



Abraham Wald and the Origins of the Sequential Probability Ratio Test

Joel B. Greenhouse & Christopher J. Phillips

To cite this article: Joel B. Greenhouse & Christopher J. Phillips (23 Dec 2025): Abraham Wald and the Origins of the Sequential Probability Ratio Test, The American Statistician, DOI: [10.1080/00031305.2025.2604805](https://doi.org/10.1080/00031305.2025.2604805)

To link to this article: <https://doi.org/10.1080/00031305.2025.2604805>



Accepted author version posted online: 23 Dec 2025.



Submit your article to this journal [↗](#)



Article views: 59



View related articles [↗](#)



View Crossmark data [↗](#)

Abraham Wald and the Origins of the Sequential Probability Ratio Test

Joel B. Greenhouse^{a,*}, Christopher J. Phillips^b

^aDepartment of Statistics & Data Science, Carnegie Mellon University

^bDepartment of History, Carnegie Mellon University

*joel@stat.cmu.edu

Abstract

Abraham Wald's formalization of the sequential probability ratio test in the crucible of World War II is one of the more famous cases in the history of statistics of the interplay of statistical theory and real-world applications. Focusing entirely on the moments around its creation, however, obscures the way in which it was also a continuation of previous work he had done in the late 1930s, and in particular, Wald's development of decision theory and his approach to using inverse probability. By situating the origins of the sequential probability ratio test in a broader history, we see not only how inverse probability initially made its way into sequential analysis but also the ongoing importance of the role of applications in motivating the development of statistical theory.

Keywords: W Allen Wallis, Inverse Probability, Statistical Research Group, History of Statistics, SPRT

1 Introduction

During World War II, civilian-scientists in the United States and Great Britain were recruited to work in centrally managed "research groups" dedicated to bringing the best scientific minds together to solve pressing military problems. Two well known examples are the Manhattan Project working on the development of the atomic bomb and the group at Bletchley Park in the UK deciphering military codes used by the Nazis. Another such collaboration, albeit not as well known, was the Statistical Research Group (SRG) based at Columbia University. The SRG consisted of a number of statisticians, economists, and mathematicians supported by the Applied Mathematics Panel (AMP) of the National Defense Research Committee (NDRC) constituted within the Office of Scientific Research and Development (OSRD) under the leadership of Vannevar Bush. W. Allen Wallis, the SRG Director of Research, preserved a history of the SRG in a personal and entertaining paper originally presented as the 1978 ASA Presidential Invited Address (Wallis 1980). According to Wallis "The sole purpose of the SRG was to serve the Army, Navy, Air Force and Marines ... Our work, however excellent, was in effect not delivered if it had no influence; so we had to understand the client's viewpoint and needs and be persuasive and accommodating." The SRG's outlook embodied a perspective long recognized and embraced by statistical scientists as central to the practice of statistics: important theoretical innovations are driven by important real-world applications.

The principal members of the SRG is a Who's Who of twentieth-century giants in the fields of statistics and economics, including Kenneth Arrow, Milton Friedman, Harold Hotelling, Frederick Mosteller, L. J. Savage, Herbert Solomon, George Stigler, Abraham Wald, Allen Wallis, and Jacob Wolfowitz. In his paper, Wallis briefly reviews several of the important problems the SRG worked on. He goes into detail, however, on one of the most far reaching and seminal developments in statistics to come out of the SRG: sequential analysis.

As Wallis tells the story he was tasked with evaluating a proposal by a Navy Captain who suggested that some sampling inspection problems could be better dealt with by adjusting the sample size as the sampling proceeded, rather than sticking rigidly to a predetermined sample size. In wartime, efficient sampling procedures were not just of mathematical interest: they mattered when materials, such as supplies of ordnance, were limited and of critical need for the war effort. The idea of a sequential inspection process had been around for a long time. For example, Dodge and Romig (1929) proposed a two-stage acceptance sampling scheme which required a decision at the end of stage 1 to accept or reject a lot, and if not, then to continue making observations during a second stage of inspection and make the final decision at the end of stage 2. As Wallis reports, there was a strong intuition that by using a sequential scheme that inspected each item, one at a time, an investigator could make a decision early rather than sticking rigidly to a design based on a predetermined sample size. Wallis reasoned "... it may become impossible for the full sample to lead to rejection, or for it to lead to acceptance, in which case there is no sense in completing the full sample," and speculated that "... it might pay to design a test in order to capitalize on this sequential feature; that is, it might pay to use a test which would not be as efficient as the classical tests if a sample of exactly N were to be taken, but which would more than offset this disadvantage by providing a good chance of terminating early when used sequentially."

The statistical question, then, was among all possible sequential designs was there an optimal one in the sense that it could reach a decision with an average sample size smaller than a pre-determined, fixed-sample size design while ensuring a specified probability of a type 1 and

type 2 error? Wallis and his SRG Deputy Director, economist Milton Friedman, thought there was but couldn't work out the theory. They approached Jacob Wolfowitz, a recent PhD in mathematics from NYU, who, according to Wallis, was skeptical and wasn't interested in pursuing this. Next, Wallis and Friedman presented the problem to Wald who initially was also skeptical. But, as Wallis tells the story, "The next day Wald phoned that he had thought some about our idea and was prepared to admit there was sense in it. . . On the second day, however, he phoned that he had found that such tests do exist and are more powerful, and furthermore he could tell us how to make them."

To be able to pinpoint to the day a novel scientific discovery and the individual responsible for it is remarkable. How was Wald able to come up with a solution to the Wallis-Friedman problem of sequential testing so quickly? Did he develop the approach and supporting theory overnight from scratch? Or, did he recognize the elements that were necessary to solve the problem and realized that he already had the tools at his disposal to solve it? These are the questions that we explore in this paper. Of course, we cannot know the explicit thought processes that led Wald to discover the sequential probability ratio test (SPRT). While we do not have extensive archival materials or personal interviews with him, we do have a copy of the contemporaneous, classified report Wald (1943) wrote detailing the derivation of the SPRT (declassified in 1945) which is nearly identical to but with some important differences to the published version of that report that appeared in the *Annals of Mathematical Statistics* (1945), and Wald's book, *Sequential Analysis* (1947).

2 Abraham Wald

Abraham Wald was born in Klausenburg, Hungary in 1902 to an observant Jewish family and died tragically at the age of 48 in a plane crash in India in December 1950. His undergraduate and graduate studies were in mathematics at the the University of Vienna where his exceptional gifts as a mathematician were quickly recognized. His interest in mathematics was at first in the fields of metric spaces, set theory and differential geometry (Menger 1952). After receiving his Ph.D. in 1933 he was unable to obtain employment at the University because of religious persecution. Luckily his mentor, the mathematician Karl Menger, introduced him to Oskar Morgenstern, the director of the *Austrian Institute for Business Cycle Research*, who had an interest in the applications of mathematics to economics and statistics. Morgenstern (who later co-authored the landmark book on game theory with John von Neumann (1944)) initially got Wald interested in a problem related to the analysis of seasonal variation in time series data (Wald 1936; Morgenstern 1951). In addition to working on problems in mathematical economics, Wald also began making contributions in probability theory.

In 1937 Wald accepted an invitation to be a visiting fellow with the Cowles Commission for Research in Economics in the United States to begin in the summer of 1938. The timing of this fellowship turned out to be critical since March 1938 was the Nazi Anschluss of Austria and Wald lost his position at the Institute and was forced to leave Austria. Shortly after arriving at the Cowles Commission, he received an invitation from Harold Hotelling, arguably the foremost statistician in the US at the time (Stigler 1996), to become a research associate in the Department of Economics at Columbia University. Starting at Columbia in the Fall of 1938, Wald spent the year immersed in learning modern statistics by reading and attending Hotelling's lectures (Wolfowitz 1952). Apparently he was a quick study. According to Wolfowitz "Wald mastered the subject" so well that his own lectures at Columbia in 1939-40 and elsewhere "were noted for their lucidity and mathematical rigor. Students not only

flocked to them, but clamored for a record of them.” (See, for example, Wald 1942.) Ingram Olkin (2013) a student at Columbia in the late 1940’s recalled,

Wald had a classic European lecture style. He started at the upper left corner of the black-board and finished at the lower right. The lectures were smooth and the delivery was a uniform distribution. Though I had a lovely set of notes, Wald treated difficult and easy parts equally, so one did not recognize pitfalls when doing homework.

By the beginning of the 1940’s Wald had successfully transitioned from working on problems in mathematical economics to making deep contributions to mathematical statistics. Given his mathematical ability, it is not a surprise that he was invited to join the SRG even though he was not yet a US citizen. Among Wald’s earliest publications in statistics, his 1939 paper published in the *Annals of Mathematical Statistics* is considered his most important and original work. In this paper, Wald established the foundations for statistical decision theory. Although the scope and impact of this paper would not be fully appreciated for a decade, it opened up a totally new field in statistics. Decision theory centers on the problem of statistical action, that is, deciding on a reasonable course of action based on incomplete information. Wald’s theory encompassed both statistical estimation and hypothesis testing. Briefly, consider a family of distributions indexed by a parameter θ where the values of θ can be thought of as states of nature. Given a random sample of size n from this family, the problem is to decide the specific distribution from which the sample is drawn, i.e., the specific value of θ . In the language of decision theory, the decision regarding a hypothesis is an action. A statistical decision function is a rule for making a decision about θ for each possible sample. As we will see, this framing gave Wald a unique perspective on the sequential inspection problem.

Following Wald’s untimely death, Wolfowitz (1952) reflected on his statistical outlook,

Wald brought to statistics a very high degree of mathematical ability and knowledge. Along with this, and in spite of his abstract and theoretical bent and predilections, he never, in any statistical investigation, lost sight of the fact that there was a question to be answered . . . Wald not only posed his statistical problems clearly and precisely, but posed them to fit the practical problem . . . This, in my opinion, was the key to his success - a high level of mathematical talent of the most abstract sort, and a true feeling for, and insight into, practical problems.

This ”feeling for and insight into practical problems” along with Wald’s exceptional mathematical ability would be key to the development of the sequential probability ratio test (SPRT).

3 Wald’s Derivation of the SPRT

We begin with some background and notation. The problem that Wallis and Friedman presented to Wald can be framed in the context of acceptance sampling where a unit, for example a manufactured object, is tested and classified into one of two categories, non-defective or defective. Typically, a random sample of a fixed number of items from the population (the lot) is tested and a criterion is specified for making a decision. For example, if the number of defective items is greater than some specified number, the lot could be considered unacceptable and would be rejected. As noted earlier, it is not hard to imagine a

scenario where the stated criterion was met before the predetermined number of items had been tested.

Let X_1, X_2, \dots, X_n be a random sample of fixed size n taken from the distribution $f(x, \theta)$ which is known other than the value of θ . (Note, we are using Wald's notation, $f(x, \theta)$, for the probability distribution for X .) Initially, Wald focused on the test of two simple hypotheses:

$$H_0 : \theta = \theta_0 \text{ or } X \sim f(x, \theta_0) \text{ vs } H_1 : \theta = \theta_1 \text{ or } X \sim f(x, \theta_1)$$

where the choice θ_0 vs θ_1 corresponds to two alternative actions, for example, accept the lot or reject it. Let $p_{0n}(x_1, \dots, x_n) = f(x_1, \theta_0) \cdots f(x_n, \theta_0)$ denote the joint probability distribution of the sample of size n given H_0 is true and let

$p_{1n}(x_1, \dots, x_n) = f(x_1, \theta_1) \cdots f(x_n, \theta_1)$ denote the joint probability distribution of the sample given H_1 is true. Neyman and Pearson (1933) showed that a region consisting of all samples (x_1, \dots, x_n) which satisfy the inequality

$$\frac{p_{1n}(x_1, \dots, x_n)}{p_{0n}(x_1, \dots, x_n)} \geq K_\alpha \quad (1)$$

is a most powerful critical region for testing H_0 against H_1 where K_α is a constant chosen so that the region will have the required probability of type I error, α . In the following, to highlight that the sample size is no longer fixed we use m to denote the size of the sample at the m -th stage, $m = \{1, 2, 3, \dots\}$.

Wald presented the derivation and basic theory of the SPRT in a classified report (henceforth referred to as the Report) dated September 1943. Because of the value of sequential analysis to the military, Wald's report was classified under the Espionage Act as "Restricted". The Report was declassified in May 1945. In June 1945, Wald published the Report along with further advances in the theory of sequential analysis in a 70 page paper that appeared in *The Annals of Mathematical Statistics*, and in his 1947 book, *Sequential Analysis*. The derivation of the SPRT presented here is based on the original classified report (Wald 1944).

The Report begins with a review of the Neyman-Pearson fixed sample size test of two simple hypotheses as described above. He then suggests an approach to the sequential testing problem.

Let d_0 and d_1 be two positive numbers less than 1 and greater than 1/2. Suppose that we want to construct a sequential test such that the conditional probability of a correct decision under the condition that H_0 is accepted is greater than or equal to d_0 , and the conditional probability of a correct decision under the condition that H_1 is accepted is greater than or equal to d_1 .

From the outset, Wald is approaching the sequential test as a decision problem where the action, i.e., choose H_0 or H_1 , is based on the probability of making a correct decision. And to be clear, what he means by "the conditional probability of a correct decision under the condition the hypothesis is accepted" is the posterior probabilities of H_0 and H_1 which he calculates using the principle of inverse probability. Inverse probability refers to the method for making inferences about the unobserved value of θ based on the data using Bayes' rule and is the approach we now recognize as Bayesian inference. The method goes back to Laplace and played a central role in the early development of statistics up through the early to mid 20th century. According to Fienberg (2006, p. 6), inverse probability was "the method of choice of the great English statisticians of the turn of the century, such as Edgeworth and Pearson" and "... when it came to the practical application of statistical ideas for inferential purposes, inverse probability ruled the day." Thus, the use of inverse probability, in this case calculating the posterior probabilities of H_0 and of H_1 , would have been a familiar and logical approach for Wald.

Following Wald, let the prior probability for H_0 be $g_0 = P(\theta = \theta_0)$ and the prior probability for H_1 be $g_1 = P(\theta = \theta_1) = 1 - g_0$. By Bayes rule, the posterior probabilities for H_0 and H_1 after the m -th observation are,

$$g_{0m} = \frac{g_0 p_{0m}(x_1, \dots, x_m)}{g_0 p_{0m}(x_1, \dots, x_m) + g_1 p_{1m}(x_1, \dots, x_m)}, \text{ and } (2)$$

$$g_{1m} = \frac{g_1 p_{1m}(x_1, \dots, x_m)}{g_0 p_{0m}(x_1, \dots, x_m) + g_1 p_{1m}(x_1, \dots, x_m)} \quad (3)$$

respectively. Based on the posterior probabilities of H_0 and H_1 given in (2) and (3)¹, Wald proposed the following

At each stage calculate g_{0m} and g_{1m} , and

- Accept H_1 if $g_{1m} = \frac{g_1 p_{1m}(x_1, \dots, x_m)}{g_0 p_{0m}(x_1, \dots, x_m) + g_1 p_{1m}(x_1, \dots, x_m)} \geq d_1$ (4)

- Accept H_0 if $g_{0m} = \frac{g_0 p_{0m}(x_1, \dots, x_m)}{g_0 p_{0m}(x_1, \dots, x_m) + g_1 p_{1m}(x_1, \dots, x_m)} \geq d_0$ (5)

- Draw an additional observation if $g_{1m} < d_1$ and $g_{0m} < d_0$.

At the heart of Wald's proposal are three possible actions based on the cumulative data taken after the m -th observation: accept H_1 , accept H_0 , or take another observation. Furthermore, Wald's criteria for taking the specified actions are based on the posterior probabilities of the hypotheses.

Wald next investigates the properties of this proposed procedure. Specifically, he focuses on the ratio of the posterior odds for H_1 to the posterior odds for H_0 , $\frac{g_{1m}}{g_{0m}}$ (the ratio which today we recognize as the Bayes Factor). By rearranging terms, the ratio of the joint distribution of the sample given H_1 to the joint distribution of the sample given H_0 , can be written in terms of the prior and posterior probabilities.

$$\frac{g_{1m}}{g_{0m}} = \frac{g_1 p_{1m}(x_1, \dots, x_m)}{g_0 p_{0m}(x_1, \dots, x_m)} \Rightarrow \frac{p_{1m}(x_1, \dots, x_m)}{p_{0m}(x_1, \dots, x_m)} = \frac{g_0 g_{1m}}{g_1 g_{0m}}$$

Wald shows that the inequalities in (4) and (5) are equivalent to

$$\frac{p_{1m}(x_1, \dots, x_m)}{p_{0m}(x_1, \dots, x_m)} \geq \left(\frac{g_0}{g_1} \right) \left(\frac{d_1}{1-d_1} \right) \quad (6)$$

$$\frac{p_{1m}(x_1, \dots, x_m)}{p_{0m}(x_1, \dots, x_m)} \leq \left(\frac{g_0}{g_1} \right) \left(\frac{1-d_0}{d_0} \right) \quad (7)$$

and observes that the constants on the right hand side of (6) and (7) do **not** depend on m . Wald then writes

If an *a priori* probability of H_0 does not exist, or if it is unknown, the inequalities [6] and [7] suggest the use of the following sequential test: At each stage calculate

$$\frac{p_{1m}(x_1, \dots, x_m)}{p_{0m}(x_1, \dots, x_m)}$$

Accept H_1 if

$$\frac{p_{1m}(x_1, \dots, x_m)}{p_{0m}(x_1, \dots, x_m)} \geq A \quad (8)$$

Accept H_0 if

$$\frac{p_{1m}(x_1, \dots, x_m)}{p_{0m}(x_1, \dots, x_m)} \leq B \quad (9)$$

Take an additional observation if

$$B < \frac{p_{1m}(x_1, \dots, x_m)}{p_{0m}(x_1, \dots, x_m)} < A \quad (10)$$

The constants A and B are chosen so that $0 < B < A$ and the sequential test has the desired value of α of an error of the first kind and the desired value of β of the probability of an error of the second kind. We shall call the test procedure defined by [8], [9] and [10] a *sequential probability ratio test*.

Wald next investigates the selection of the constants A and B to ensure the sequential test does not exceed the desired type I and type II error rates. Again, he approaches this derivation based on the posterior probabilities of the hypotheses given in (2) and (3) above. We skip the details of the derivation which the interested reader can find in Wald (1945, 1947). Wald shows for a given α and β ,

$$A \leq \frac{1-\beta}{\alpha} \text{ and } B \geq \frac{\beta}{1-\alpha} \quad (11)$$

and further, he argues that a very good approximations for A and B are

$$A = \frac{1-\beta}{\alpha} \text{ and } B = \frac{\beta}{1-\alpha}.$$

Wald also addresses the central question about the optimality of the SPRT. He makes a strong, intuition-based argument that sequential tests are "more efficient and more economical . . . in the sense that they give equally sound conclusions with, on average, a smaller amount of data" (Wald 1944, p. v). Later, in a 1948 paper, Wald and Wolfowitz formally prove that a SPRT with prescribed error probabilities does indeed minimize the expected sample size of all tests having error probabilities that do not exceed those of the SPRT.

Commenting on the role of prior probabilities in the derivation of the SPRT Wald makes the following observation, "Since the inequalities under consideration are valid irrespective of the existence of an *a priori* probability of the hypothesis H_0 , it is obviously possible to eliminate the notion of an *a priori* probability from the proof altogether." In fact, Wald notes that a later section of the Report contains a proof of these inequalities with no use whatsoever of the notion of an *a priori* probability.

Given that in the end the SPRT does not depend on the specification of, or the existence of, prior probabilities why go through the derivation based on the method of inverse probability? In a footnote that appears in the Report on page 38, Wald explains

Since the alternative proof, given in Section 4.1, of the fundamental inequalities [i.e., eqs 8-11] is simple, it would have been possible to omit the approach based on *a priori* probabilities. The proof in terms of *a priori* probabilities has been included mainly because it gives an excellent intuitive indication of why the best sequential test should be based on the

probability ratio $\frac{p_{1m}(x_1, \dots, x_m)}{p_{0m}(x_1, \dots, x_m)}$.

Like others before him, it appears that Wald found the use of the principle of inverse probability a natural starting off point for developing inferential methods even if he didn't believe in the existence of prior probabilities.

4 On the Origins of the SPRT

Because of Wallis' documentation of the work of the SRG we have a remarkable window into the why and when of the beginnings of the field of sequential analysis. Still, Wallis didn't explain how Wald came up with the derivation of the sequential probability ratio test so quickly. Years later, however, when Wallis was asked whether he thought Wald had perhaps developed some notions of the SPRT and its properties previously, Wallis responded "As far as I know he had just constructed the test" in the 48 hours after the problem was posed to him (Olkin 1991). Although he may not have previously developed a solution specifically for the sequential sampling inspection problem, it seems to us that Wald's initial insight was framing the SPRT as a decision problem where after each observation the action was either accept H_0 , accept H_1 or take another observation. He conceived of partitioning the sample space after the m -th observation into 3 mutually exclusive sets: samples that supported H_0 (accept H_0), samples that supported H_1 (accept H_1), and samples for which the accumulative evidence was insufficient to make a decision (take another observation). Establishing the criteria for choosing which action to take based on the posterior probability of the hypotheses would have been a familiar and, as Wald himself stated, "intuitive" approach.

We noted earlier that among Wald's earliest publications in statistics, his 1939 paper published in the *Annals of Mathematical Statistics* established the foundations for statistical decision theory. This framing of a decision problem, allowing for multiple states of nature, was a generalization of a Neyman-Pearson hypothesis test consisting of two actions, i.e., accept or reject H_0 . We believe the sequential acceptance sampling problem presented to Wald by Wallis and Friedman provided Wald with a real-world problem to which to apply his theory of statistical decision functions and was the theoretical foundation for the development of the SPRT. Where did Wald's ideas for statistical decision theory come from? It is likely that Wald's outlook on statistical inference was informed by his work in economics with Morgenstern and others in Vienna. Support for this conjecture can be found in Hunter (2004) who discusses Wald's collaborations in mathematics and economics in inter-war Vienna and the influence of this intellectual community on his contributions to statistics, and Giocoli (2013).

Was Wald a Bayesian? Although Wald used the method of inverse probability, it is clear that he was not a Bayesian in the modern sense. As we have seen in his derivation of the SPRT the motivation for his use of Bayes rule was that it aligned with his intuition and led to good frequentist properties. His perspective on the use of prior probabilities is clearly stated in Wald (1939, p. 302):

... [R]egarding the introduction of an a priori probability distribution of θ ... First, the objection can be made against it, as Neyman has pointed out, that θ is merely an unknown constant and not a variate, hence it makes no sense to speak of the probability distribution of θ . Second, even if we may assume that θ is a variate, we have in general no possibility of determining the distribution of θ and any assumptions regarding this distribution are of hypothetical character ... The reason why we introduce here a hypothetical probability distribution of θ is simply that it proves to be useful in deducing certain theorems and in the calculation of the best system of regions of acceptance.

Wolfowitz, Wald's friend and close collaborator, observed in a memorial tribute (1952, p. 4) that "Wald made use of Bayes solutions purely as a mathematical tool and without invoking *objectionable connotations*" (emphasis is ours). It is interesting to note that the climate by 1952 was so hostile to the use of prior probabilities that in memorializing Wald, Wolfowitz felt it was necessary to use an euphemism for prior probabilities. Nevertheless, as Wald's contributions to the foundations of statistics began to be appreciated after his death, his contributions to the emergence of Bayesian inference in the latter half of the 20th century were elevated (see, e.g., Lindley (1953) and Spanos (2017)). Shortly after Wald's death George Barnard (1953 p.65) wrote

What Wald did from the practical point of view was to show that even if we disbelieve in the existence of a prior distribution, we should none the less behave as if it did exist. He thus reinstated Bayes's theorem to the central place in the theory of decisions which it had occupied in the time of Gauss and Laplace

... Wald helped to rescue Bayes's theorem from the obscurity into which it had been driven in recent years.

It is clear that Wald's work was grounded in the use of prior probabilities as a means to an end. Yet, if he had not died prematurely, it is interesting to wonder whether over time and the influence of, say, his SRG colleague L. J. Savage, whether Wald would have come around to embracing the Bayesian outlook.

5 Discussion

Our goal was to review the derivation of the SPRT and to investigate how Wald originally approached the problem, specifically focusing on which mathematical and statistical tools he based the discovery of the SPRT. For a history of the development of sequential analysis see, for example, Wald (1947), Ghosh (1991) and Lai (2001).

The story of the origins of the SPRT is emblematic of the history of statistics in the first half of the 20th century. Leading up to the 20th century, statistical techniques and approaches were developed by scientists who were primarily concerned with addressing problems of data reduction. The needs of the applied sciences motivated the development of statistical methods. As Porter (1986, p.10) notes ". . . practice was decidedly ahead of theory during the early history of statistics, and "pure" or abstract statistics was the offspring, not the parent, of its applications." Indeed, there were theoretical developments during this period, but scientific and practical problems drove the development of statistical methods (Stigler 1999, p. 87).

The beginning of the 20th century saw the continued development of practical methods to address the needs of the applied sciences, but it was also recognized that the theoretical foundations upon which these methods were based "had been neglected" and that there were "fundamental problems [that] had been ignored and fundamental paradoxes left unresolved" (Fisher 1922). Mathematicians who had become deeply involved in motivating subject-matter problems not only continued to develop methods to advance practice but also began to make significant advances to the theoretical foundations of statistics. Examples include the familiar founders of modern mathematical statistics such as K. Pearson, R.A. Fisher, J. Neyman and E. S. Pearson. Their work not only helped advance the particular subject-matter problems in which they were involved but the methods they developed were also generalizable to similar

classes of problems across disciplines. Furthermore, and significantly, the theory they developed opened up the field of mathematical statistics. Arguably, the statistician-scientist engaging in real-world problems was the model upon which the foundations of mathematical statistics was built.

The story of Abraham Wald and the development of sequential analysis is a case in point. The SPRT was motivated by the war-time problem of optimizing sampling inspections which led to wider applications, such as, for example, the design and analysis of sequential clinical trials. In addition, and significantly, Wald's work opened up the new theoretical field of sequential analysis. Wald was the epitome of the statistician-scientist, a gifted mathematician carefully framing statistical theory to fit the practical problem with which he was working (Wolfowitz 1952).

Although Wallis' article about the SRG provides the why and when of the beginnings of sequential analysis, he did not provide the how. Eighty years after the publication of Wald's derivation of the SPRT it is illuminating to recall that the foundations of statistical science are based on the critically important interplay between theory and applications. As Stigler noted decades ago in his account of the origins of the field, mathematical statistics emerged very clearly in a period of dialog and conversation between scientists and mathematical statisticians (Stigler 1996, p. 250). The SRG and Wallis' commitment to ensuring that the work of the SRG was relevant and solved real-world military problems during WWII is an exemplar of statistical practice then and now. Telling the story of the origins of the SPRT has also been an opportunity to recall the significant and impactful contributions of one of the giants of 20th century mathematical statistics, Abraham Wald.

Acknowledgements

The authors gratefully acknowledge the support for this work funded, in part, by the Carnegie Mellon University Dietrich College Inter-departmental Research Seed Grant program.

Disclosures

The authors report there are no competing interests including financial to declare.

Notes

¹ For clarity, equation numbers are consecutive within our paper and do not align with Wald's equation numbers.

References

- Barnard, G. (1953), Discussion of "Statistical Inference," by D. Lindley, *Journal of the Royal Statistical Society, Series B*, 15, 65.
- Dodge, H. F., and Romig, H. G. (1929), "A method of sampling inspection," *The Bell System Tech Journal*, 8, 613-631.
- Fienberg, S. E. (2006), "When did Bayesian inference become "Bayesian"?" *Bayesian Analysis*, 1, 1-40.
- Fisher, R. A. (1922), "On the mathematical foundations of theoretical statistics," *Philosophical Transactions of the Royal Society A*, 222, 594-604.
doi:<https://doi.org/10.1098/rsta.1922.0009>
- Ghosh, B. K. (1991), "A brief history of sequential analysis," in *Handbook of Sequential Analysis*, editors B. K. Ghosh and P. K. Sen. New York: Marcel Dekker.
- Giocoli, N. (2013), "From Wald to Savage: *Homo Economicus* becomes a Bayesian statistician," *Journal of the History of the Behavioral Sciences*, 49, 63-95. DOI: 10.1002/jhbs.21579
- Hunter, P. W. (2004), "Connections, Context, and Community: Abraham Wald and the Sequential Probability Ratio Test," *The Mathematical Intelligencer*, 26, 25-33.
<https://doi.org/10.1007/BF02985397>.
- Lai, T. L. (2001), "Sequential analysis: Some classical problems and new challenges," *Statistica Sinica*, 11, 303-408.
- Lindley, D. V. (1953), "Statistical Inference (with discussion)," *Journal of the Royal Statistical Society, Series B*, 15, 30-76.
- Menger, K. (1952), "The formative years of Abraham Wald and his work in geometry," *The Annals of Mathematical Statistics*, 23, 14-20.
- Morgenstern, O. (1951), "Abraham Wald, 1902-1950," *Econometrica*, 19, 361-367.
- Neyman, J and Pearson, E. S. (1933), "On the problem of the most efficient tests of statistical hypotheses," *Philosophical Transactions of the Royal Society of London, Series A*, 231, 263-294.
- Olkin, I. (1991), "A Conversation with W. Allen Wallis," *Statistical Science*, 6, 121-140.
- Olkin, I. (2013), "Reminiscences of the Columbia University statistics department in the late 1940s," <https://stat.columbia.edu/wp-content/uploads/2021/07/Olkin-Ingram-Dept-Statistics-article.pdf>
- Porter, T. M. (1986), *The Rise of Statistical Thinking, 1820-1900*. Princeton NJ: Princeton University Press.

Spanos, A. (2017), "Why the decision-theoretic perspective misrepresents frequentist inference: Revisiting Stein's paradox," in *Advances in Statistical Methodologies and their Application to Real Problems*, editor T. Hokimoto, DOI: 10.5772/65720.

Stigler, S. M. (1996), "The history of statistics in 1933," *Statistical Science*, 11, 244-252.

Stigler, S. M. (1999), *Statistics on the Table: The History of Statistical Concepts and Methods*. Cambridge MA: Harvard University Press.

von Neumann, J. and Morgenstern, O. (1944), *Theory of games and economic behavior*. Princeton, New Jersey: Princeton University Press.

Wald, A. (1936), *Berechnung und Ausschaltung von Saisonschwankungen, Beitrage zur Konjunktur-forschung* (Calculation and elimination of seasonal fluctuations, contributions to the economic cycle research), Vol. 9, Vienna: Verlag von Julius Springer.

Wald, A. (1939), "Contributions to the theory of statistical estimation and testing hypotheses," *The Annals of Mathematical Statistics*, 10, 299-326.

Wald, A. (1942), "On the principles of statistical inference: Four lectures delivered at the University of Notre Dame, February 1941," *Notre Dame Mathematical Lectures, Number 1*. <https://projecteuclid.org/ebooks/notre-dame-mathematical-lectures/On-the-Principles-of-Statistical-Inference/toc/ndml/1175196978>

Wald, A. (1944), *Sequential analysis of statistical data: Theory*. A report submitted by the Statistical Research Group, Columbia University to the Applied Mathematics Panel, National Defense Research Committee. (Declassified May 1945).

Wald, A. (1945), "Sequential tests of statistical hypotheses," *The Annals of Mathematical Statistics*, 16, 117-186

Wald, A. (1947), *Sequential Analysis*. New York: John Wiley & Sons.

Wald, A. and Wolfowitz, J. (1948), "Optimum character of the sequential probability ratio test," *Annals of Mathematical Statistics*, 19, 326-339.

Wallis, W. A. (1980), "The Statistical Research Group, 1942-1945," *Journal of the American Statistical Association*, 75, 320-330.

Wolfowitz, J. (1952), "Abraham Wald, 1902-1950," *The Annals of Mathematical Statistics*, 23, 1-13.

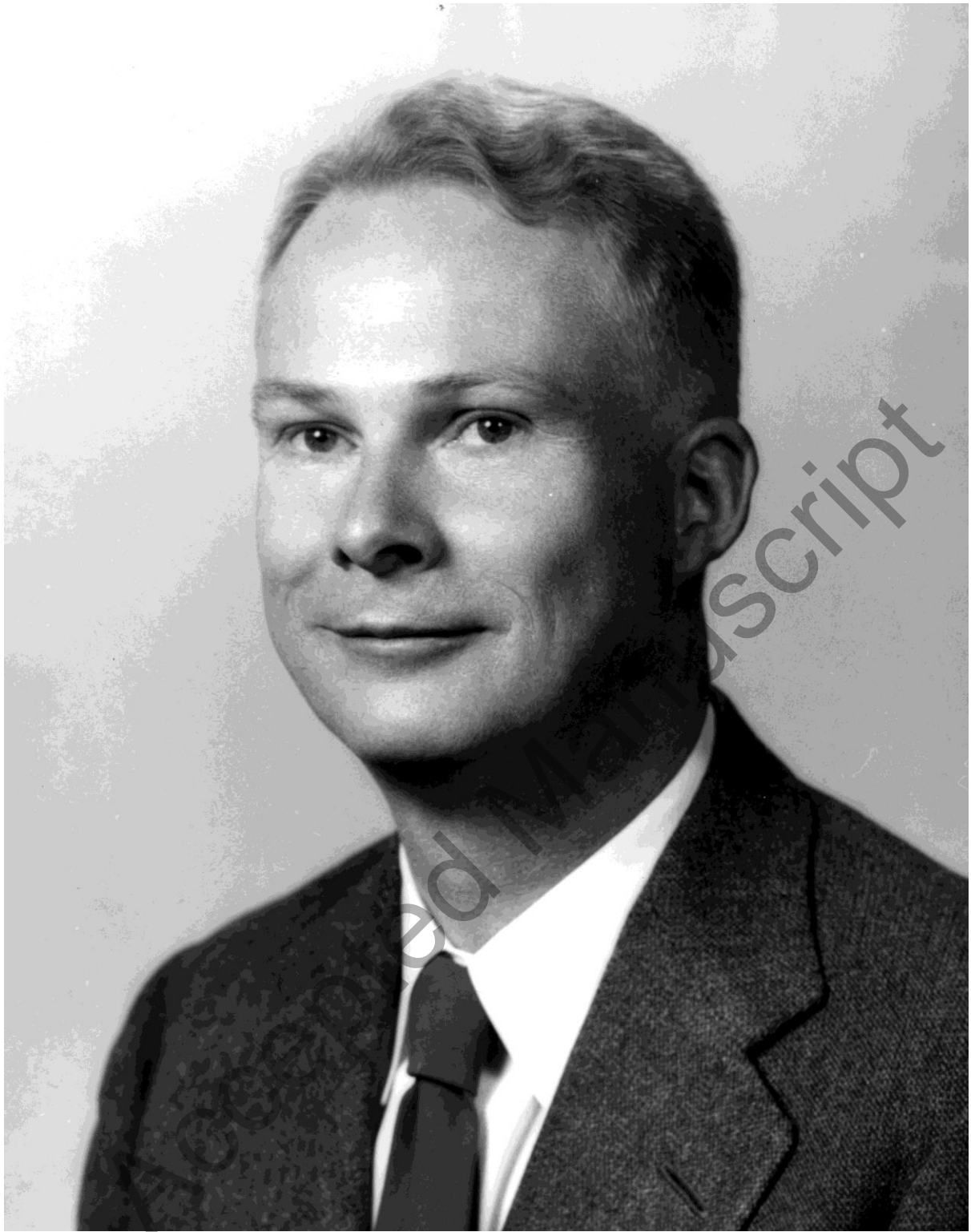


Fig.1 Photo of Wallis



Fig. 2 Photo of Friedman



Fig 3. Photo of Wald about here

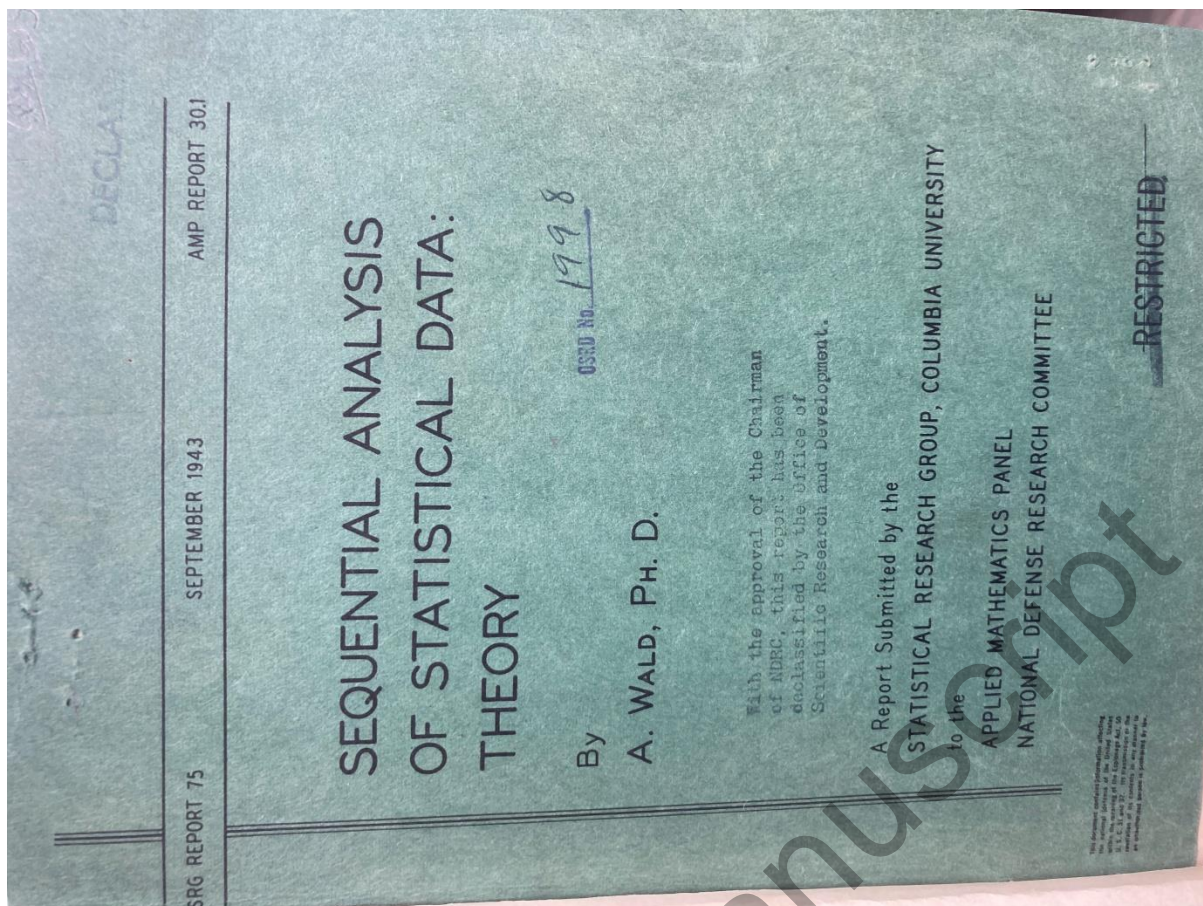


Fig. 4 Photo of cover of the Report