R. A. Fisher—The Father of Modern Statistics

By Yundi Kong and Ezra Becker
Stat 522—Spring 2021

Synopsis

Ronald Aylmer Fisher was a British statistician active in the first half of the twentieth century. He is universally considered to be the father of modern statistics. In this paper we discuss the following contributions Fisher made to the discipline.

- Proof of the distribution of the t-statistic using a multidimensional geometric approach
- Expansion of the t-statistic concept to the development of the F-statistic
- Introduction of the concept of *degrees of freedom*
- Development of the *maximum likelihood* framework
- Development of the modern theory of the design of experiments, including *randomization*, *blocking*, *factorial design*, and *ANOVA*
- Development of *discriminant analysis*
- Establishment of p = 0.05 as the benchmark for statistical significance
- Establishment of a formal, rigorous, and connected structure for mathematical statistics

In addition, Ronald Fisher was one of the giants in the field of genetics. We discuss briefly his reconciliation of a fierce debate between two schools of thought in genetics in the 20^{th} century.

We conclude with a discussion of Fisher's advocacy for the pseudoscientific and inherently racist concept of *eugenics*.

Introduction

Ronald Aylmer Fisher (1890-1962) was a British statistician and geneticist. But to call him just that is like describing the world's tallest skyscraper, the Burj Khalifa in Dubai, as being merely an office building; it completely misses the size and significance of the man. Fisher has been called by many the "Father of Modern Statistics," and truly, his ideas and methodology form the core framework of what every statistics student learns today. Our project assignment asks us to describe the connections between Fisher's contributions and the course *Stat 522*, but, in fact, it would be easier to enumerate those aspects of our course that Fisher did *not* influence than those that he did. Hence, we do not propose to discuss here in any exhaustive fashion Fisher's many innovations, because doing so would require several volumes. Instead, we will attempt to briefly describe a handful of the foundational ideas Fisher contributed to the discipline. To attempt anything else would far exceed the length limits we must follow. Of course, Fisher's contributions are all interrelated, and together weave much of the fabric of modern statistics. For the sake of clarity, we present this selection of Fisher's achievements as individual threads; naturally, the reader will see the connections between them.

The best place to start is not with Fisher himself, but with William Sealy Gosset (1876-1937), known to statisticians everywhere under the pseudonym "Student," for Gosset provided the launching pad that allowed Fisher to begin his chain of contributions. In 1908 Gosset published his famous paper[1] on the t-test, which provided what is now the standard approach to estimating the error of a sample mean when the population variance is unknown. As with many tremendous innovations, Gosset's 1908 paper was not initially recognized by the scientific community for the breakthrough that it represented. But Fisher saw the importance of Gosset's work, and built upon it fruitfully in several different ways.

Rigorous mathematical proofs and the use of geometric arguments

One of the interesting aspects of Gosset's paper is that he formulated the t-distribution through comparison of data to known distributions, and while his ideas were on target, he was unable to derive a rigorous proof for the distribution of the t-statistic. In Gosset's words, "The law of formation [of the moment generating function] appears to be a simple one, but I have not seen my

[1] Student 1908, pp. 1-25.

way to a general proof."[2] It was Fisher who found a rigorous proof for the distribution of the t-statistic. It is a famous (and true) story that Gosset did not in fact understand Fisher's proof when Fisher sent it to him, and with no ego in the matter confessed as much to Karl Pearson, at the time considered the leading statistician of his generation[3]. Fisher's proof was based on multidimensional geometry; and while the method of least squares was purportedly used by Gauss circa 1795 and published by Legendre in 1805[4], geometric approaches to developing broad statistical concepts were generally not used before Fisher. It was one of Fisher's many insights that a dataset could be envisaged in multidimensional space, and he used this approach to rigorously derive the distribution of the t-statistic[5]. The impact of Fisher's first correspondence to Gosset reverberates to this day. The connection between statistics and geometry appears often in *Stat 522*, from the ellipsoid nature of a multivariate parameter estimate's confidence region to the interpretation of the directions of principal component vectors, and more. It can be well said that what Fisher started with Gosset became a lifetime of achievement. Just as he led the way in using geometric proofs for statistical concepts, he also set a higher standard for the rigor of proofs in the discipline.

Extension of the t-test Concept

Fisher did not just appreciate Gosset's work on the t-test statistic in and of itself. He recognized that Gosset's t-test for estimating the mean of a single sample could be expanded to test the differences between the means of two samples of different sizes, which was a far more powerful application of the concept[6]. Beyond that, Fisher realized that this form of test statistic—a ratio of two independent random variables under specific circumstances—could be applied to a diverse range of important tests of significance. For example, by applying the t-test concept to evaluate the significance of regression coefficients, Fisher developed the F-distribution[7], which we have used extensively in *Stat 522*, including in tests of significance for ANOVA models as well as multivariate tests using a modified F-distribution under Hotelling's T^2 statistic.

[2] Student 1908, p. 4.
[3] Fisher Box 1978, pp. 84-85.
[4] Stigler 1986, pp. 145-146.
[5] Stigler 2016, pp. 95-96.
[6] Stigler 2016, pp. 96-97
[7] Stigler 2016, p.96.

Degrees of Freedom

Fisher did not correspond with Gosset just to share his proof of the distribution of the t-test statistic. He also wrote to Gosset to point out what Fisher perceived to be (and in fact was) an error in Gosset's derivation of the independence of the sample mean and the sample standard deviation. Specifically, in calculating the sample standard deviation, Gosset followed the accepted practice of the time to divide the sum of the distances from each data point to the mean by the number of data points n. Fisher pointed out that the denominator should actually be n-1 in the case where the sample mean is used instead of the population mean, and thereby demonstrated the significance of the concept of *degrees of freedom* for statistics. In this manner, Fisher gave us the notion that the information contained in a set of data can be distinguished from the data points themselves. He demonstrated that each linear constraint on a given model (such as the calculation of a sample mean) reduced the degrees of freedom available to the statistician by one. He also applied the concept of degrees of freedom to correct the method of application of Karl Pearson's χ^2 test[8], which previously had required special and complicated adjustments dependent upon circumstances, but for which adjustments there was no clear supporting theoretical framework. Fisher's concept of degrees of freedom provided the rigorous theoretical underpinnings for the correct calculation of the critical values of the χ^2 test, the t-test, the F-test, etc. and is a fundamental consideration in all estimation and testing exercises today.

Maximum Likelihood

The idea of *maximum likelihood* is intuitively attractive. It is reasonable to believe that the underlying, unknown parameters of a model are those most likely to produce the data one has at hand in a given sample. This concept was not developed by Fisher; it can be traced back at least as far as the work of Daniel Bernoulli in the late 1770s[9]. But Fisher revolutionized its application. He introduced the modern nomenclature and structure of the likelihood function $L(\theta|X)$[10], which he defined as the probability density function of the parameter θ given the observed data X. The importance and versatility of the likelihood function in modern statistics is difficult to overstate. Fisher's maximum likelihood approach to parameter estimation—the ubiquitous MLE—provided

[8] Fisher 1924, pp. 805-813
[9] Stigler 1986, p. 110.
[10] Stigler 2016, p. 83.

a direct and more powerful way of estimating model parameters than the Method of Moments, which had been the standard before Fisher. It allowed for the rigorous development of likelihood ratio tests, which Jerzy Neyman and Egon Pearson (the son of Karl Pearson) proved to be the uniformly most powerful ("UMP") test for simple hypotheses[11], and which can be extended to be UMP for one-sided composite tests when the likelihood function is monotonic. The attractiveness of the likelihood ratio test has made it hugely popular for the construction of test statistics, and is usually the first choice for statisticians in doing so. Without Fisher's development of maximum likelihood, the progress of statistics would have been delayed by decades.

But Fisher wasn't done there. He also found that, if the proposed model were well-behaved (i.e., was twice-differentiable in the desired region), then one could express all of the information contained in the data in terms of the second derivative of the likelihood function (or its log, for the sake of computational convenience)—what we are familiar with as the *Fisher Information*.[12] The Fisher Information is central to our understanding of variance, and again shows the power of Fisher's recognition that the information contained in a dataset can be distinguished from the data points themselves. It is this Fisher Information that C. R. Rao[13], Harald Cramér and others used to find the theoretical lower bound on the variance estimator for a data sample (interestingly, it so happens they performed this work independently but almost concurrently). It is also not too far a stretch to say that the Fisher Information matrix is the single most important element of optimal design theory.

Design of Experiments—Randomization, Blocking, Factorial Design, ANOVA

Moving on—Fisher's contributions to the field of the design of experiments were nothing short of revolutionary. Prior to Fisher, experiments were designed according to the "expert" judgment of the experimenter. As such, there were no rigorous controls over unknown sources of variation. Moreover, experiments could only be conducted to investigate a single factor at a time, as there was no method available to disentangle the confounding effects of testing multiple factors simultaneously—in fact, there was no ability to even determine when and how confounding might

[11] Neyman *et al* 1933, pp. 289-337.
[12] Stigler 2016, pp. 83-84.
[13] Rao 1945, pp. 81-89

occur. As well, there was no method available to handle the variation in measurements across multiple experiments. Thus, any such variation—if considered at all—was incorrectly attributed to the parameter of interest.

Fisher's work changed everything. In the words of Stephen Stigler, "Fisher consign[ed] a large part of two thousand years of experimental philosophy and practice to the dustbin."[14] Fisher introduced the idea of randomization to wash out the effects of unknown sources of variation. This single concept permeates everything we do as statisticians. For example, randomization is the foundation of so much of linear statistical inference and is a basic condition of every statistical scenario we have discussed in *Stat 522*. For championing randomization alone, Fisher should be considered a hero.

Fisher also introduced blocking into the design process, thereby providing a way to account for and control the variance associated with nuisance parameters. He additionally developed factorial designs which, in combination with blocking, allowed researchers to study the effects of various levels of multiple parameters simultaneously and across multiple experimental runs. And to tie this all together, Fisher developed the Analysis of Variance, allowing for rigorous hypothesis testing of experimental results using well understood distributions such as the F-distribution. It was also Fisher who coined the term *null hypothesis*[15]. In short, Fisher single-handedly brought design theory into the modern era. His textbook *The Design of Experiments*, first published in 1935, is a *tour de force* of analytical structure and insight.

Discriminant Analysis

The ability to classify objects into two or more groups is an incredible analytic and computing achievement. Modern advances via machine learning methods have led to such marvelous tools as facial recognition and artificial intelligence-based medical diagnosis. The roots of analytical classification methods can be traced to Fisher's *linear discriminant*, a concept he proposed in 1936[16]. His approach (again a topic in *Stat 522*) was to find the separation threshold that

[14] Stigler 2016, p. 153.
[15] Fisher 1935, pp. 15-17.
[16] Fisher 1936, pp. 179-188.

maximized the ratio of between-class variation to within-class variation. Although the computational approach to classification problems has certainly advanced over the past century, the fundamental concept—that the best classification system is one where the classes are heterogeneous across but homogeneous within—has remained essentially unchanged since Fisher's original work in this area.

P-Values and Significance

Fisher did not develop the concept of the p-value, nor was it he who popularized its use. But Fisher did have an immense impact on the interpretation of p-values, an impact whose benefit has become a topic of fierce debate. Specifically, Fisher suggested that a p-value of 0.05 should be used as the cut-off for statistical significance. He did so in an entirely casual way, in a discussion of a standardized distribution:

> "The value for which p = .05, or 1 in 20, is 1.96 or nearly 2; it is convenient to take this point as a limit in judging whether a deviation is to be considered significant or not. Deviations exceeding twice the standard deviation are thus formally regarded as significant."[17]

With these two short sentences, Fisher set a standard for significance testing that has been followed almost without question by throngs of researchers for almost 100 years. Yet we believe that Fisher would be upset at how near-universal the use of $p = 0.05$ has become as the measure of statistical significance. He says "it is convenient" to use $p = 0.05$—but we take that to mean convenient in the absence of other indicators. Fisher was keenly aware that different scientific questions require different standards for what might be considered a significant dynamic. The misuse and misunderstanding of the nature of the p-value has become so pervasive that the American Statistical Association has an entire website dedicated to the conversation of whether p-values should even be used anymore in academic journals[18]. Fisher was an advocate for understanding the underlying dynamics of an experiment, and for thoughtfully applying all relevant statistical tools to its analysis. It is unfortunate that we cannot ask his opinion of the current debate, but we feel confident he would spurn the mechanical use of $p = 0.05$ as lazy.

An Overarching Structure

---

[17] Fisher 1925, p. 44.

[18] https://amstat.tandfonline.com/doi/full/10.1080/00031305.2016.1154108#.YH0LQuhKiUm as of April 18, 2021.

One might think of each of Fisher's contributions as huge achievements individually, but when taken together they comprise something more. Fisher built and formalized a rigorous, connected structure for mathematical statistics, whereas previously there had been only a juxtaposition of diverse ideas. This is perhaps best summed up by Joan Fisher Box, Fisher's daughter, in her biography of her father, a portion of which we present here edited for clarity.

> "It is difficult now to imagine the field of mathematical statistics as it existed in the first decades of the twentieth century. By modern standards the terms of discourse appear crude and the discussion extremely confused. The whole field was like an unexplored archeological site, its structure hardly perceptible. There existed a great endeavor from which flowed a profusion of statistical measures and tables by which to describe and distinguish the masses of observations being collected in the field. However, the principles that could justify the use of these measures had not been closely examined.
>
> "[These fundamental problems] were difficult both philosophically and mathematically. More clearly than others, [Fisher] recognized the fundamental issues and set himself the task of finding solutions. In the process he defined the aims and scope of modern statistics and introduced many of the important concepts, together with much of the vocabulary we use to frame them."[19]

Fisher gave us not just many of the most important tools of mathematical statistics, he gave us a *philosophy* of statistics, and the rigorous mathematical foundations upon which to build more and more powerful statistical tools. It is perhaps this contribution, while subtle, that might be considered Fisher's greatest in the field of statistics.

Direct and Indirect Influence

Fisher's ideas appear in one form or another in every statistics course taught today. Literally millions of students have been exposed to and benefitted from his innovations. Fisher's textbook *Statistical Methods for Research Workers* was the first textbook to provide a rigorous treatment of mathematical statistics when it was originally published in 1925, and has gone through at least 14 editions to date. As such, it is arguably the most widely read statistical textbook in existence. Fisher went on to write two other textbooks in statistics: *The Design of Experiments* (first published in 1935) and *Statistical Methods of Scientific Inference* (first published in 1956).

[19] Fisher Box 1978, pp. 62-63.

Beyond his textbooks, Fisher wrote hundreds of scholarly articles and conducted an immense scholarly correspondence. In 1971 professor J. H. Bennet of the University of Adelaide published some 294 of Fisher's articles in a five-volume collection. Yet, while this collection is extensive, it is still not considered exhaustive.

According to the *Mathematical Genealogy Project*[20], Fisher had six direct Ph.D. students and an impressive 1,318 academic descendants (i.e., Ph.D. students of his students, and of their students, etc.). The most famous of Fisher's students is Calyampudi Radhakrishna (C.R.) Rao, who alone has fostered at least 688 of Fisher's academic descendants. Beyond that, Rao's seminal 1965 textbook on linear statistical inference is a classic in the field, and we have referred to its 1973 edition often in *Stat 522*.

Genetics and Agriculture

Amazingly, statistics was not the only field where Fisher made enormous contributions. He was also one of the most important figures in the field of analytical genetics, and modernized the discipline of agricultural experimentation. Indeed, we doubt Fisher would have described himself as a statistician, in that he did not seem to study statistics for its own sake. Rather, he likely would have called himself a geneticist, and probably considered all of his statistical achievements as a means to an end, i.e., the development of the tools he needed for the study of genetics, or for more rigorous experimentation in agriculture.

His contributions to the study of genetics are too numerous to recount here. But to give a sense for his brilliance in the field, we note that in 1930 Fisher published a book called *The Genetical Theory of Natural Selection*, which presented a unifying theory that reconciled a large and bitter debate between the Charles Darwin school of thought, which espoused the theory of natural selection, and that of Gregor Mendel, whose premise was the laws of trait inheritance. As much as Fisher is revered in the field of statistics, his stature is at least as high in the field of genetics.

Ironically, in considering the role of probability in the field of statistics Fisher was a staunch frequentist, and was fiercely opposed to the Bayesian philosophy. Unlike in genetics, Fisher was

[20] https://genealogy.math.ndsu.nodak.edu/id.php?id=46924 as of April 18, 2021.

unable—or uninterested—in reconciling the two competing schools of thought, a competition which continues (although perhaps not quite so fiercely) to the present day.

Eugenics

If one only took into account Fisher's immense contributions, one might be forgiven for considering him something of a superman. But Fisher was human, and also had definite shortcomings—some of which were exacerbated by the very fact of his professional prestige and authority. We cannot write a report discussing all the things Fisher did to advance science without also discussing what he did to damage its progress. We refer here primarily to Fisher's endorsement of the pseudoscientific, racially-tinged concept of *eugenics*, one heavily favored by the pro-imperialist, anti-Semitic and even fascist parties of his day. And while Fisher would have argued (as did his daughter in her biography of him[21]) that his aims were noble, that his goal was, via selective breeding, nothing more than the improvement of the quality of life of all mankind, the concept of selecting "superior" humans for "desirable" qualities inherently leads to the persecutory creation of categories of "inferiors" based on arbitrarily chosen standards of race, heritage and ethnicity. The natural question which, despite his great acumen, Fisher never clearly posed to himself, is who decides what qualities are desirable? And what happens to those people now deemed as possessing traits considered undesirable? Had Fisher voiced such notions while sitting in a tavern, they still would have done a certain amount of harm. But what makes Fisher's support for eugenics so damaging was his willingness to put them into print and advocate for eugenics so publicly, thereby lending it his considerable prestige and authority. No sincere admirer of Fisher's accomplishments—as we are—can fail to reflect with sadness that he devoted considerable energies in an effort to give scientific legitimacy to an unworthy and damaging cause.

Yet, without apologizing for or distorting the record, it still is possible to understand Fisher's beliefs in the context of his times. Fisher came from a relatively well-to-do background and upbringing. He was not the only one of his class in the early part of the twentieth century to parrot then-fashionable notions of "Anglo-Saxon superiority." These racists beliefs, which came in a variety of forms, were driven in large part by a combination of anxiety over mounting challenges to the British Empire, combined with an exaggerated belief, imbued in childhood, of the

[21] Fisher Box 1978, p. 2.

indispensability of the supposed “British race” to human progress. Putting it this way is not meant to condone Fisher’s own contribution, but merely to point out that he, like all of us, was a product of his times. Every scientist is a combination of insight and ignorance; and if Fisher was a giant in terms of the former, we should also acknowledge that he failed to live up to his own high intellectual standards when it came to social-political beliefs. Consequently, though we can understand a daughter’s loyalty, we see no need to follow Joan Fisher in her apologetic effort to portray her father’s views as somehow being noble at their root. Rather, we can see Fisher’s ideology for the ugliness it represents. But having done so, we can acknowledge this failing without vitiating the value of Fisher’s enormous contributions to the discipline of statistics.

Summary

We have only provided the broadest of brush strokes describing what we would consider the most important of Fisher’s accomplishments—and even so, it is hard to believe that one theorist could accomplish so much. It sometimes feels that five lifetimes would not be enough to do what Fisher did. Fisher was a once-in-a-generation talent who gave statistics the coherent structure and analytical rigor that has allowed it to flourish as an independent field of study for 100 years. With computing power continually growing and the amount of available data skyrocketing, the discipline of statistics has evolved and expanded in new and exciting directions. And while the future of statistics will certainly spawn its share of intellectual visionaries, it will be on Ronald Fisher’s shoulders that they stand.

References

Bennett, J. H., editor (1971-1974). Collected papers of R. A. Fisher. University of Adelaide.

Fisher, R. A. (1924) “On a distribution yielding the error functions of several well known statistics”. *Proceedings International Mathematical Congress,* Toronto. 2, pp. 805-813.

Fisher, R. A. (1925). Statistical Methods for Research Workers. The University of Adelaide.

Fisher, R. A. (1930). The Genetical Theory of Natural Selection. Oxford at the Clarendon Press.

Fisher, R. A. (1935). The Design of Experiments. The University of Adelaide.

Fisher, R. A. (1936). "The Use of Multiple Measurements in Taxonomic Problems". *Annals of Eugenics*. **7** (2): 179–188.

Fisher, R. A. (1956). Statistical Methods and Scientific Inference. The University of Adelaide.

Fisher Box, Joan (1978). R. A. Fisher: The Life of a Scientist. New York: John Wiley & Sons.

Mathematical Genealogy Project. https://genealogy.math.ndsu.nodak.edu/id.php?id=46924 as of April 18, 2021.

Neyman, J. and E. S. Pearson. (1933). "IX. On the problem of the most efficient tests of statistical hypotheses". *Phil. Trans. R. Soc. Lond. A.* **231** (694–706): 289–337.

Rao, C. Radhakrishna (1945). "Information and the accuracy attainable in the estimation of statistical parameters". *Bulletin of the Calcutta Mathematical Society*. **37**: 81–89.

Rao, C. Radhakrishna. (1965) Linear Statistical Inference and its Applications. New York: John Wiley & Sons.

Stigler, Stephen M. (1986). The History of Statistics: The Measurement of Uncertainty before 1900. The Belknap Press of Harvard University Press.

Stigler, Stephen M. (2016). The Seven Pillars of Statistical Wisdom. Harvard University Press.

Student. (1908) "The probable error of a mean". *Biometrika*. **6** (1): 1–25.