-rates of two groups of animals are equal, without specifying what these death-rates usually are. In such cases it is evidently the equality rather than any particular values of the death-rates that the experiment is designed to test, and possibly to disprove.

In cases involving statistical "estimation" these ideas may be extended to the simultaneous consideration of a series of hypothetical possibilities. The notion of an error of the so-called "second kind," due to accepting the null hypothesis "when it is false" may then be given a meaning in reference to the quantity to be estimated. It has no meaning with respect to simple tests of significance, in which the only available expectations are those which flow from the null hypothesis being true.

RANDOMISATION; THE PHYSICAL BASIS OF THE VALIDITY OF THE TEST

We have spoken of the experiment as testing a certain null hypothesis, namely, in this case, that the subject possesses no sensory discrimination whatever of the kind claimed; we have, too, assigned as appropriate to this hypothesis a certain frequency distribution of occurrences, based on the equal frequency of the 70 possible ways of assigning 8 objects to two classes of 4 each; in other words, the frequency distribution appropriate to a classification by pure chance. We have now to examine the physical conditions of the experimental technique needed to justify the assumption that, if discrimination of the kind under test is absent, the result of the experiment will be wholly governed by the laws of chance. It is easy to see that it might well be otherwise. If all those cups made with the milk first had sugar added, while those made with the tea first had none, a very obvious difference in flavour would have been introduced which might well ensure that all those made with sugar should be classed alike. These groups might either be classified all right or all wrong, but in such a case the frequency of the critical event in which all cups are classified correctly would not be 1 in 70, but 35 in 70 trials, and the test of significance would be wholly vitiated. Errors equivalent in principle to this are very frequently incorporated in otherwise well-designed experiments.

It is no sufficient remedy to insist that "all the cups must be exactly alike" in every respect except that to be tested. For this is a totally impossible requirement in our example, and equally in all other forms of experimentation. In practice it is probable that the cups will differ perceptibly in the thickness or smoothness of their material, that the quantities of milk added to the different cups will not be exactly equal, that the strength of the infusion of tea may change between pouring the first and the last cup, and that the temperature also at which the tea is tasted will change during the course of the experiment. These are only examples

of the differences probably present; it would be impossible to present an exhaustive list of such possible differences appropriate to any one kind of experiment, because the uncontrolled causes which may influence the result are always strictly innumerable. When any such cause is named, it is usually perceived that, by increased labour and expense, it could be largely eliminated. Too frequently it is assumed that such refinements constitute improvements to the experiment. Our view, which will be much more fully exemplified in later sections, is that it is an essential characteristic of experimentation that it is carried out with limited resources, and an essential part of the subject of experimental design to ascertain how these should be best applied; or, in particular, to which causes of disturbance care should be given, and which *ought* to be deliberately ignored. To ascertain, too, for those which are not to be ignored, to what *extent* it is worth while to take the trouble to diminish their magnitude. For our present purpose, however, it is only necessary to recognise that, whatever degree of care and experimental skill is expended in equalising the conditions, other than the one under test, which are liable to affect the result, this equalisation must always be to a greater or less extent incomplete, and in many important practical cases will certainly be grossly defective. We are concerned, therefore, that this inequality, whether it be great or small, shall not impugn the exactitude of the frequency distribution, on the basis of which the result of the experiment is to be appraised.

THE EFFECTIVENESS OF RANDOMISATION

The element in the experimental procedure which contains the essential safeguard is that the two modifications of the test beverage are to be prepared "in random order." This, in fact, is the only point in the experimental procedure in which the laws of chance, which are to be in exclusive control of our frequency distribution, have been explicitly introduced. The phrase "random order" itself, however, must be regarded as an incomplete instruction, standing as a kind of shorthand symbol for the full procedure of randomisation, by which the validity of the test of significance may be guaranteed against corruption by the causes of disturbance which have not been eliminated. To demonstrate that, with satisfactory randomisation, its validity is, indeed, wholly unimpaired, let us imagine all causes of disturbance—the strength of the infusion, the quantity of milk, the temperature at which it is tasted, etc.—to be predetermined for each cup; then since these, on the null hypothesis, are the only causes influencing classification, we may say that the probabilities of each of the 70 possible choices or classifications which the subject can make are also predetermined. If, now, after the disturbing causes are fixed, we assign, strictly at random, 4 out of the 8 cups to each of our

experimental treatments, then every set of 4, whatever its probability of being so classified, will certainly have a probability of exactly 1 in 70 of *being* the 4, for example, to which the milk is added first. However important the causes of disturbance may be, even if they were to make it certain that one particular set of 4 should receive this classification, the probability that the 4 so classified and the 4 which ought to have been so classified should be the same, must be rigorously in accordance with our test of significance.

It is apparent, therefore, that the random choice of the objects to be treated in different ways would be a complete guarantee of the validity of the test of significance, if these treatments were the last in time of the stages in the physical history of the objects which might affect their experimental reaction. The circumstance that the experimental treatments cannot always be applied last, and may come relatively early in their history, causes no practical inconvenience; for subsequent causes of differentiation, if under the experimenter's control, as, for example, the choice of different pipettes to be used with different flasks, can either be predetermined before the treatments have been randomised, or, if this has not been done, can be randomised on their own account; and other causes of differentiation will be either (*a*) consequences of differences already randomised, or (*b*) natural consequences of the difference in treatment to be tested, of which on the null hypothesis there will be none, by definition, or (*c*) effects supervening by chance independently from the treatments applied. Apart, therefore, from the avoidable error of the experimenter himself introducing with his test treatments, or subsequently, other differences in treatment, the effects of which the experiment is not intended to study, it may be said that the simple precaution of randomisation will suffice to guarantee the validity of the test of significance, by which the result of the experiment is to be judged.

THE SENSITIVENESS OF AN EXPERIMENT. EFFECTS OF ENLARGEMENT AND REPETITION

A probable objection, which the subject might well make to the experiment so far described, is that only if every cup is classified correctly will she be judged successful. A single mistake will reduce her performance below the level of significance. Her claim, however, might be, not that she could draw the distinction with invariable certainty, but that, though sometimes mistaken, she would be right more often than not; and that the experiment should be enlarged sufficiently, or repeated sufficiently often, for her to be able to demonstrate the predominance of correct classifications in spite of occasional errors.

An extension of the calculation upon which the test of significance was

based shows that an experiment with 12 cups, six of each kind, gives, on the null hypothesis, 1 chance in 924 for complete success, and 36 chances for 5 of each kind classified right and 1 wrong. As 37 is less than a twentieth of 924, such a test could be counted as significant, although a pair of cups have been wrongly classified; and it is easy to verify that, using larger numbers still, a significant result could be obtained with a still higher proportion of errors. By increasing the size of the experiment, we can render it more sensitive, meaning by this that it will allow of the detection of a lower degree of sensory discrimination, or, in other words, of a quantitatively smaller departure from the null hypothesis. Since in every case the experiment is capable of disproving, but never of proving this hypothesis, we may say that the value of the experiment is increased whenever it permits the null hypothesis to be more readily disproved.

The same result could be achieved by repeating the experiment, as originally designed, upon a number of different occasions, counting as a success all those occasions on which 8 cups are correctly classified. The chance of success on each occasion being 1 in 70, a simple application of the theory of probability shows that 2 or more successes in 10 trials would occur, by chance, with a frequency below the standard chosen for testing significance; so that the sensory discrimination would be demonstrated, although, in 8 attempts out of 10, the subject made one or more mistakes. This procedure may be regarded as merely a second way of enlarging the experiment and, thereby, increasing its sensitiveness, since in our final calculation we take account of the aggregate of the entire series of results, whether successful or unsuccessful. It would clearly be illegitimate, and would rob our calculation of its basis, if the unsuccessful results were not all brought into the account.

QUALITATIVE METHODS OF INCREASING SENSITIVENESS

Instead of enlarging the experiment we may attempt to increase its sensitiveness by qualitative improvements; and these are, generally speaking, of two kinds: (*a*) the reorganisation of its structure, and (*b*) refinements of technique. To illustrate a change of structure we might consider that, instead of fixing in advance that 4 cups should be of each kind, determining by a random process how the subdivision should be effected, we might have allowed the treatment of each cup to be determined independently by chance, as by the toss of a coin, so that each treatment has an equal chance of being chosen. The chance of classifying correctly 8 cups randomised in this way, without the aid of sensory discrimination, is 1 in 2^8, or 1 in 256 chances, and there are only 8 chances of classifying 7 right and 1 wrong; consequently the sensitiveness of the experiment has been increased, while still using only 8 cups, and it is possible to score a

significant success, even if one is classified wrongly. In many types of experiment, therefore, the suggested change in structure would be evidently advantageous. For the special requirements of a psycho-physical experiment, however, we should probably prefer to forego this advantage, since it would occasionally occur that all the cups would be treated alike, and this, besides bewildering the subject by an unexpected occurrence, would deny her the real advantage of judging by comparison.

Another possible alteration to the structure of the experiment, which would, however, decrease its sensitiveness, would be to present determined, but unequal, numbers of the two treatments. Thus we might arrange that 5 cups should be of the one kind and 3 of the other, choosing them properly by chance, and informing the subject how many of each to expect. But since the number of ways of choosing 3 things out of 8 is only 56, there is now, on the null hypothesis, a probability of a completely correct classification of 1 in 56. It appears in fact that we cannot by these means do better than by presenting the two treatments in equal numbers, and the choice of this equality is now seen to be justified by its giving to the experiment its maximal sensitiveness.

With respect to the refinements of technique, we have seen above that these contribute nothing to the validity of the experiment, and of the test of significance by which we determine its result. They may, however, be important, and even essential, in permitting the phenomenon under test to manifest itself. Though the test of significance remains valid, it may be that without special precautions even a definite sensory discrimination would have little chance of scoring a significant success. If some cups were made with India and some with China tea, even though the treatments were properly randomised, the subject might not be able to discriminate the relatively small difference in flavour under investigation, when it was confused with the greater differences between leaves of different origin. Obviously, a similar difficulty could be introduced by using in some cups raw milk and in others boiled, or even condensed milk, or by adding sugar in unequal quantities. The subject has a right to claim, and it is in the interests of the sensitiveness of the experiment, that gross differences of these kinds should be excluded, and that the cups should, not as far as *possible*, but as far as is practically convenient, be made alike in all respects except that under test.

How far such experimental refinements should be carried is entirely a matter of judgment, based on experience. The validity of the experiment is not affected by them. Their sole purpose is to increase its sensitiveness, and this object can usually be achieved in many other ways, and particularly by increasing the size of the experiment. If, therefore, it is decided that the sensitiveness of the experiment should be increased, the experi-

menter has the choice between different methods of obtaining equivalent results; and will be wise to choose whichever method is easiest to him, irrespective of the fact that previous experimenters may have tried, and recommended as very important, or even essential, various ingenious and troublesome precautions.

COMMENTARY ON

The Scientific Aptitude of Mr. George Bernard Shaw

BERNARD SHAW was not at his best as a scientific thinker. Science interested him but he was inclined to be erratic. Though science offered a fertile field for the exercise of his talents as a controversialist, a foe of pretense and a joker, he was often unable to distinguish between the stuffed robe and the honest scientist, between theories that merited serious attention and theories that were pure humbug. Moreover he himself espoused the most incredible nonsense. He fought vivisection and vaccination; he had a low opinion of medical knowledge and an even lower opinion of its practitioners; he had his own astonishing theories of biology, physiology, bacteriology and hygiene, and nothing would persuade him that the sun was burning itself out (since he expected to live longer than Methuselah he felt he had a personal stake in the catastrophe); he dismissed laboratory experiments generally as mere "put-up jobs," performances rigged for the purpose of proving preconceived theories regardless of the weight of evidence.

But for all his prejudices and eccentric notions, Shaw did not close his mind to the important works of science. He followed the advances of research in fields as varied as Pavlov's work on dogs and the Michelson-Morley interferometer experiments on ether drift. He "liked visiting laboratories and peeping at bacteria through the microscope." [1] He was curious about how things work: automobiles, radios, machine tools, motorcycles, phonographs. He was an enthusiastic photographer and camera tinkerer. Every efficient labor-saving device won his admiration but "for old-fashioned factory machinery his contempt was boundless: he said a louse could have invented it all if it had been keen enough on profits." [2]

Shaw and Jonathan Swift were much alike in their attitudes to science. Both men lived in periods of great scientific advance; both respected science; neither had any special aptitude for it. Both approached the subject as social reformers and satirists; both despised pretentiousness; neither had much use for science as a purely speculative activity. Swift aimed his wit at mathematics, which in its advanced forms seemed to him completely trivial; Shaw waged war on biological practices which he thought

[1] Hesketh Pearson, *G.B.S., A Full Length Portrait*, New York, 1942, p. 270. I have drawn on this biography for many of the details of this sketch; also on Bernard Shaw, *Sixteen Self-Sketches*, New York, 1949.

[2] Pearson, *op. cit.*, p. 270.

cruel and stupid. That he exaggerated is understandable. He enjoyed exaggeration and he regarded it as an essential tool of reform. "If you do not say a thing in an irritating way, you may just as well not say it at all, since nobody will trouble themselves about anything that does not trouble them." (The grammar *is* bizarre, even for G.B.S.)

There was, however, one branch of science which Shaw neither tilted at nor enlarged with theories of his own. The subject he spared was mathematics. He did not minimize its importance and he admitted, dropping at least this once the pose of omniscience, that he knew very little about it. He blamed his ignorance on the wretched instruction he received at the Wesleyan Connexional School. "Not a word was said to us about the meaning or utility of mathematics: we were simply asked to explain how an equilateral triangle could be constructed by the intersection of two circles, and to do sums in a, b, and x instead of in pence and shillings, leaving me so ignorant that I concluded that a and b must mean eggs and cheese and x nothing, with the result that I rejected algebra as nonsense, and never changed that opinion until in my advanced twenties Graham Wallas and Karl Pearson convinced me that instead of being taught mathematics I had been made a fool of."

The influence of these distinguished men was highly beneficial. To be sure Shaw never became unduly proficient as a calculator: "I never used a logarithm in my life, and could not undertake to extract the square root of four without misgivings." But he learned to appreciate the importance of at least one division of higher mathematics, the theory of probability and statistics. The following selection presents a Shavian version of the development and practical application of the calculus of chance. It is a delightful account and very sensible. No one else, so far as I know, has treated the history of mathematics in this way. I suspect that if there were more Shaws teaching the subject it would become popular. But the mathematical probability of this compound circumstance is admittedly small.

Let the king prohibit gambling and betting in his kingdom, for these are vices that destroy the kingdoms of princes. —THE CODE OF MANU (*c. 100*)

In play there are two pleasures for your choosing—
The one is winning, and the other losing. —BYRON

7 The Vice of Gambling and the Virtue of Insurance

By GEORGE BERNARD SHAW

INSURANCE, though founded on facts that are inexplicable, and risks that are calculable only by professional mathematicians called actuaries, is nevertheless more congenial as a study than the simpler subjects of banking and capital. This is because for every competent politician in our country there must be at least a hundred thousand gamblers who make bets every week with turf bookmakers. The bookmaker's business is to bet against any horse entered for a race with anybody who thinks it will win and wants to bet that it will. As only one horse can win, and all the rest must lose, this business would be enormously lucrative if all the bets were for even money. But the competition among bookmakers leads them to attract customers by offering "odds," temptingly "long," against horses unlikely to win: whilst giving no odds at all on the most likely horse, called the favorite. The well-known cry, puzzling to novices, of "two to one bar one" means that the bookmaker will bet at odds of two to one against any horse in the race except the favorite. Mostly, however, he will bet at odds of ten to one or more against an "outsider." In that case, if, as sometimes happens, the outsider wins, the bookmaker may lose on his bet against it all that he gained on his bets against the favorites. On the scale between the possible extremes of gain and loss he may come out anywhere according to the number of horses in the race, the number of bets made on each of them, and the accuracy of his judgment in guessing the odds he may safely offer. Usually he gains when an outsider wins, because mostly there is more money laid on favorites and fancies than on outsiders; but the contrary is possible; for there may be several outsiders as well as several favorites; and, as outsiders win quite often, to tempt customers by offering too long odds against them is gambling; and a bookmaker must never gamble, though he lives by gambling. There are practically always enough variable factors in the game to tax the bookmaker's financial ability to the utmost. He must budget so as to come out at worst still solvent. A bookmaker who gambles will ruin himself as cer-

tainly as a licensed victualler (publican) who drinks, or a picture dealer who cannot bear to part with a good picture.

The question at once arises, how is it possible to budget for solvency in dealing with matters of chance? The answer is that when dealt with in sufficient numbers matters of chance become matters of certainty, which is one of the reasons why a million persons organized as a State can do things that cannot be dared by private individuals. The discovery of this fact nevertheless was made in the course of ordinary private business.

In ancient days, when travelling was dangerous, and people before starting on a journey overseas solemnly made their wills and said their prayers as if they were going to die, trade with foreign countries was a risky business, especially when the merchant, instead of staying at home and consigning his goods to a foreign firm, had to accompany them to their destination and sell them there. To do this he had to make a bargain with a ship owner or a ship captain.

Now ship captains, who live on the sea, are not subject to the terrors it inspires in the landsman. To them the sea is safer than the land; for shipwrecks are less frequent than diseases and disasters on shore. And ship captains make money by carrying passengers as well as cargo. Imagine then a business talk between a merchant greedy for foreign trade but desperately afraid of being shipwrecked or eaten by savages, and a skipper greedy for cargo and passengers. The captain assures the merchant that his goods will be perfectly safe, and himself equally so if he accompanies them. But the merchant, with his head full of the adventures of Jonah, St. Paul, Odysseus, and Robinson Crusoe, dares not venture. Their conversation will be like this:

CAPTAIN. Come! I will bet you umpteen pounds that if you sail with me you will be alive and well this day year.

MERCHANT. But if I take the bet I shall be betting you that sum that I shall die within the year.

CAPTAIN. Why not if you lose the bet, as you certainly will?

MERCHANT. But if I am drowned you will be drowned too; and then what becomes of our bet?

CAPTAIN. True. But I will find you a landsman who will make the bet with your wife and family.

MERCHANT. That alters the case of course; but what about my cargo?

CAPTAIN. Pooh! The bet can be on the cargo as well. Or two bets: one on your life, the other on the cargo. Both will be safe, I assure you. Nothing will happen; and you will see all the wonders that are to be seen abroad.

MERCHANT. But if I and my goods get through safely I shall have to pay you the value of my life and of the goods into the bargain. If I am not drowned I shall be ruined.

CAPTAIN. That also is very true. But there is not so much for me in it as you think. If you are drowned I shall be drowned first; for I must be the last man to leave the sinking ship. Still, let me persuade you to venture. I will make the bet ten to one. Will that tempt you?

MERCHANT. Oh, in that case—

The captain has discovered insurance just as the goldsmiths discovered banking.

It is a lucrative business; and, if the insurer's judgment and information are sound, a safe one. But it is not so simple as bookmaking on the turf, because in a race, as all the horses but one must lose and the bookmaker gain, in a shipwreck all the passengers may win and the insurer be ruined. Apparently he must therefore own, not one ship only, but several, so that, as many more ships come safely to port than sink, he will win on half a dozen ships and lose on one only. But in fact the marine insurer need no more own ships than the bookmaker need own horses. He can insure the cargoes and lives in a thousand ships owned by other people without his having ever owned or even seen as much as a canoe. The more ships he insures the safer are his profits; for half a dozen ships may perish in the same typhoon or be swallowed by the same tidal wave; but out of a thousand ships most by far will survive. When the risks are increased by war the odds on the bets can be lowered.

When foreign trade develops to a point at which marine insurers can employ more capital than individual gamesters can supply, corporations like the British Lloyds are formed to supply the demand. These corporations soon perceive that there are many more risks in the world than the risk of shipwreck. Men who never travel nor send a parcel across the seas, may lose life or limb by accident, or have their houses burnt or robbed. Insurance companies spring up in all directions; and the business extends and develops until there is not a risk that cannot be insured. Lloyds will bet not only against shipwreck but against almost any risk that is not specifically covered by the joint stock companies, provided it is an insurable risk: that is, a safe one.

This provision is a contradiction in terms; for how can a safe transaction involve a risk or a risk be run safely?

The answer takes us into a region of mystery in which the facts are unreasonable by any method of ratiocination yet discovered. The stock example is the simplest form of gambling, which is tossing a coin and betting on which side of it will be uppermost when it falls and comes to rest. Heads or tails they call it in England, head or harp in Ireland. Every time the coin is tossed, each side has an equal chance with the other of winning. If head wins it is just as likely to win the next time and the next and so on to the thousandth; so that on reasonable grounds a thousand heads in succession are possible, or a thousand tails; for the fact that

head wins at any toss does not raise the faintest reasonable probability that tails will win next time. Yet the facts defy this reasoning. Anyone who possesses a halfpenny and cares to toss it a hundred times may find the same side turning up several times in succession; but the total result will be fifty-fifty or as near thereto as does not matter. I happen to have in my pocket ten pennies; and I have just spilt them on the floor ten times. Result: forty nine heads and fifty one tails, though five-five occurred only twice in the ten throws, and heads won three times in succession to begin with. Thus though as between any two tosses the result is completely uncertain, in ten throws it may be six-four or seven-three often enough to make betting a gamble; but in a hundred the result will certainly be close enough to fifty-fifty to leave two gamblers, one crying heads and the other tails every time, exactly or very nearly where they were when they started, no richer and no poorer, unless the stakes are so high that only players out of their senses would hazard them.

An insurance company, sanely directed, and making scores of thousands of bets, is not gambling at all; it knows with sufficient accuracy at what age its clients will die, how many of their houses will be burnt every year, how often their houses will be broken into by burglars, to what extent their money will be embezzled by their cashiers, how much compensation they will have to pay to persons injured in their employment, how many accidents will occur to their motor cars and themselves, how much they will suffer from illness or unemployment, and what births and deaths will cost them: in short, what will happen to every thousand or ten thousand or a million people even when the company cannot tell what will happen to any individual among them.

In my boyhood I was equipped for an idle life by being taught to play whist, because there were rich people who, having nothing better to do, escaped from the curse of boredom (then called ennui) by playing whist every day. Later on they played bezique instead. Now they play bridge. Every gentleman's club has its card room. Card games are games of chance; for though the players may seem to exercise some skill and judgment in choosing which card to play, practice soon establishes rules by which the stupidest player can learn how to choose correctly: that is, not to choose at all but to obey the rules. Accordingly people who play every day for sixpences or shillings find at the end of the year that they have neither gained nor lost sums of any importance to them, and have killed time pleasantly instead of being bored to death. They have not really been gambling any more than the insurance companies.

At last it is discovered that insurers not only need not own ships or horses or houses or any of the things they insure, but that they need not exist. Their places can be taken by machines. On the turf the bookmaker, flamboyantly dressed and brazenly eloquent, is superseded by the Total-

izator (Tote for short) in which the gamblers deposit the sums they are prepared to stake on the horses they fancy. After the race all the money staked on the winner is divided among its backers. The machine keeps the rest. On board pleasure ships young ladies with more money than they know what to do with drop shillings into gambling machines so constructed that they occasionally return the shilling ten or twentyfold. These are the latest successors of the roulette table, the "little horses," the dice casters, and all other contraptions which sell chances of getting money for nothing. Like the Tote and the sweepstake, they do not gamble: they risk absolutely nothing, though their customers have no certainty except that in the lump they must lose, every gain to Jack and Jill being a loss to Tom and Susan.

How does all this concern the statesman? In this way. Gambling, or the attempt to get money without earning it, is a vice which is economically (that is, fundamentally) ruinous. In extreme cases it is a madness which persons of the highest intelligence are unable to resist: they will stake all they possess though they know that the chances are against them. When they have beggared themselves in half an hour or half a minute, they sit wondering at the folly of the people who are doing the same thing, and at their own folly in having done it themselves.

Now a State, being able to make a million bets whilst an individual citizen can afford only one, can tempt him or her to gamble without itself running the slightest risk of losing financially; for, as aforesaid, what will happen in a million cases is certain, though no one can foresee what will happen in any one case. Consequently governments, being continually in pressing need of money through the magnitude of their expenses and the popular dislike of taxation, are strongly tempted to replenish the Treasury by tempting their citizens to gamble with them.

No crime against society could be more wickedly mischievous. No public duty is more imperative than the duty of creating a strong public conscience against it, making it a point of bare civic honesty not to spend without earning nor consume without producing, and a point of high civic honor to earn more than you spend, to produce more than you consume, and thus leave the world better off than you found it. No other real title to gentility is conceivable nowadays.

Unfortunately our system of making land and capital private property not only makes it impossible for either the State or the Church to inculcate these fundamental precepts but actually drives them to preach just the opposite. The system may urge the energetic employer to work hard and develop his business to the utmost; but his final object is to become a member of the landed gentry or the plutocracy, living on the labor of others and enabling his children to do the same without ever having worked at all. The reward of success in life is to become a parasite and

found a race of parasites. Parasitism is the linchpin of the Capitalist applecart: the main Incentive without which, we are taught, human society would fall to pieces. The boldest of our archbishops, the most democratic of our finance ministers, dares not thunder forth that parasitism, for peers and punters alike, is a virus that will rot the most powerful civilization, and that the contrary doctrine is diabolical. Our most eminent churchmen do not preach very plainly and urgently against making selfishness the motive power of civilization; but they have not yet ventured to follow Ruskin and Proudhon in insisting definitely that a citizen who is neither producing goods nor performing services is in effect either a beggar or a thief. The utmost point yet reached in England is the ruling out of State lotteries and the outlawing of the Irish sweepstakes.

But here again the matter is not simple enough to be disposed of by counsels of Socialist perfection in the abstract. There are periods in every long lifetime during which one must consume without producing. Every baby is a shamelessly voracious parasite. And to turn the baby into a highly trained producer or public servant, and make its adult life worth living, its parasitism must be prolonged well into its teens. Then again old people cannot produce. Certain tribes who lay an excessive stress on Manchester School economics get over this difficulty easily by killing their aged parents or turning them out to starve. This is not necessary in modern civilization. It is quite possible to organize society in such a manner as to enable every ablebodied and ableminded person to produce enough not only to pay their way but to repay the cost of twenty years education and training, making it a first-rate investment for the community, besides providing for the longest interval between disablement by old age and natural death. To arrange this is one of the first duties of the modern statesman.

Now childhood and old age are certainties. What about accidents and illnesses, which for the individual citizen are not certainties but chances? Well, we have seen that what are chances for the individual are certainties for the State. The individual citizen can share its certainty only by gambling with it. To insure myself against accident or illness I must make a bet with the State that these mishaps will befall me; and the State must accept the bet, the odds being fixed by the State actuaries mathematically. I shall at once be asked Why with the State? Why not with a private insurance company? Clearly because the State can do what no private company can do. It can compel every citizen to insure, however improvident or confident in his good luck he may happen to be, and thus, by making a greater number of bets, combine the greatest profit with the greatest certainty, and put the profit into the public treasury for the general good. It can effect an immense saving of labor by substituting a single organization for dozens of competing ones. Finally it can insure at cost

price, and, by including the price in the general rate of taxation, pay for all accidents and illnesses directly and simply without the enormous clerical labor of collecting specific contributions or having to deal in any way with the mass of citizens who lose their bets by having no accidents nor illnesses at any given moment.

The oddity of the situation is that the State, to make insurance certain and abolish gambling, has to compel everyone to gamble, becoming a Supertote and stakeholder for the entire population.

As ship insurance led to life insurance, life insurance to fire insurance and so on to insurance against employer's liability, death duties, and unemployment, the list of insurable risks will be added to, and insurance policies will become more comprehensive from decade to decade, until no risks that can worry a reasonably reckless citizen are left uncovered. And when the business of insurance is taken on by the State and lumped into the general taxation account, every citizen will be born with an unwritten policy of insurance against all the common risks, and be spared the painful virtues of providence, prudence, and self-denial that are now so oppressive and demoralizing, thus greatly lightening the burden of middle-class morality. The citizens will be protected whether they like it or not, just as their children are now educated and their houses now guarded by the police whether they like it or not, even when they have no children to be educated nor houses to be guarded. The gain in freedom from petty cares will be immense. Our minds will no longer be crammed and our time wasted by uncertainty as to whether there will be any dinners for the family next week or any money left to pay for our funerals when we die.

There is nothing impossible or even unreasonably difficult in all this. Yet as I write, a modest and well thought-out plan of national insurance by Sir William Beveridge, whose eminence as an authority on political science nobody questions, is being fiercely opposed, not only by the private insurance companies which it would supersede, but by people whom it would benefit; and its advocates mostly do not understand it and do not know how to defend it. If the schooling of our legislators had included a grounding in the principles of insurance the Beveridge scheme would pass into law and be set in operation within a month. As it is, if some mutilated remains of it survive after years of ignorant squabbling we shall be lucky, unless, indeed, some war panic drives it through Parliament without discussion or amendment in a few hours. However that may be, it is clear that nobody who does not understand insurance and comprehend in some degree its enormous possibilities is qualified to meddle in national business. And nobody can get that far without at least an acquaintance with the mathematics of probability, not to the extent of making its calculations and filling examination papers with typical equa-

tions, but enough to know when they can be trusted, and when they are cooked. For when their imaginary numbers correspond to exact quantities of hard coins unalterably stamped with heads and tails, they are safe within certain limits; for here we have solid certainty and two simple possibilities that can be made practical certainties by an hour's trial (say one constant and one variable that does not really vary); but when the calculation is one of no constant and several very capricious variables, guesswork, personal bias, and pecuniary interests, come in so strongly that those who began by ignorantly imagining that statistics cannot lie end by imagining, equally ignorantly, that they never do anything else.

PART IX

The Supreme Art of Abstraction: Group Theory

COMMENTARY ON

Certain Important Abstractions

THE Theory of Groups is a branch of mathematics in which one does something to something and then compares the result with the result obtained from doing the same thing to something else, or something else to the same thing. This is a broad definition but it is not trivial. The theory is a supreme example of the art of mathematical abstraction. It is concerned only with the fine filigree of underlying relationships; it is the most powerful instrument yet invented for illuminating structure.

The term *group* was first used in a technical sense by the French mathematician Évariste Galois in 1830. He wrote his brilliant paper on the subject at the age of twenty, the night before he was killed in a stupid duel.[1] The concept was strongly developed in the nineteenth century by leading mathematicians, among them Augustin-Louis Cauchy (1789–1857),[2] Sir Arthur Cayley (see p. 341), Camille Jordan (1838–1922), and two eminent Norwegians, Ludwig Sylow (1832–1918) and Marius Sophus Lie (1842–1899). In a little more than a century it has effected a remarkable unification of mathematics, revealing connections between parts of algebra and geometry that were long considered distinct and unrelated. "Wherever groups disclosed themselves, or could be introduced, simplicity crystallized out of comparative chaos." [3] Group theory has also helped physicists penetrate to the basic structure of the phenomenal world, to catch glimpses of innermost pattern and relationship. This is as deep, it should be observed, as science is likely to get. Even if we do not accept the idea that the ultimate essence of things *is* pattern, we may conclude with Bertrand Russell that any other essence is an individuality "which always eludes words and baffles description, but which, for that very reason, is irrelevant to science."

Let us return briefly to our somewhat dreamy definition and make it more concrete and explicit. The best plan, perhaps, is to give a specific example of a group and then to erase most of its details until nothing is left but essentials. This is the famous Carrollian method of defining a grin as what remains after the Cheshire cat, the vehicle of the grin, has vanished.

The class or set of all the positive and negative integers, including zero,

[1] The story of his tragic life is dramatically told by E. T. Bell in *Men of Mathematics*, New York, 1937.

[2] "To Cauchy has been given the credit of being the founder of the theory of groups of finite order, even though fundamental results had been previously reached by J. L. Lagrange, Pietro Abbati, P. Ruffini, N. H. Abel, and Galois." Florian Cajori, *A History of Mathematics*, New York, 1919, p. 352.

[3] E. T. Bell, *Mathematics, Queen and Servant of Science*, New York, 1951, p. 164.

in conjunction with the ordinary arithmetic operation of addition, constitutes a familiar group. Its defining properties are these: (1) The sum of any two integers of the set is an integer of the set; (2) in adding three (or more) integers, any set of them may be added first without varying the result; you will recognize this as the associative rule of arithmetic (e.g., $(3 + 7) + 9 = 3 + (7 + 9)$); (3) the set contains an "identity" or "unit" element (namely zero) such that the sum of this element and any other element in the set is again the latter element (e.g., $4 + 0 = 4$, $0 + 8 = 8$, etc.); (4) every integer in the set has an inverse or reciprocal, such that the sum of the two is the identity element (e.g., $2 + (-2) = 0$, $-77 + 77 = 0$, etc.). These are the attributes of our particular group. Now for some erasures. (1) The elements of the set may be arithmetic objects (e.g., numbers), geometric objects (e.g., points), physical entities (e.g., atoms), or they may be undefined;[4] (2) their number may be finite or infinite; (3) the operation or rule of combining the elements may be an arithmetic process (e.g., addition, multiplication), a geometric process (e.g., rotation, translation), or it may be undefined. Two further conditions are essential: (4) the combining rule must be associative; (5) every element of the set must have an inverse. Besides these five conditions, a set may be Abelian or non-Abelian according as the combining rule is commutative or noncommutative (i.e., for addition, either $2 + 3 = 3 + 2$ or $2 + 3 \neq 3 + 2$, and, for multiplication, either $2 \cdot 3 = 3 \cdot 2$ or $2 \cdot 3 \neq 3 \cdot 2$). These are the bare bones of the group concept. It is hard to believe how much unification of bewildering details has been achieved by the theory; "what a wealth, what a grandeur of thought may spring from what slight beginnings."[5]

A few words should be added concerning two other fundamental mathematical terms which arise frequently in group theory. The first is *transformation*, which embodies the idea of change or motion. An algebraic expression is transformed by changing it to another having different form, by substituting for the variables their values in terms of another set of variables; a geometric figure is transformed by changing its co-ordinates, by mapping one space on another, by moving the figure pursuant to a procedural rule, e.g., projection, rotation, translation.[6] More generally,

[4] It should be emphasized that while both the elements and operations of a group may theoretically be undefined, if the group is to be useful in science they must in some way correspond to elements and operations of observable experience. Otherwise manipulating the group amounts to nothing more than a game, and a pretty vague and arid game at that, suitable only for the most withdrawn lunatics.

[5] The British geometer H. F. Baker, as quoted by Florian Cajori in *A History of Mathematics*, New York, 1919, p. 283.

[6] This definition of transformation suffices for our purposes, but the concept is much more comprehensive than I have indicated. Any problem, process or operation, as Keyser says (see next note), having to do with ordinary functions is a problem, process or operation having to do with relations or transformations. A pairing or coupling is a transformation; so is a relation, a function, a mathematical calculation, a deductive inference.

any object of thought may be transformed by associating it with or converting it into another object of thought.

The second important term is *invariance*. The invariant properties of an algebraic expression, geometric figure, class, or other object of thought are those which remain the same under transformations. Suppose the elements of the class of positive integers (1, 2, 3 · · ·) are transformed by the rule of doubling each element

$$\begin{array}{cccccc} 1 & 2 & 3 & 4 & \cdot & \cdot \cdot \cdot \\ \downarrow & \downarrow & \downarrow & \downarrow & & \\ 2 & 4 & 6 & 8 & \cdot & \cdot \cdot \cdot \end{array}$$

The transforms constitute the class of even integers: 2, 4, 6, · · ·. Since the integers are transformed into integers, the property of being an integer is evidently preserved; which is to say it is an invariant under the doubling transformation. However, the *value* of each integer of the original class is not preserved; it is doubled; thus value is not an invariant under this transformation. Another example of an invariant in our case are the ratios of the elements of each class: if the elements of the first class are called x's and those of the transformed class y's then under the rule $y = 2x$,

$$\frac{y_1}{y_2} = \frac{2x_1}{2x_2} = \frac{x_1}{x_2}.$$[7]

A geometric invariant is similar. Take a rigid object such as a glass paper-weight and move it by sliding from one end of the table to the other. This is a transformation. The paper-weight retains its shape and dimensions; the retained properties are therefore invariants. The mathematician describes these facts by saying that the metric properties of rigid bodies are invariant under the transformation of motion. Since the paper-weight's position and distance from an object such as the mirror on the wall or the Pole Star are changed by the transformation, they are not invariant. If the object moved were a blob of mercury it is unlikely that it would retain its shape or dimensions, but its mass would probably be invariant and certainly its atomic structure. (For a further discussion of this point see pp. 581–598, selection on topology.)

Group theory has to do with the invariants of groups of transformations. One studies the properties of an object, the features of a problem unaffected by changes of condition. The more drastic the changes, the fewer the invariants. What better way to get at the fundamentals of structure than by successive transformations to strip away the secondary properties. It is a method analogous to that used by the archaeologist who clears away hills to get at cities, digs into houses to uncover ornaments, utensils and potsherds, tunnels into tombs to find sarcophagi, the winding

[7] The example is from Cassius J. Keyser, *Mathematical Philosophy*, New York, 1922, pp. 183–185.

sheets they hold and the mummies within. Thus he reconstructs the features of an unseen society; and so the mathematician and scientist create a theoretical counterpart of the unseen structure of the phenomenal world. Whitehead has characterized these efforts in a famous observation: "To see what is general in what is particular and what is permanent in what is transitory is the aim of scientific thought."

* * * * *

I have selected two essays to illustrate group theory. The first is a chapter from Cassius J. Keyser's lectures on mathematical philosophy.[8] Keyser, a prominent American mathematician, was born in Rawson, Ohio, in 1862. He was educated in Ohio schools and at Missouri University; for a time he studied law but then turned to mathematics and earned his graduate degrees at Columbia. After five years as superintendent of schools in Ohio and Montana (1885–90), he became professor of mathematics at the New York State Normal School, and in 1897 joined the staff of Columbia. He was appointed Adrain professor in 1904, serving in this post until 1927 when he was made emeritus. He died in 1947, aged eighty-five. Keyser had broad interests in mathematics, as a geometer, historian and philosopher. He was much admired as a teacher for the care he took to his lectures, their breadth, clarity and honesty. He is a little old-fashioned in his style and a trifle long-winded. Keyser was not the man to drop a point until he had squeezed it dry both as to its scientific content and cultural bearings; the reader may also come to feel a little squeezed. But he had an unfailingly interesting and reflective mind, and I have nowhere found a better survey of the group concept than in the selection below. The second essay, by Sir Arthur Eddington, is one of the Messenger Lectures given at Cornell University, appearing in a book titled *New Pathways of Science*. The discussion of groups exhibits the usual dazzling Eddington virtuosity; it is one of his best pieces of popularization in one of his best books.

[8] *Mathematical Philosophy, A Study of Fate and Freedom*, New York, 1922.

Mazes intricate.
Eccentric, interwov'd, yet regular
Then most, when most irregular they seem. —MILTON

1 The Group Concept

By CASSIUS J. KEYSER

I INVITE your attention during the present hour to the notion of group. Even if I were a specialist in group theory,—which I am not,—I could not in one hour give you anything like an extensive knowledge of it, nor facility in its technique, nor a sense of its intricacy and proportions as known to its devotees, the priests of the temple. But the hour should suffice to start you on the way to acquiring at least a minimum of what a respectable philosopher should know of this fundamental subject; and such a minimum will include: a clear conception of what the term "group" means; ability to illustrate it copiously by means of easily understood examples to be found in all the cardinal fields of interest—number, space, time, motion, relation, play, work, the world of sense-data and the world of ideas; a glimpse of its intimate connections with the ideas of transformation and invariance; an inkling of it both as subject-matter and as an instrument for the delimitation and discrimination of doctrines; and discernment of the concept as vaguely prefigured in philosophic speculation from remote antiquity down to the present time.

I believe that the best way to secure a firm hold of the notion of group is to seize upon it first in the abstract and then, by comparing it with concrete examples, gradually to win the sense of holding in your grasp a living thing. In presenting the notion of group in the abstract, it is convenient to use the term *system*. This term has many meanings in mathematics and so at the outset we must clearly understand the sense in which the term is to be employed here. The sense is this: as employed in the definition of group, the term system means some definite class of things together with some definite rule, or way, in accordance with which any member of the class can be combined with any member of it (either with itself or any other member). For a simple example of such a system we may take for the class the class of ordinary whole numbers and for the rule of combination the familiar rule of addition. You should note that there are three and only three respects in which two systems can differ: by having different classes, by having different rules of combination, and by differing in both of these ways.

The definition of the term "group" is as follows.

Let S denote a system consisting of a class C (whose members we will denote by a, b, c and so on) and of a rule of combination (which rule we will denote by the symbol $\circ$, so that by writing, for example, $a \circ b$, we shall mean the result of combining b with a). The system S is called a group if and only if it satisfies the following four conditions:

(a) If a and b are members of C, then $a \circ b$ is a member of C; that is, $a \circ b = c$, where c is some member of C.

(b) If a, b, c are members of C, then $(a \circ b) \circ c = a \circ (b \circ c)$; that is, combining c with the result of combining b with a yields the same as combining with a the result of combining c with b; that is, the rule of combination is associative.

(c) The class C contains a member i (called the identical member or element) such that if a be a member of C, then $a \circ i = i \circ a = a$; that is, C has a member such that, if it be combined with any given member, or that member with it, the result is the given member.

(d) If a be a member of C, then there is a member a' (called the reciprocal of a) such that $a \circ a' = a' \circ a = i$; that is, each member of C is matched by a member such that combining the two gives the identical member.

Other definitions of the term "group" have been proposed and are sometimes used. The definitions are not all of them equivalent but they all agree that to be a group a system must satisfy condition (a).

Systems satisfying condition (a) are many of them on that account so important that in the older literature of the subject they are called groups, or closed systems, and are now said to have "the group property," even if they do not satisfy conditions (b), (c) and (d). The propriety of the term "closed system" is evident in the fact that a system satisfying (a) is such that the result of combining any two of its members is itself a member—a thing *in* the system, not out of it.

Various Simple Examples of Groups and of Systems that Are Not Groups.—You observe that by the foregoing definition of group every group is a system; groups, as we shall see, are infinitely numerous; yet it is true that relatively few systems are groups or have even the group property—so few relatively that, if you select a system at random, it is highly probable you will thus hit upon one that is neither a group nor has the group property.

Take, for example, the system S_1 whose class C is the class of integers from 1 to 10 inclusive and whose rule of combination is that of ordinary multiplication $\times$; $3 \times 4 = 12$; 12 is not a member of C, and so S_1 is not closed—it has not the group property.

Let S_2 have for its C the class of all the ordinary integers, 1, 2, 3, . . . *ad infinitum*, and let $\circ$ be $\times$ as before; as the product of any two integers

is an integer, (a) is satisfied—S_2 is closed, has the group property; (b), too, is evidently satisfied, and so is (c), the identity element being 1 for, if n be any integer, $n \times 1 = 1 \times n = n$; but (d) is not satisfied—none of the integers (except 1) composing C has a reciprocal in C—there is, for example, no integer n such that $2 \times n = n \times 2 = 1$; and so S_2, though it has the group property, is not a group.

Let S_3 be the system consisting of the class C of all the positive and negative integers including zero and of addition as the rule of combination; you readily see that S_3 is a group, zero being the identical element, and each element having its own negative for reciprocal.

A group is said to be finite or infinite according as its C is a finite or an infinite class and it is said to be Abelian or non-Abelian according as its rule of combination is or is not commutative—according, that is, as we have or do not have $a \circ b = b \circ a$, where a and b are arbitrary members of C. You observe that the group S_3 is both infinite and Abelian.

For an example of a group that is finite and Abelian it is sufficient to take the system S_4 whose C is composed of the four numbers, $1, -1, i, -i$, where i is $\sqrt{-1}$, and whose rule of combination is multiplication; you notice that the identical element is 1, that 1 and -1 are each its own reciprocal and that i and $-i$ are each the other's reciprocal.

Let S_5 have the same C as S_3 and suppose $\circ$ to be subtraction instead of addition; show that S_5 has the group property but is not a group. Show the like for S_6 in which C is the same as before and $\circ$ denotes multiplication. Show that S_7 where C is the same as before and $\circ$ means the rule of division, has not even the group property.

Consider S_8 where C is the class of all the rational numbers (that is, all the integers and all the fractions whose terms are integers, it being understood that zero can not be a denominator) and where $\circ$ denotes $+$; you will readily find that S_8 is a group, infinite and Abelian. Examine the systems obtained by keeping the same C and letting $\circ$ denote subtraction, then multiplication, then division. Devise a group system where $\circ$ means division.

If S and S' be two groups having the same rule of combination and if the class C of S be a proper part of the class C' of S' (*i.e.*, if the members of C are members of C' but some members of C' are not in C), then S is said to be a sub-group of S'. Observe that S_3 is a sub-group of S_8.

Show that S_9 is a group if its C is the class of all real numbers and its $\circ$ is $+$; note that S_8 is a sub-group of S_9 and hence that S_3 is a sub-group of a sub-group of a group. Is S_9 itself a sub-group? If so, of what group or groups? Examine the systems derived from S_9 by altering the rule of combination.

The most difficult thing that teaching has to do is to give a worthy sense of the meaning and scope of a great idea. A great idea is always

generic and abstract but it has its living significance in the particular and concrete—in a countless multitude of differing instances or examples of it; each of these sheds only a feeble light upon the idea, leaving the infinite range of its significance in the dark; whence the necessity of examining and comparing a large number of widely differing examples in the hope that many little lights may constitute by union something like a worthy illustration; but to present these numerous examples requires an amount of time and a degree of patience that are seldom at one's disposal, and so it is necessary to be content with a selected few. And now here is the difficulty: if the examples selected be complex and difficult, they repel; if they be simple and easy, they are not impressive; in either case, the significance of the general concept in question remains ungrasped and unappreciated. I am going, however, to take the risk—to the foregoing illustrations of the group concept I am going to add a few further ones,—some of them very simple, some of them more complex,—trusting that the former may not seem to you too trivial nor the latter too hard.

Every one has seen the pretty phenomenon of a grey squirrel rapidly rotating a cylindrical wire cage enclosing it. It may rotate the cage in either of two opposite ways, senses or directions. Let us think of rotation in only one of the ways, and let us call any rotation, whether it be much or little, a *turn*. Each turn carries a point of the cage along a circle-arc of some length, short or long. Denote by R the special turn (through 360°) that brings each point of the cage back to its starting place. Let S_{10} be the system whose C is the class of all possible turns and whose o is addition of turns so that $a \circ b$ shall be the whole turn got by following turn a by turn b. You see at once that S has the group property for the sum of any two turns is a turn; it is equally evident that the associative law—condition (b)—is satisfied. Note that R is equivalent to no turn,—equivalent to rest,—equivalent to a zero turn, if you please; note that, if a be a turn greater than R and less than $2R$, then a is equivalent to a's excess over R; that, if a be greater than $2R$ and less than $3R$, then a is equivalent to a's excess over $2R$; and so on; thus any turn greater than R and not equal to a multiple of R is equivalent to a turn less than R; let us regard any turn that is thus greater than R as identical with its equivalent less than R; we have, then, to consider no turns except R and those less than R—of which there are infinitely many; you see immediately that, if a be any turn, $a \circ R = R \circ a = a$, which means that condition (c) is satisfied with R for identical element. Next notice that for any turn a there is a turn a' such that $a \circ a' = a' \circ a = R$. Hence S_{10} is, as you see, a group. Show it to be Abelian. You will find it instructive to examine the system derived from S_{10} by letting C be the class of all turns (forward or backward).

Perhaps, you will consider the system suggested by the familiar spectacle of a ladybug or a measuring-worm moving round the rim or edge of a

THE WORLD OF MATHEMATICS

Volume 3

EDITED BY JAMES NEWMAN

" . . . promises to be the most frequently used reference book on mathematics, as well as a delight to readers with a wide range of backgrounds."—The *New York Times*

The World of Mathematics, a monumental four-volume reference 15 years in the making, was specially designed to make mathematics more accessible to the layman. It comprises nontechnical essays on every aspect of the subject, including articles by and about scores of eminent mathematicians, as well as literary figures, economists, biologists, and many other thinkers. Included are selections from the works of Archimedes, Galileo, Descartes, Newton, Gregor Mendel, Edmund Halley, Jonathan Swift, John Maynard Keynes, Henri Poincaré, Lewis Carroll, George Boole, Bertrand Russell, Alfred North Whitehead, John von Neumann, and many others. In addition, an informative commentary by noted mathematics scholar James R. Newman precedes each essay or group of essays, explaining their relevance and context in the history and development of mathematics.

Volume 3 (Parts VIII–XVII) contains essays on such topics as statistics and the design of experiments, group theory, the mathematics of infinity, the unreasonableness of mathematics, the vocabulary of mathematics, and mathematics as an art. The essays gathered here include "The Law of Large Numbers," by Jacob Bernoulli; "The Vice of Gambling and the Virtue of Insurance," by George Bernard Shaw; "Mathematics and the Metaphysicians," by Bertrand Russell; and 30 more.

Unabridged republication of Vol. III of the 4-volume edition published by Simon and Schuster, New York, 1956. Indexes. Numerous text figures. 624pp. 5⅜ x 8½. Paperbound.

OTHER VOLUMES AVAILABLE

THE WORLD OF MATHEMATICS, James R. Newman (ed.). Vol. 1. 768pp. 5⅜ x 8½. 41153-2 Pa. $14.95

THE WORLD OF MATHEMATICS, James R. Newman (ed.). Vol. 2. 720pp. 5⅜ x 8½. 41150-8 Pa. $14.95

THE WORLD OF MATHEMATICS, James R. Newman (ed.). Vol. 4. 464pp. 5⅜ x 8½. 41152-4 Pa. $14.95

Free Dover Mathematics and Science Catalog (59065-8) available upon request.

ISBN 0-486-41151-6

IN USA