PUBLISHER'S NOTE

Salem Press is pleased to add *Principles of Scientific Research* as the eighth title in the *Principles of series* that includes *Chemistry, Physics, Astronomy, Computer Science, Physical Science, Biology, and Scientific Research.* This new resource introduces students and researchers to the fundamentals of scientific research using easy-to-understand language, giving readers a solid start and deeper understanding and appreciation of this complex subject.

- The 105 entries include entries that explain basic principles of scientific research, ranging from Abductive Reasoning to Type I and Type II Errors,as well as biographies of key figures in scientific research that include a description of their significant contributions to the field, ranging from Robert Brown to Chien-Shiung Wu. All of the entries, and are arranged in an A to Z order, making it easy to find the topic of interest.

Entries related to basic principles and concepts include the following:

- Fields of study to illustrate the connections between the scientific research and the various branches science research theory and design to experimental design and statistical analysis;
- An abstract that provides brief, concrete summary of the topic and how the entry is organized;
- Principal terms that are fundamental to the discussion and to understanding the concepts presented;
- Basic principles that clarify the essentials of the topic
- Text that gives an explanation of the principles and its importance to scientific research, including theory and practice, benefits and drawbacks, and practical applications;
- Formulas and equations related to the principle;
- Illustrations that clarify difficult concepts via models, diagrams, and charts of such key topics as longitudinal sampling, nested designs, and probabilistic sampling;
- Further reading lists that relate to the entry.

Entries related to important figures in scientific research include the following:

- A brief overview of the individual and their contributions;
- Key dates and biographical data;
- Primary field(s) and specialties;
- Sidebars explaining their significant advances, inventions, or discoveries;
- Text that provides information about the scientist's Early Life, Life's Work, and Impact;
- Further reading lists that relate to the entry.

This reference work begins with a comprehensive introduction to the field, written by editor Donald E. Franceschetti, professor emeritus at the University of Memphis.

The book includes helpful appendixes as another valuable resource, including the following:

- Time Line of Inventions and Scientific Advancements
- Glossary;
- Bibliography; and
- Subject Index.

Salem Press and Grey House Publishing extend their appreciation to all involved in the development and production of this work. The entries have been written by experts in the field. Their names and affiliations follow the Editor's Introduction.

Principles of Scientific Research, as well as all Salem Press reference books, is available in print and as an e-book. Please visit www.salempress.com for more information.

EDITOR'S INTRODUCTION

FROM THE *NEW ATLANTIS* TO WHAT LIES AHEAD

Many people in modern society would describe their main activity in life as research. The Random House Collegiate Dictionary defines research as "a diligent and systematic inquiry into a subject in order to discover or revise facts, theories," etc. In this sense research is an activity that is characteristic of modern western societies. It has existed for a long time. More recently, emphasis has been placed on research-based methods, in science, in teaching, in legislation, in economic planning and so on. Exactly what is meant by the modern emphasis on research?

RESEARCH AND THE SCIENTIFIC METHOD

The world changed in the seventeenth century. The ancient Greeks and Romans had refined deductive logic to a precise science. In deductive logic one sets out to draw all the conclusions inherent in a set of premises. By far the most impressive success was found in the geometry of Euclid (325 BCE- ?) and some of the philosophical tracts of Aristotle (384 BCE - 322 BCE). Euclid mainly collected results known to the Babylonians and other inhabitants of the Ancient Neat East. His great contributions lay in systematizing results of others so that they followed from a simple set of axioms and postulates. Aristotle was a collector of facts; he a great was also a great organizer, who writings on a great variety of subjects have been preserved for us. His teaching on deductive logic has come down to us as the *Organon*, and had a great influence on philosophy.

The western world put philosophy at the service of religion. The rediscovery of Aristotle's writings presented the now monotheistic world with a number of challenges. Jewish, Christian, and Moslem scholars all adapted the thought of the pagan Aristotle to their own religious systems.

As the Dark Ages ended Aristotle was considered the greatest of philosophers and his ideas could not be brought into question.

The Greeks considered the celestial realm to be perfect but the terrestrial realm was far from perfect. The Greeks had applied mathematics to the motions of celestial objects. The orbits of the planets were perfect circles, because circles were perfect. When it became apparent that the planetary orbits were not perfect circles, a complicated system of epicycles was introduced. The need for a more robust methodology began to be felt. The polish astronomer Copernicus (1473-1543) introduced the heliocentric solar system and doubts began to grow that the earth upon which men walked was the center of the universe.

The hypothetico-deductive method is a variant of deductive logic used particularly by some physicists. A good example is found in the *Philosophiæ Naturalis Principia Mathematica* of Isaac Newton (1642-1727). Newton's famous dictum: *Hypothesis non fingo* or "I do not make hypotheses," turned physics away from physical speculation and toward a precise mathematical description of what occurs leaving the "why" an open question.

Earnest Rutherford (1871-1937) has quipped that science was either Physics or stamp collecting. The stamp collecting side is represented by Carol Linnaeus (1707-1778), who was the first biologist to classify the members of the animal and plant kingdoms. Grouped with Linnaeus must be the early Thomas Edison (1837-1931) the American inventor who found a workable filament material by a process of trial and error. The trial and error method does work sometimes, but does not provide insight as to what might be the most promising theory. A small amount of physical theory, if available, could have greatly simplified Edison's search.

Rutherford's comment became the basis of a classification scheme proposed in a highly influential book entitled *Pasteur's Quadrant: Basic Science and Technological Innovation*, written by Donald E Stokes and published in 1997. Stoke suggested that the prevalent classification of research as pure or applied was misleading and that scientific activity could be better plotted in two dimensions. Along one axis based on whether considerations of use motivated it, and a perpendicular axis by whether it was driven by a quest for fundamental knowledge. In this scheme traditional pure basic research occupied one quadrant. The work of scientists such as Louis Pasteur, which was almost always driven by the need to treat one disease or another, fell into a separate quadrant while the work of Thomas Edison occupied the non-theory driven quadrant.

SIR FRANCIS BACON

The greatest propagandist for the inductive method was Sir Francis Bacon (1561-1626), contemporary of Galileo (1564-1642) and Newton (1642-1727). Unlike Galileo but like Newton, he was a Protestant, a prolific writer, and a man who aspired to a position of influence. Like Newton he attached far more importance to his social status and his role in government than to his fundamental contributions to science or philosophy. Unlike Newton he was born to the nobility but the early death of his father greatly reduced his financial prospects. He entered the service of Queen Elizabeth I, but the Queen did not favor him. He fared much better under her son, King James I, rising through various administrative posts, becoming Lord Chancellor of England and being raised to the peerage. In the process he made a number of influential enemies. Bacon was forced to resign as Lord Chancellor, but allowed to retain his titles as a peer of the realm: Baron Verulam and Viscount St. Albans. He died shortly thereafter having caught a chill during an early experiment with frozen food.

Bacon's value as a scientist is open to question as he seems unaware the scientific achievements of his time including William Gilbert's modeling of terrestrial magnetism or with Sir William Harvey's discovery of the circulation of blood. Nonetheless his written books—*The Advancement and Proficience of Learning Divine and Humane* (1605), *Novum Organum Scientarium* (1620) and *New Atlantis* (1627) gave voice to the methods of inductive logic, so that it is generally referred to as Baconian induction even though Bacon's methods have been replaced by more modern ones. *New Atlantis* is an incomplete utopian novel that describes a mythical island where a great many citizens are involved in scientific research. This reflects the growing involvement in professional societies and scientific journals. A century later, the mode of existence on the island was parodied by Jonathan Swift in the third book of *Gulliver's Travels.* While some research papers are written in a manner that invites parody, the image of a culture based on scientific research appealed to a great many people. A great many historians cite Bacon's work as a turning point in scientific research.

Immanuel Kant, possibly the greatest philosopher of modern times, includes a motto from Bacon's preface to *Instauratio Magna* in his work, *Critique of Pure Reason.* More recently, Loren Eiseley, popular writer and anthropologist, described Bacon in *Francis Bacon* and the Modern Dilemma (1962) as "the man who saw through time."

Bacon's time saw the emergence of the first scientific societies in Europe; The Accademia del Cimento, to which Galileo belonged, and the Royal Society of London. These learned societies sponsored the first scientific journals and provided national forums where the best scientists could meet to discuss and debate their work. They also sponsored scientific research which until that time, had been the prerogative of royalty.

As the Industrial Revolution (1760–1840) progressed, it became the case that ordinary citizens were capable of amassing fortunes. Governments found this accumulation of wealth to be irresistible and so began to impose inheritance taxes so that wealth could not pass down untaxed from one generation to the next. Wealthy individuals learned that nonprofit foundations were one way to preserve their interests when they could not pass their money directly to their heirs. Industrialists often found their interests best-served by scientific research, leading to the establishment of such foundations as the Rockefeller Foundation, the Ford Foundation, and more recently the Bill and Malinda Gates Foundation.

The way in which science and wealth interacted to some extent reflected national style. In England, the best universities, originally established to serve the nobility, witnessed the growth of new institutions meant to serve those students who were disinclined to devote years to the study of Latin and Greek. In Germany, organic chemistry would flourish. The popular notion of the scientist toiling away, night after night on his latest invention to serve no higher purpose then his own pure curiosity is a trope that actually derives from nineteenth-century German romanticism rather than out of the traditions of the universities of the Middle Ages.

SCIENCE AND INVENTION: EDISON'S INVENTION FACTORY, BELL TELEPHONE LABS

The founding fathers of the American nation made provision in Article 8 of the United States Constitution for the new National Government to issue letters patent which granted inventors monopoly rights to their inventions for a limited number of years.

In doing so they were following the example of the British Statute on Monopolies, adopted in England in 1623. Over the centuries this law became very important in the advancement of technology as barriers to information exchange were lowered.

Thomas Edison was a largely self-taught inventor who accumulated over 1000 patents in his lifetime. In 1876, he established a research laboratory which became known as Edison's invention factory.

In the same year, 1876, the Scottish-educated engineer, Alexander Graham Bell invented the telephone. Bell Telephone Labs was established in 1925. Its technical staff would be responsible for numerous patents, also.

As time went on industrial research organizations acquired a secondary mission: filing patents claims in broad areas of interest to the parent companies, and as a result a novel technology often involves the use of patents owned by the parent organization. Before the technology can be brought to market, developers would have to acquire a license to use the patent. Failure to secure the appropriate license can stall or halt development. Today, particularly in the medical realm, a new product may list several hundred items of prior art, requiring the services of specialist in intellectual property law, before it hits the market. In the United States, and most industrialized countries, intellectual property law has become a major area of legal practice, with a separate bar exam, additional educational requirements, and increasingly subject to international treaties as well as federal law.

EUROPEAN EDUCATIONAL INSTITUTIONS AND AMERICAN PRAGMATISM

During the Middle Ages, European universities generally took pride in their traditions, many having outlasted several national governments. Universities were administered by their faculties, or by the church, and often enjoyed a measure of extraterritoriality. Bound together by the church and the Latin language they provided a refuge for scholars during the later Middle Ages and into the Renaissance and early Modern Era.

The university system had to adapt to the emergence of the sciences and fields such as engineering. The system of academic degrees: bachelor's, master's, and doctoral was basically retained but the number of subjects in which one could earn a degree was greatly expanded. The once universal use of Latin was discontinued.

The discovery of the New World and the protestant reformation brought many new styles to higher education.

In 1861 Yale University became the first American institution to award the Doctor of Philosophy degree. Until then Americans seeking an advanced degree had to travel to Europe to complete their education.

The main source of training for American scientists remained the British and German universities, until the Second World War.

SCIENCE BUILDS THE BOMB

Perhaps no period in the recent history of science has left as great a mark on the organization of science as the Manhattan Project. The story has the elements of high-level science as well as high drama: A group of European scientists, transplanted to the United States as political refugees, realize at the beginnings of the Second World War that it was possible to build an extremely powerful weapon based on discoveries made less than ten years earlier. At the same time, powerful and ruthless enemies had already begun their evil work.

President Franklin Roosevelt makes the decision to build a bomb at the imploring of Albert Einstein, a European refugee fleeing the drumbeat of anti-Semitism spreading throughout Europe as well as (ironically, as it turns out) a dedicated pacifist. The project is confronted with both practical and ethical conundrums–The bomb may not work. If it does work, it might result in the deaths of a hundred thousand civilians. The bomb might work too well and eliminate all human life on earth. A European-trained leader is selected for the project. The team leader, J. Robert Oppenheimer, has a politically questionable past. But in the opinion of General Groves, a key Army officer, Oppenheimer is one of the few men who could make the project work. Is it ethical to exempt so many men at a time of military crisis? Is it ethical to plan on the deaths of so many thousands to bring the war to a conclusion?

The American-led effort was successful and the bombs constructed but the decision to use the bombs fell to Harry Truman, who assumed office on Roosevelt's death and who as vice president had no knowledge of the bomb's existence

Historians and philosophers argue about the bombs' use to this day. Did it save more lives than it destroyed? Did it shorten the course of the war? Did it make the world a safer or more dangerous place? The threat of nuclear war has become an accepted fact the post-WWII world, although the United States' bombing of Japan is the only nuclear act of war to date.

Following the War, the United States, geographically isolated from Europe and Asia, found itself engaged in a debated concerning whether or not it should assume a leadership role or return to isolationism. The Allied nations, left so decimated by WWII, faced an immense rebuilding task. They worried that if the United States abandoned its leadership role, other nations might fall sway to the allure of communism.

Vannevar Bush and the Endless Frontier

The question of what to do with the research facilities built in the United States during the Second World War and dedicated to developing nuclear weapons came up even before the war ended. Franklin Roosevelt asked Vannevar Bush, head of wartime research and development, for suggestions. His findings were reported in the document Science: The Endless Frontier, which described the rewards of investing in research in glowing terms. The book is a bit painful to read now, its optimism reminiscent of the *New Atlantis*. It recommended the establishment of a National Science Foundation to fund basic research in the sciences. It further recommended that the National Science Foundation and other governmental agencies charged with the conduct of research should use the available university faculties to conduct the research and train the next generation of researchers, with the government bearing the full cost of the research. A great many universities embraced the ideal of the research university. Unfortunately, there was the hidden assumption that the exponential growth of the scientific establishment could continue indefinitely. It did from the early years under President Eisenhower until America achieved President Kennedy's dream of landing men on the Earth's moon in 1969.

By the mid-1970s the situation had changed markedly. University science and engineering departments and faculties had grown, while there were fewer American citizens with interest in science and engineering.

Developments since the Second World War

To understand how research has developed since the Second World War we will consider the advances made in communications and computer science. One could argue that the current state of affairs comes about as the result of four critical discoveries. The first was thermionic emission from a heated metal, an observation made by Thomas Edison himself in 1883. Edison could think of no practical application of the effect, so he dutifully recorded his observations in a notebook and went on to other things. Fortunately for electrical science, he was visited by the British Sir John Ambrose Fleming in 1884. In 1889, Fleming began a collaboration with Italian inventor Gugliemo Marconi. In the process, he obtained a patent for vacuum tube rectifier. A few years later the American Lee De Forest added a third element to make a vacuum tube triode. Advances in vacuum tube design and applications certainly fell in Pasteur's Quadrant. Entire new fields of enquiry including many aspects of nuclear and particle physics were opened by the ability to achieve a high vacuum.

The second critical invention has been the digital computer and has its origins in the purely deductive mathematical logic of Aristotle. The mathematical question is subtle: Is there an effective procedure to calculate a number of interest? By the nineteenth century, mathematicians had begun to wonder if all the theorems of mathematics could be derived from a single set of postulates. To answer this question Alan Turing had to refine the notion of an effective procedure. He reduced each effective procedure to a set of instructions for a so-called Turing machine, a machine which read numbers or symbols from a tape, compared the symbol just read, with a symbol stored in a one slot memory, and based on the symbol would write the result and move the tape forward or backward. Turing showed that it was possible to formulate a universal Turing machine that could read in a description of any Turing machine and then emulate it. The surprising results of Turing's analysis were that it was generally impossible to feed the description of a Turing machine into another to determine that the first machine would come to a halt. The important practical result was the design, at least in concept, of a programmable digital computer. The first computers were actually humans who could do calculations by hand with speed and accuracy. During the Second World War these were augmented by vacuum tube devices which, if large, frequently failed due to filament burn out. Electronic computers would make a great leap in power and reliability with the invention, in 1947, of the transistor.

The third major development was the transistor, a substitute for the vacuum tube triode, which was far more reliable and could be miniaturized. Once the basic theory was understood the device quickly moved from the quest for fundamental understanding to application driven research.

The final stage in the evolution of computer and communication technology was that of miniaturization. Perhaps no one of Edison's time could have envisioned that using photolithography on slices of ultra- refined silicon one could manufacture integrated circuits with thousands of transistors per square centimeter and do so cheaply enough that a calculator or cell telephone could become a throwaway item. The giants of high-tech manufacturing yesterday; Oppenheimer and the polymath John von Neuman, who did much of the engineering on the first computers, would have to move over to accommodate teenage college drop-outs, like Bill Gates and Steve Jobs, who not incidentally listed J. Robert Oppenheimer as one of his heroes.

THE CHANGING FRONTIER?

The digital computer, with its birth in pure theory and its development in wartime necessity, is now ubiquitous. Computers can analyze data at very high speeds. Computers can solve some problems, actually a relatively small fraction, but can always be applied to simulate a complex system. Computing still represents an open frontier.

There are other obvious frontiers to be explored. Research areas once cloaked in military secrecy are now open to private enterprise. The space race has given way to cooperative missions. The aspirations of the physics community in the United States, to build a superconducting super collider were scrapped, and the United States must join with European nations to conduct particle physics research. For the moment, the ultimate consideration in how quickly research gets done will be economic. The players may change but scientific discovery remains an endless frontier.

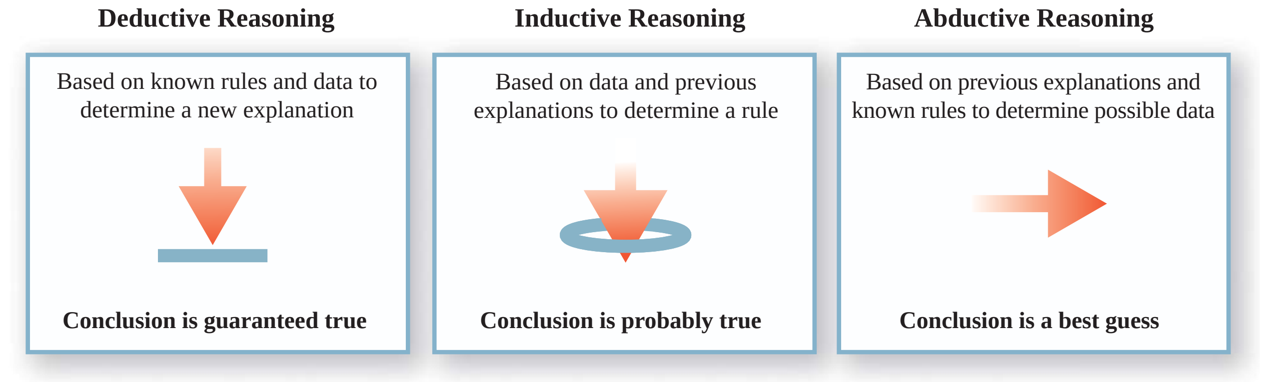
FROM INDUCTION TO ABDUCTION

One more trend is worthy of mention. In *The Rise of Statistical Thinking, 1820-1900*, Theodore M. Porter shows how the idea of uncertainty has come to permeate the sciences, both in the laboratory and in the social realm. With the development of the computer, there has been a realization that the majority of physical problems have no exact solution, and further, that for most purposes a highly accurate solution may not be worth the required effort to obtain it. Thus, the search for accurate solutions has given way to a search for solutions which are, in the words of artificial intelligence pioneer, and Nobel Laureate Herbert Simon, are good enough. Coming to terms with this reality is one of the major challenges facing researchers in all fields.

—Donald R. Franceschetti

A

ABDUCTIVE REASONING



FIELDS OF STUDY

Hypothesis Testing; Research Theory

ABSTRACT

Abductive reasoning is the form of logical inference that is used when available information is incomplete. A hypothesis is developed to explain some observed phenomenon as best as possible. Though less certain than other forms of reasoning, it is valuable and frequently used in both science and everyday life.

PRINCIPAL TERMS

- **abductive logic programming:** an extension of logic programming that permits some predicates to have undefined values rather than definite ones.
- **deductive reasoning:** a type of reasoning in which general rules are applied to a specific situation in order to draw a narrower conclusion; if the premises on which the conclusion is based are true, the conclusion is also deemed to be true.
- **inductive reasoning:** a type of reasoning in which knowledge of specific situations is used to draw a broader and more general conclusion; if the premises on which the conclusion is based are true, the conclusion is deemed to be highly probable but not certain.

ABDUCTION, DEDUCTION, AND INDUCTION

Abductive reasoning, or abduction, is a logical process of inference in which incomplete data is used to draw the simplest and most likely possible conclusion. It is an important concept in fields ranging from law to computer science and is often used by people in everyday life. Yet it is often poorly understood, especially compared with the better-known deductive reasoning and inductive reasoning. While these types of logical inference rely on a systematic building up of premises that lead to a certain or probable conclusion, abduction involves educated guessing based on the best available observations or facts.

In the context of a scientific experiment, deductive and inductive reasoning may be used to analyze facts or phenomena generated by the experiment. Abductive reasoning comes before this, when the researcher forms the hypothesis that the experiment is designed to test. It is therefore a crucial part of the research design process. Abduction bridges the gap between what is already known about the universe, such as natural laws and other forms of information that can be relied on or assumed to be true, and new information that a researcher wishes to obtain. In other words, deductive and inductive reasoning are best used to analyze the results of an experiment, while abductive reasoning is best used to devise a possible explanation for the phenomenon being studied and determine how best to gather data to test that explanation.

ABDUCTION IN RESEARCH

Abductive reasoning is a crucial part of research design that all researchers should be familiar with. However, while abductive reasoning always makes do with incomplete data, it is important to recognize that this does not condone ill-informed research or total guesswork. Errors in research are common when researchers do not have enough information about the context of the phenomenon they are studying. This limits the options available when forming a hypothesis about the most likely explanation for an observed event. Researchers must thoroughly investigate all of the potential factors that may be at work in a given scenario before designing an experiment to test an explanation for that scenario. When used correctly, abductive reasoning is a valuable tool for both scientific inquiry and everyday decision making.

ABDUCTION IN PRACTICE

A classic example of abductive reasoning is in a criminal trial. The jury must come to a conclusion based on presented evidence, even if they do not have access to all of the relevant information. Another common example is medical diagnosis. The complexity of health factors on both general and individual levels means that full information regarding a patient is often unavailable. A doctor typically asks questions of the patient to gather as much data as possible, but the patient's complaints may not align exactly with examples in medical literature, or the patient may be totally unresponsive. In such cases the doctor must make an educated guess about the patient's condition based on symptoms and any other known factors.

For example, a patient with a skin irritation might be asked whether they have been exposed to any chemicals or bitten by an insect or snake. The answers, along with the doctor's knowledge, will help narrow down the likely cause of the rash. However, even if the patient says they were bitten by a snake, they may not know what kind of snake it was. If only rattlesnakes live in the area and the patient's symptoms are consistent with a rattlesnake bite, the doctor can safely guess that the patient was bitten by a rattlesnake and provide the appropriate treatment. It is still logically possible that a different kind of snake is to blame, or even that the bite was from a nonvenomous snake and the rash is completely unrelated. A rattlesnake bite is simply the most likely explanation under the circumstances.

ARTIFICIAL INTELLIGENCE

Abductive reasoning has become especially relevant in the field of computer science, particularly with regard to abductive logic programming (ALP). Research into developing artificial intelligence (AI) has shown that guessing or estimating is one of the most difficult tasks to program a computer to perform. For a computer to make such a logical leap, it must evaluate and assign relative importance to multiple rules that have been defined to help guide it. This process is meant to approximate how humans use their awareness of the rules that govern the world around them—water is wet, gravity pulls objects toward the earth, and so on—to conceive of possible explanations for observed phenomena.

The challenge for AI programmers is that there are so many rules that human beings take for granted, it is difficult to define them all so that a computer can refer to them. Even after this is done, programmers must then devise a mechanism that the computer can use to decide which rules are more likely to account for a specific phenomenon. Still, when implemented successfully, ALP allows intelligent systems to better determine and correct faults, plan and make decisions, and revise themselves.

—Scott Zimmer, JD

FURTHER READING

Dunne, Danielle D., and Deborah Dougherty. "Abductive Reasoning: How Innovators Navigate in the Labyrinth of Complex Product Innovation." *Organization Studies*, vol. 37, no. 2, 2016, pp. 131–59.

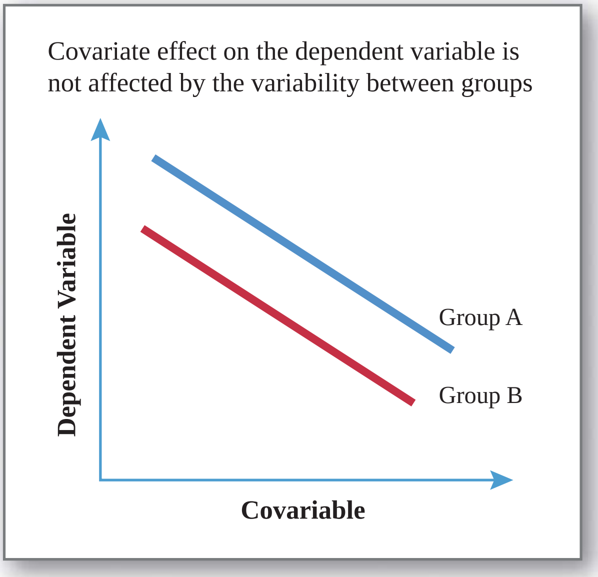
Holyoak, Keith J., and Robert G. Morrison, editors. *The Oxford Handbook of Thinking and Reasoning*. Oxford UP, 2012.

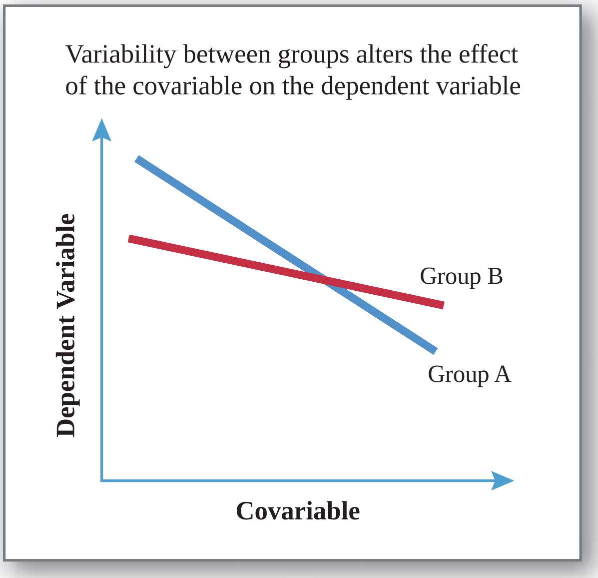
Mirza, Noeman A., et al. "A Concept Analysis of Abductive Reasoning." *Journal of Advanced Nursing*, vol. 70, no. 9, 2014, pp. 1980–94. *Academic Search Complete*, search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=97431963&site=ehost-live. Accessed 9 Mar. 2017.

Tavory, Iddo, and Stefan Timmermans. *Abductive Analysis: Theorizing Qualitative Research*. U of Chicago P, 2014.

Walton, Douglas. *Abductive Reasoning*. 2005. U of Alabama P, 2014.

ANCOVA





FIELDS OF STUDY

Statistical Analysis; Research Design; Experimental Design

ABSTRACT

Analysis of covariance (ANCOVA) is a statistical method used to adjust group means in order to control for the effects of variables other than the ones being studied. It depends on the use of pretests to establish a baseline that can then be taken into account when analyzing observations.

PRINCIPAL TERMS

- **covariate:** a continuous independent variable other than the one being manipulated that may influence, predict, or explain the effects of the manipulated variable on the dependent variable.
- **interaction:** in statistics, an effect by which the impact of an independent variable on the dependent variable is altered, in a nonadditive way, by the influence of another independent variable or variables.
- **posttest:** an examination that takes place after a treatment or other experimental intervention.
- **pretest:** an examination that takes place before a treatment or other experimental intervention.
- **treatment:** the process or intervention applied or administered to members of an experimental group.

ANCOVA Defined and Explained

Analysis of covariance (ANCOVA, sometimes ANACOVA or ANOCOVA) is used to analyze and control for the effects of covariates among multiple sample groups in a research study. It combines elements of two other statistical tests, regression and analysis of variance (ANOVA). Regression is used to determine whether and how certain variables are related. ANOVA is used to analyze differences in the means of different sample groups. Accordingly, ANCOVA is used to analyze differences in means of different sample groups after controlling for the effects of any other independent variables that might influence the dependent variable. It is a method of statistical control, rather than experimental design control.

A researcher who plans to use ANCOVA must first identify one or more potential covariates in their study. There are several criteria for selecting effective covariates. First, there must be a linear relationship between the chosen covariate and the dependent variable and that relationship must be the same among all sample groups. Second, the covariate must not be influenced by any other independent variables. If there is any interaction between the covariate and the independent variable being studied, then the covariate cannot be controlled for without significantly altering the results. Third, the covariate must be a continuous, not categorical, variable. In other words, it must be numerically measurable in intervals. If sufficient measurements of multiple potential covariates are taken at the outset, the researcher can choose which best fits the criteria later.

ANCOVA is most often used in pretest-posttest experimental designs. A pretest is conducted before the study to measure the chosen covariate(s). In a clinical trial to determine which dosage of a drug to treat high blood pressure is most effective, for example, subjects might be divided into three sample groups: one receiving 0.5 milligram of the drug, one receiving 1.0 milligram, and one receiving 1.5 milligrams. The independent variable would be the dosage given to each sample group, and the covariate might be each patient's baseline blood pressure. Then, after the treatment has been administered, a posttest is used to measure the dependent variable. In the clinical trial example, the dependent variable would be the patient's post-treatment blood pressure. Given the value of the independent variable, dependent variable, and covariate for each subject, the researcher can conduct ANCOVA to adjust the means for each group to reflect what the results would be if all subjects had the same pretreatment blood pressure. Using the adjusted means, they can then determine whether the null hypothesis should be rejected. In ANCOVA, the null hypothesis states that there is no significant difference among the means of the sample groups after the means have been adjusted to control for the covariate.

Advantages and Disadvantages

ANCOVA, like ANOVA, is governed by certain assumptions. The following assumptions apply to both:

- Residuals (that is, the difference between an observed value in a group and the group mean) in each sample group are normally distributed.
- Variance is homogenous among sample groups.
- Sample groups are fully independent.

Furthermore, ANCOVA also assumes that there is a linear relationship between the covariate(s) and the dependent variable, that the slopes of the lines expressing this relationship are equal among all sample groups, and that the covariate is measured exactly. If any of these assumptions are false, ANCOVA may produce incorrect or misleading results. Moreover, the effects of violating these assumptions are not always clear, which can cause problems during data analysis.

ANCOVA in Practice

The validity of ANCOVA has continued to be upheld in a number of studies. A 2017 study by Jennifer L. Reeves and colleagues used ANCOVA to test the impact of using iPads in prekindergarten (pre-K) classrooms, for example. The scores from the school year's first pre-K assessment were the covariate, and scores from the third assessment were the dependent variable. With that data, the study found that students who used iPads for learning achieved significantly better results in math and phonological awareness than those who did not.

Like ANOVA, ANCOVA can be calculated using sums of squares. With one-way ANCOVA, there are three variation sources: the sum of squares between groups (or "treatment sum of squares"); the sum of squares within groups (or "error sum of squares" or "residual sum of squares"); and the covariance sum of squares. However, because ANCOVA involves adjusting means—and, by extension, adjusting the sums of squares—the calculations involved are far more complex than in ANOVA. As a result, these calculations are most often performed using statistical software. Once the sums of squares are adjusted, they can be used to calculate an F statistic, as in ANOVA, in order to compare the adjusted means.

Significance

Like ANOVA, ANCOVA can be adapted for use in experimental, quasi-experimental, and observational research designs. Both are also used across a wide range of fields, making it important to understand their assumptions, techniques, and use. They are widely available in popular software packages and online applications.

—Elizabeth Rholetter Purdy, PhD

Further Reading

Huitema, Bradley E. *The Analysis of Covariance and Alternatives: Statistical Methods for Experiments, Quasi-Experiments, and Single-Case Studies.* 2nd ed., John Wiley & Sons, 2011.

Jennings, Megan A., and Robert A. Cribbie. "Comparing Pre-Post Change across Groups: Guidelines for Choosing between Difference Scores, ANCOVA, and Residual Change Scores." *Journal of Data Science*, vol. 14, no. 2, 2016, pp. 205–29. *Academic Search Complete*, search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=116345929&ehost-live. Accessed 4 Apr. 2017.

Lai, Keke, and Ken Kelley. "Accuracy in Parameter Estimation for ANCOVA and ANOVA Contrasts: Sample Size Planning via Narrow Confidence Intervals." *British Journal of Mathematical and Statistical Psychology*, vol. 65, no. 2, 2012, pp. 350–70. *Academic Search Complete*, search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=74280439&ehost-live. Accessed 4 Apr. 2017.

Randolph, Karen A., and Laura L. Myers. *Basic Statistics in Multivariate Analysis*. Oxford UP, 2013.

Reeves, Jennifer L., et al. "Mobile Learning in Pre-kindergarten: Using Student Feedback to Inform Practice." *Journal of Educational Technology and Society*, vol. 20, no. 1, 2017, pp. 37–44. *Academic Search Complete*, search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=120706100&ehost-live. Accessed 4 Apr. 2017.

Rutherford, Andrew. *ANOVA and ANCOVA: A GLM Approach*. 2nd ed., John Wiley & Sons, 2011.

Trochim, William M. K. "Covariance Designs." *Research Methods Knowledge Base*, 20 Oct. 2006, www.socialresearchmethods.net/kb/expcov.php. Accessed 4 Apr. 2017.

Warner, Rebecca M. *Applied Statistics: From Bivariate through Multivariate Techniques*. 2nd ed., SAGE Publications, 2013.

ANCOVA SAMPLE PROBLEM

Two professors are teaching the same introductory college course in American government. Some students have taken previous classes in American government, but others have not. At the end of the semester, both classes will take the same final exam. A researcher wants to conduct a pretest-posttest study to find out which professor does a better job of teaching the material that will be on the exam. In this study, what would be the independent and dependent variables, and what should the researcher use as a covariate?

Answer:

Each class represents a sample group, so the independent variable is which professor is teaching the class. The dependent variable will be the students' scores on the final exam. Because some students have studied the subject before, the researcher wants to control for prior knowledge. To do this, the researcher will ask both classes to take a pretest at the beginning of the semester that will cover the same range of material as the final exam. The covariate will be the students' scores on this pretest.

ANOVA

FIELDS OF STUDY

Statistical Analysis

ABSTRACT

Analysis of variance (ANOVA) is a method of analyzing the differences between the means of multiple data sets and testing the validity of those differences. It is generally considered the most popular statistical test.

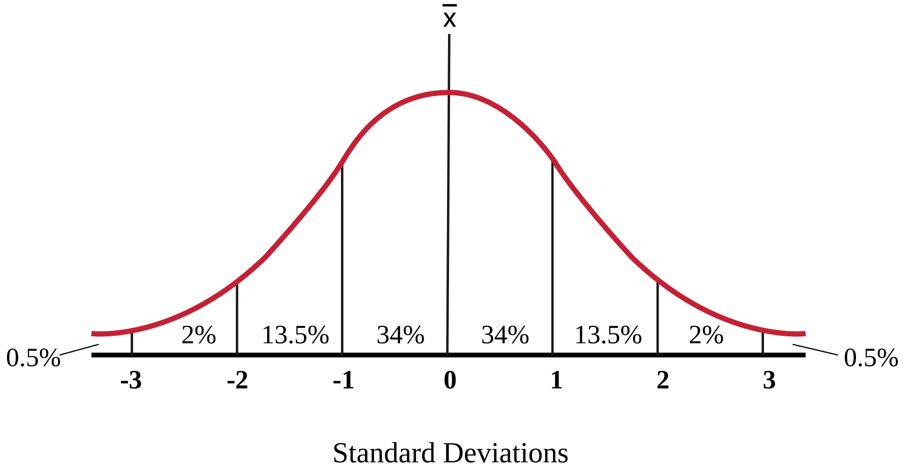
PRINCIPAL TERMS

- **normal distribution:** a probability or frequency distribution in which plotting the values contained in a data set, according to either the probability of their occurring or the frequency with which they occur, results in the appearance of a symmetrical bell-shaped curve, with the majority of values clustered around the middle; also called a bell curve.
- **sum of squares between:** the sum of the squares of the deviation of each group mean from the grand (overall) mean, multiplied by the number of data points in each group; also called the treatment sum of squares.
- **sum of squares within:** the overall sum of the sums of the squares of the deviation of each data point within a group from the mean of that group; also called the error sum of squares or the residual sum of squares.
- **variance:** a measure of how widely data points within a group are dispersed from the group mean, expressed as the average squared distance of each data point from the mean.

ANOVA Defined and Explained

Analysis of variance (ANOVA) is a common method for analyzing differences among the means of multiple data sets. ANOVA is used to determine if two or more groups are statistically distinct by testing how individual data points vary from group means and from the overall mean. In ANOVA, the null hypothesis is that there are no differences among the means of each group except for those that randomly occur by chance. If the researcher disproves the null hypothesis, then the groups are statistically distinct and that distinction must be explained. ANOVA does not identify any differences that exist among groups; it only shows that they exist.

ANOVA tests may be either one-way or two-way. One-way ANOVA is used to test the impact of one independent variable on two or more groups. Two-way ANOVA tests two or more independent variables. There are two main ANOVA models: fixed effects and random effects. The fixed-effects model is used when the values of the independent variable(s) being tested are deliberately selected by the researchers, either because those are the only possible values or



they are the only values of interest. For example, a clinical trial that administered different drug dosages to different experimental groups would be analyzed using the fixed-effects model. The random-effects model is used when the values of the independent variable(s) being tested are not chosen by the researchers, typically because subjects were randomly selected from a larger population and the variable(s) being studied are preexisting in that population. For example, a study of the blood-pressure levels of randomly selected individuals would be analyzed using the random-effects model.

Advantages and Disadvantages

Three main assumptions govern the use of ANOVA. The first is that the residuals—that is, the difference between an observed value in a group and the group mean—of each sample group follow a normal distribution. Second, variance (the square of the standard deviation) is homogenous among the sample groups being tested. Third, the sample groups from which the data are gathered are truly independent of one another. While these assumptions must apply in most cases, in a randomized controlled experiment, only the homogeneity of variance is assumed.

ANOVA is often compared to its predecessor, the Student *t*-test. The *t*-test was also designed to compare mean scores of groups and shares the assumptions of normal distributions and independence of groups. Since the *t*-test performs a group-by-group comparison, it is better than ANOVA at identifying details about group differences. However, a single *t*-test can only test the differences between two groups, and running multiple *t* tests greatly increases the risk of type I errors. In general, a *t*-test is preferable when comparing two groups, and ANOVA is the tool of choice for comparisons of three or more groups.

ANOVA in Practice

The validity of ANOVA depends on the findings of the *F* test, which tests the equality of variances. To conduct the *F* test, one must determine the sum of squares between groups (SS_B) and the sum of squares within groups (SS_W). SS_B is a measure of the variability between the means of different sample groups and the grand (overall) mean. It is also called the "treatment sum of squares" because this variation is attributable to differences in the treatment or other factor being tested. SS_W is a measure of the variability between data points in each sample group and the mean of that group. It is also called the "error sum of squares" because this variation is attributable to random error. (Adding SS_B and SS_W together gives the total sum of squares, or SS_T, but this is not necessary for ANOVA.)

Next, SS_B and SS_W must each be divided by their respective degrees of freedom. SS_B has $k - 1$ degrees of freedom, where k is the total number of sample groups. SS_W has $N - k$ degrees of freedom, where N is the total number of data points from all sample groups. This calculation produces the mean square between groups (MS_B) and the mean square within groups (MS_W), respectively.

Finally, MS_B is divided by MS_W. The resulting value is called the *F* statistic or *F* ratio. An *F* distribution table gives the *F* critical values based on the degrees of freedom in the numerator ($k - 1$) and in the denominator ($N - k$). If the *F* ratio is greater than the *F* critical value, the null hypothesis is rejected. Recall that the null hypothesis in ANOVA is that there are no statistically significant differences among the means of each group. An *F* ratio above the critical value indicates that the differences are, in fact, statistically significant.

Significance

ANOVA is adaptable to a wide range of fields, including medicine, natural science, social science, and business. It may be used in experimental, quasi-experimental, and observational studies. For these reasons, it has become the most popular statistical method of comparing groups by testing the validity of samples taken from overall populations. Assistance in performing ANOVA is also easily accessible, since it is a common function of most statistical software applications. Researchers should keep in mind that ANOVA does not explain any differences that occur but only shows whether differences exist.

—Elizabeth Rholetter Purdy, PhD

FURTHER READING

Cardinal, Rudolf N., and Michael R. F. Aitken. *ANOVA for the Behavioural Sciences Researcher*. Lawrence Erlbaum Associates, 2006.

Madrigal, Lorena. *Statistics for Anthropology*. 2nd ed., Cambridge UP, 2012.

"One-Way ANOVA." *Laerd Statistics*, Lund Research, 2013, statistics.laerd.com/statistical-guides/one-way-anova-statistical-guide.php. Accessed 4 Apr. 2017.

Plansky, M. "Analysis of Variance—One-Way." *Psychological Statistics*, U of Wisconsin–Stevens Point, 1997–2016, www4.uwsp.edu/psych/stat/12/anova-1w.htm. Accessed 4 Apr. 2017.

Randolph, Karen A., and Laura L. Myers. *Basic Statistics in Multivariate Analysis*. Oxford UP, 2013.

Rutherford, Andrew. *ANOVA and ANCOVA: A GLM Approach*. 2nd ed., John Wiley & Sons, 2011.

Tarlow, Kevin R. "Teaching Principles of Inference with ANOVA." *Teaching Statistics*, vol. 38, no. 1, 2016, pp. 16–21. *Academic Search Complete*, search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=112335846&site=ehost-live. Accessed 4 Apr. 2017.

Walker, Jeffrey T., and Sean Maddan. *Statistics in Criminology and Criminal Justice*. Jones and Bartlett Publishers, 2012. Print.

Wall Emerson, Robert. "ANOVA and *t*-Tests." *Journal of Visual Impairment & Blindness*, vol. 111, no. 2, 2017, pp. 193–96. *Academic Search Complete*, search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=121669183&site=ehost-live. Accessed 4 Apr. 2017.

ANOVA SAMPLE PROBLEM

A researcher wants to know if there are any significant differences in the length of periwinkle shells gathered from three different beaches. The researcher has twenty-four shells total: eight from Russia, eight from France, and eight from the eastern United States. The sum of squares between (SS_B) is 13.332, and the sum of squares within (SS_W) is 225.183. Using this information, calculate the F ratio, and then use an F distribution table to determine if there are any significant differences among the groups. Assume a P value of 0.05.

Answer:

Substitute the number of sample groups (k) and total number of data points (N) into the formulas for the degrees of freedom for SS_B and SS_W:

$$SS_B\text{: } k - 1 = 3 - 1 = 2$$

$$SS_W\text{: } N - k = 24 - 3 = 21$$

Divide SS_B and SS_W by their respective degrees of freedom to determine the mean square between groups (MS_B) and the mean square within groups (MS_W):

$$MS_B = 13.332 / 2 = 6.666$$

$$MS_W = 225.283 / 21 = 10.723$$

Finally, divide MS_B by MS_W to find the F ratio:

$$F = 13.332 / 225.283 = 0.622$$

The F ratio is 0.662. According to an F distribution table, the F critical value for 2 degrees of freedom in the numerator, 21 degrees of freedom in the denominator, and a P value of 0.05 is 3.47. The F ratio is well below the F critical value. Thus, there are no statistically significant differences between the means of the sample groups, and the null hypothesis is accepted.