History of Statistics Paper I

Yingke(Kiki) He

May 19th 2025

Week 1: Overview

For Week 1, I focused on the foundations of statistical thinking by revisiting how core ideas such as probability, error measurement, and data combination gradually emerged over centuries.

Stigler, The History of Statistics, Chapters 1–2

1.1 The Method of Least Squares

Stigler begins his historical narrative with the rise of the method of least squares, which he identifies as one of the first fully developed statistical techniques. Though commonly associated with Adrien-Marie Legendre, who formally introduced the method in 1805 in an appendix to his

work on cometary orbits, Stigler shows that its intellectual roots go much deeper (Stigler, 1986, pp. 11–13).
Prior to Legendre, scientists like Tobias Mayer and Roger Boscovich had proposed methods to combine inconsistent observations, particularly in astronomy, though these lacked formal optimization criteria (Stigler, 1986, pp. 15–18). Mayer's practice involved grouping similar observational equations to reduce discrepancies, while Boscovich proposed minimizing the sum of absolute errors—an approach useful in one-dimensional problems but hard to generalize.

Legendre's major contribution was to articulate a general method for solving overdetermined systems of equations by minimizing the sum of squared residuals. This not only made the method mathematically tractable but also applicable across many scientific domains. Stigler notes that Legendre's clarity and practical demonstration using French meridian arc data accelerated the method's acceptance, especially in fields like geodesy and celestial mechanics (Stigler, 1986, pp. 19–24).

The story of least squares does not end with Legendre. Carl Friedrich Gauss, who later claimed to have used the method as early as 1795, provided a probabilistic justification grounded in the normal distribution. Though the priority dispute between Gauss and Legendre was significant, Stigler emphasizes that Gauss's theoretical work helped transform least squares from a computational rule into a statistically principled approach to inference. This intersection of measurement, error, and probability set the tone for the modern statistical paradigm.

1.2 Combining Observations and Handling Error

Following his introduction of the method of least squares, Stigler (1986, Ch. 1) then talks about how scientists historically dealt with variation in repeated measurements. Prior to the 18th century, it was common for astronomers to treat variation as a nuisance to be removed by averaging only when conditions were highly controlled. There was little sense of combining observations made under different circumstances. The method of least squares represented not only a computational solution but also a conceptual breakthrough: a method of integrating multiple imperfect data points to infer an underlying truth.

A central issue was how to formally treat measurement error—a problem that became important as observational sciences like astronomy and geodesy relied on more complex and precise data. Although early figures like Roger Cotes and Tobias Mayer offered practical examples for data combination, it was Legendre's formulation that provided a unified framework. Later, Gauss's probabilistic model, assuming normally distributed errors, justified least squares as the optimal method under certain conditions, thus connecting error theory and probability formally(Stigler, 1986, pp. 12–20).

This development marks a shift from only describing observations to modeling the uncertainty within them. As Kendall (1960) emphasizes in his historiographical critique, this statistical thinking did not emerge from ancient inventories or bureaucratic record-keeping but rather from the demands of scientific inquiry in post-1660 Europe. The motivation to quantify uncertainty and extract signal from noise became central to what would later be called inferential statistics.

Furthermore, Stigler's account shows that treating measurement error probabilistically laid the groundwork for statistical estimation. Instead of assuming a single "true" value obscured by error, scientists began to ask how confident they could be in their estimates given the observed variation. This shift from deterministic to probabilistic reasoning matches Hacking's (1975/2006) argument that probability emerged not merely as a mathematical tool but as a new conceptual space—a way to formalize belief and evidence under uncertainty. Hacking describes this as a transformation in the nature of epistemology itself, one that allowed statisticians and philosophers alike to reason about imperfect knowledge in a structured way.

Stigler also explains how the method of least squares became widely adopted by the early 19th century, partly because it could transform inconsistent observational equations into a coherent estimate. This practical value, combined with its growing theoretical justification, ensured its dominance in the physical sciences. The method's success signaled that measurement error was no longer an obstacle to be feared, but a manageable feature of scientific data, thus making the structure of empirical analysis uncertain (Stigler, 1986, pp. 22–27).

1.3 Early Probability and the Foundations of Inference

In Chapter 2, Stigler (1986) turns from the combination of measurements to the quantification of uncertainty—how thinkers began to mathematically model and reason about randomness, leading to the emergence of statistical inference. This shift began in the late 17th and early 18th centuries, especially with the work of Jacob Bernoulli, Abraham De Moivre, and Thomas Simpson.

Jacob Bernoulli's Ars Conjectandi (1713) contained the first major theoretical breakthrough in this direction: the Law of Large Numbers. This result mathematically demonstrated that the proportion of successes in repeated trials would converge toward the true probability as the number of trials increased. While it seemed simple, it was transformative —it gave formal expression to the intuition that more data leads to more reliable inference (Stigler, 1986, pp. 34–36). As Fienberg (1992) also notes, Bernoulli's contribution marked the moment when probabilistic reasoning was first connected to actual empirical observations.

Abraham De Moivre, building on Bernoulli's work, introduced an approximation that brought the normal distribution into statistical reasoning. His 1733 derivation of the normal approximation to the binomial distribution gave a powerful tool for making probabilistic statements about sums of random variables. Though still framed within forward-looking frequentist logic, De Moivre's result allowed for practical inferences in the presence of large sample sizes and inspired later formalizations of the Central Limit Theorem (Stigler, 1986, pp. 39–42).

However, neither Bernoulli nor De Moivre fully explained what is now recognized as inverse inference—reasoning from data back to an unknown parameter. That task fell more squarely on thinkers like Simpson, who in 1755 began to consider the errors of observation themselves and to quantify the uncertainty of the average as a random variable. Stigler (1986, pp. 43–45) credits Simpson with introducing a conceptual shift: the realization that averages themselves have uncertainty, and that their error could be estimated through assumptions about the distribution of individual errors.

This framework represented a precursor to later ideas about sampling distributions and estimation. As Kendall (1960) underscores, this growing interest in modeling variation probabilistically, rather than just eliminating it through averaging, was a mark of the emerging statistical mindset. It was no longer sufficient to describe what had been observed; thinkers were beginning to ask how likely it was that those observations reflected a deeper truth.

In philosophical terms, this transition reflects what Hacking (1975/2006) describes as the "emergence of probability"—not just as a branch of mathematics, but as a tool for reasoning under uncertainty. Rather than simply organizing past data, early probabilists were beginning to construct rules for belief, decision-making, and estimation, anticipating the foundational ideas of both frequentist and Bayesian inference.

1.4 Philosophical and Historical Context

While Stigler's chapters emphasize technical and conceptual breakthroughs, a richer understanding of the origins of statistical reasoning requires stepping back to ask: When did statistics actually begin? This is the central question posed by M.G. Kendall (1960) in his essay Where Shall the History of Statistics Begin?, where he challenges the tendency to locate the origins of statistics in ancient administrative record-keeping, such as the Domesday Book, Roman censuses, or biblical population counts.

Kendall argues that these early documents, while often cited as "proto-statistics," lacked any statistical intent or methodology. They were inventories, not inferences—designed for static record-keeping, not for drawing conclusions about variability or uncertainty (Kendall, 1960, p. 447). In his view, modern statistics only begins when data are used analytically, not merely collected descriptively. For Kendall, the true origins lie in the 17th-century political arithmetic movement, notably with John Graunt's (1662) Bills of Mortality and William Petty's economic analyses. These thinkers weren't just tallying data—they were attempting to reason with it, making estimations and predictions in a way that anticipates the logic of statistical inference.

This historical boundary-setting aligns with the philosophical argument by Ian Hacking (1975/2006) in The Emergence of Probability. As summarized by Barnouw (1979), Hacking contends that probability as a distinct epistemic category did not exist prior to the 1660s. While the elements—opinion, risk, signs, and uncertainty—were already present in earlier thought (e.g., in Aristotle's distinction between scientia and opinione), the transformation came when scholars began to treat degrees of belief as measurable quantities, grounded in empirical data and rules of inference. This marked a change in how people reasoned about knowledge, causation, and evidence.

What both Kendall and Hacking stress is that statistics is not just a body of techniques, but a way of thinking—a response to a new understanding of the world as uncertain, variable, and amenable to quantification. Stigler's chapters show how this mindset crystallized in specific tools like least squares and probabilistic error models, but Kendall and Hacking help us see the conceptual terrain that made such developments possible.

Fienberg, A Brief History of Statistics in Three and One-Half Chapters

2.1 Pre-History and Early Probability (1650–1750)

Stephen Fienberg begin his overview of statistical history by characterizing the period between 1650 and 1750 as the pre-history of statistics, during which the essential mathematical foundation for later statistical reasoning was established—but without the full conceptual apparatus of inference (Fienberg, 1992, p. 209). This era, which overlaps significantly with the developments described in Stigler's Chapter 2, centered on games of chance, life tables, and astronomical observations, but lacked a unifying statistical methodology.

The key scholars during this time—John Graunt, Christiaan Huygens, Jacob Bernoulli, and Abraham De Moivre—approached problems that involved uncertainty, but largely within the realm of classical probability rather than statistical inference. Fienberg highlights Graunt's 1662 Bills of Mortality as the first major attempt to derive conclusions from demographic data. This work marks an inflection point where empirical records were used not just descriptively but analytically, laying the groundwork for political arithmetic (Fienberg, 1992, pp. 210–211).

The period also saw the emergence of the Law of Large Numbers through Bernoulli's Ars Conjectandi, which formalized how observed frequencies would converge to true probabilities with more trials. Yet, as Stigler also emphasizes, Bernoulli was still working within a frequentist mindset, and neither he nor De Moivre provided a full framework for making inferences about unknown parameters from data (Stigler, 1986, pp. 34–39; Fienberg, 1992, p. 212).

Fienberg's framing here parallels Kendall's concern that modern statistical thinking should not be traced back to ancient inventories or medieval record-keeping. Instead, both authors identify the mid-17th century as an important period in which thinkers began to analyze variation systematically rather than merely document it. Moreover, Hacking's notion of "emergence" reinforces the idea that this era did not simply extend earlier traditions—it marked the creation of a new intellectual object: probability as a formal tool for reasoning about uncertainty (Barnouw, 1979; Hacking, 1975).

In this phase, then, the essential building blocks of statistical science—probability theory, life tables, error modeling—were laid down. But these tools were used to address specific, often narrow problems, such as evaluating annuities, predicting lifespans, or estimating the likelihood of gambling outcomes. A general theory of inference, applicable across scientific and social domains, had not yet been formulated.

Fienberg appropriately characterizes this period as foundational but incomplete: a time of great innovation in quantitative thinking, but still far from what we now recognize as statistical science. As such, this phase was not about statistics as a field, but about equipping future statisticians with the tools they would need to formalize inference in the centuries ahead.

2.2 Formal Inference and Mathematical Consolidation (1750–1820)

In the second phase of his historical outline, Fienberg (1992) identifies the period from 1750 to 1820 as the birth of formal statistical inference, during which probability theory began to be systematically applied to the combination of observations and the estimation of unknown quantities. This era builds directly upon the intellectual groundwork laid in the previous century, but introduces key methodological innovations that move statistics beyond descriptive summaries and into the realm of estimation, error quantification, and prediction.

One of the major turning points in this period is the method of least squares, formally introduced by Legendre in 1805, and further developed by Gauss, who later supplied it with a probabilistic foundation grounded in the normal distribution (Fienberg, 1992, pp. 213–214; Stigler, 1986, Ch. 1). The least squares method not only offered a practical solution for fitting models to data, but also embodied a broader idea: that one could construct formal rules for inference that take into account the imperfection of data and allow for the estimation of true values from noisy observations.

At the same time, Laplace was working toward a more unified treatment of probability, extending inverse probability methods and producing what would later be recognized as foundational work in Bayesian inference. His application of probabilistic reasoning to diverse areas—including planetary motion, population statistics, and error estimation—represented a major step toward viewing probability as a general tool for scientific reasoning, rather than a curiosity limited to gambling or actuarial science (Fienberg, 1992, pp. 214–215).

This period also witnessed the development and early application of the Central Limit Theorem—a result for the viability of statistical inference. Originally formulated by De Moivre and expanded by Laplace, the theorem showed that the distribution of the sum (or average) of many independent random variables approaches a normal distribution, regardless of the underlying distribution. This showed the reliability of inference methods based on the normal distribution and legitimized the growing use of probabilistic models in empirical science (Stigler, 1986, pp. 39–42).

In highlighting these developments, Fienberg paints this era as the first full synthesis of mathematics and uncertainty, where inference was no longer limited to ad hoc approximations or descriptive logic but formalized through mathematical tools and the concept of likelihood. This trend reflects what Hacking (1975) describes as the transition to a new epistemic regime in which probability—and not certainty—became the language through which scientific claims were evaluated (Barnouw, 1979, p. 439).

While the tools were still limited and often rooted in specific applications (e.g., astronomy, demography), this period marked the beginning of generalizable statistical methodology. The distinction between frequentist and Bayesian reasoning was not yet sharply drawn, but the groundwork for both traditions was in place. As Kendall (1960) noted, this was the moment when thinkers began reasoning about data as we do today, asking not just what the data say, but how confidently we can say it.

2.3 Expansion into the Social and Biological Sciences (1820–1900)

The third phase in Fienberg's narrative marks the diffusion of statistical thinking into new domains, particularly the social sciences and biology, between 1820 and 1900. This period saw the refinement and application of the mathematical tools developed earlier—such as least squares, correlation, and the normal distribution—in service of new empirical questions. These innovations not only expanded the reach of statistics but also reshaped its identity, as it became both a scientific method and a tool of governance.

A central figure in this expansion was Adolphe Quetelet, who introduced the concept of the "average man" (l'homme moyen), applying the normal distribution to human traits such as height, crime, and mortality. Quetelet's work exemplified the belief that social phenomena could be studied with the same statistical rigor as physical phenomena, and he helped promote the idea that variability in human behavior followed lawful, predictable patterns (Fienberg, 1992, p. 215).

While his application of the normal curve to moral and social traits was controversial, it represented an important shift toward quantifying human behavior and using statistical aggregates to draw conclusions about society.

This period also saw the emergence of regression and correlation, thanks largely to Francis Galton, who studied heredity and introduced the concepts of regression toward the mean and bivariate normal distribution in the 1880s. His ideas were further formalized by Karl Pearson, who developed mathematical definitions of correlation coefficients and the chi-square test, and by George Udny Yule, who contributed to the development of multiple regression and time series analysis (Fienberg, 1992, pp. 216–217; Stigler, 1986, Ch. 7–8).

While these tools were methodological breakthroughs, their conceptual and political implications were far-reaching. The statistical study of heredity and intelligence became tightly linked to the eugenics movement, a fact acknowledged by both Stigler and Fienberg, and which modern statisticians must confront as part of the discipline's legacy. As Hacking and later historians have argued, the epistemological power of statistics—to quantify uncertainty, compare populations, and infer causality—was not just used for scientific enlightenment, but also for social control and classification (Barnouw, 1979, p. 440).

Moreover, this phase reinforced a division noted by Kendall (1960): while descriptive statistics remained dominant in state bureaucracies and censuses, the inferential branch—rooted in probability theory—was increasingly adopted within scientific and academic circles. It wasn't until the end of the 19th century that these threads began to converge into a more unified discipline, aided by institutional developments such as the founding of Biometrika in 1901.

2.4 Institutionalization and Modern Statistics (1900–1950)

Fienberg's "half chapter" covers the period from 1900 to 1950, which he characterizes as the time when statistics crystallized into its modern form—not only as a mathematical discipline but as a professionalized and institutionalized field. This phase saw the emergence of key concepts that now define classical statistics: likelihood, sufficiency, hypothesis testing, estimation theory, and the separation of frequentist and Bayesian paradigms.

At the center of this transformation was Ronald A. Fisher, whose work in the 1920s and 30s revolutionized how statisticians approached experimental data. Fisher introduced the concept of maximum likelihood estimation, formalized the use of the analysis of variance (ANOVA), and established rigorous methods for designing experiments to control for confounding factors. His insistence on randomization and replication set the standard for scientific methodology across disciplines (Fienberg, 1992, pp. 218–219).

Fienberg emphasizes that Fisher's contributions were both foundational and polarizing. While his techniques became widely adopted, they also sparked debates—particularly with Jerzy Neyman and Egon Pearson, who challenged Fisher's informal use of p-values and instead proposed a framework of hypothesis testing based on Type I and Type II errors, power, and confidence intervals. Together, these competing approaches came to define the frequentist school, which dominated mid-20th century statistics (Fienberg, 1992, pp. 219–220).

Meanwhile, Bayesian methods, rooted in the earlier work of Bayes and Laplace, were largely sidelined during this period. As Hacking (1975/2006) points out, the early 20th century saw a cultural and philosophical shift that favored objective, long-run frequency interpretations of probability over subjective degrees of belief. Bayesian approaches persisted in limited circles—particularly through the work of Harold Jeffreys and Bruno de Finetti—but would not regain widespread prominence until the computational progresses of the late 20th century (Barnouw, 1979, pp. 441–442).

This period also saw the rise of statistical institutions and publications that solidified the field's identity. The founding of Biometrika (1901), the Royal Statistical Society's expanding influence, and the application of statistics during World War II (especially in operations research) elevated the discipline's practical relevance. As Fienberg notes, these institutional supports transformed statistics from a set of techniques into a profession with standards, pedagogy, and global reach (Fienberg, 1992, p. 221).

Yet even as the tools of modern statistics were maturing, their philosophical foundations remained unsettled. The debate between Fisher and Neyman–Pearson highlighted enduring tensions: Should inference be about decision-making or evidence? Should probability represent frequency in a population or belief in a hypothesis? These questions, first glimpsed in the works of Bernoulli, Laplace, and Quetelet, had now become defining concerns for the discipline.

By 1950, statistics had institutional authority, theoretical rigor, and broad interdisciplinary applications. But as Fienberg and Hacking both suggest, the epistemological tensions embedded in the field—from objectivity to subjectivity, from inference to decision—remained very much alive, setting the stage for future developments in both Bayesian revival and computational statistics.

2.5 Framing the Narrative

While Fienberg provides a structured and chronological account of statistics' development, the perspectives of M.G. Kendall and Ian Hacking invite us to rethink what defines the discipline's origins and character. Their ideas sharpen and deepen the story Fienberg tells.

Kendall (1960) cautions against tracing statistics back to ancient data collection efforts like censuses or inventories. These were descriptive tools, not analytical ones. For Kendall, statistics begins only when data are used to make inferences, a criterion that places its origin around the mid-17th century, with the emergence of political arithmetic in the work of Graunt and Petty. While Fienberg implicitly agrees by beginning his timeline in the same period, Kendall's framing emphasizes a conceptual threshold—that statistics must involve reasoning under uncertainty, not just recording facts.

Hacking's (1975/2006) view, as reviewed by Barnouw (1979), adds a philosophical dimension. He argues that probability emerged in the 1660s not as a solution to old problems, but as a new epistemological tool that redefined how people understood uncertainty and evidence. Where Fienberg presents a cumulative story of progress, Hacking highlights a conceptual rupture—a shift from certainty to probabilistic thinking as a foundation for knowledge.

Together, these perspectives show that statistics is more than a technical toolkit; it is a way of thinking, shaped by historical context and philosophical change. While Fienberg offers a timeline of growth, Kendall and Hacking prompt us to question when and why statistics became possible at all, and what assumptions continue to underpin its practice today.

What I Learned

This week's readings really changed how I think about statistics. I used to see it mostly as a set of methods for analyzing data, but now I understand that those methods have a long and complicated history. From Stigler, I learned that ideas like least squares and the normal distribution didn't just appear out of nowhere—they took decades of trial, error, and refinement. Scientists were trying to figure out how to deal with errors and how to make the best use of uncertain or inconsistent data. These techniques were built to solve real problems, especially in astronomy and measurement, and eventually became standard tools that we still use today.

Fienberg's reading helped me see how statistics developed in phases—from early work on probability and life tables, to more formal ideas about inference, and then into the wide range of applications we see now. What I found most interesting was how statistics moved into fields like biology and the social sciences. It wasn't just about numbers—it became a way to understand people, populations, and behavior. I also got a better sense of the debates that shaped the field, like the difference between frequentist and Bayesian approaches.

Kendall's piece made me think more carefully about what really counts as statistics. Just collecting data isn't enough—what matters is using that data to learn something or make decisions. And from Hacking, I got the idea that the way people think about uncertainty actually changed over time. Probability became a new way of understanding and reasoning about the world, and that shift helped make modern statistics possible.

Overall, I've come to see statistics as something that developed through both practical needs and changes in how people thought. It's not just about dealing with numbers—it's a way of looking at problems, making sense of limited information, and drawing conclusions responsibly. Understanding the history behind it all gives me a new appreciation for the tools we use and the thinking that goes into them.

Week 2: The 1700s

For Week Two, I chose to focus on the topic "**The Gauss–Laplace Synthesis: From Inverse Probability to the Birth of Modern Inference**", based on Chapters 3 and 4 of Stephen Stigler's The History of Statistics.

I. Introduction

These chapters explore a crucial moment in the development of statistical reasoning, when two leading figures—Pierre-Simon Laplace and Carl Friedrich Gauss—approached the problem of estimation and uncertainty from different angles but ultimately arrived at a unified framework. Their work brought together inverse probability, error modeling, and the method of least squares, forming the intellectual foundation for what we now think of as statistical inference. Stigler (1986) describes this convergence as the Gauss–Laplace synthesis, highlighting its role in transforming scattered probabilistic ideas into a general method for drawing conclusions from data.

This synthesis was not just a technical achievement; it marked a shift in how scientists understood the relationship between theory and observation. Laplace extended Bayes' theorem to real-world problems using prior and posterior probabilities, while Gauss offered a justification for least squares based on the assumption of normally distributed errors. Their different philosophical perspectives—Laplace's more Bayesian view and Gauss's more frequentist reasoning—nonetheless led to the same estimation procedure, demonstrating the growing power and flexibility of statistical methods. Understanding how these ideas came together helps explain how modern inference was born, and why statistics today still carries traces of both traditions.

II. Bayes and Laplace: The Rise of Inverse Probability

The roots of inverse probability can be traced to Thomas Bayes, whose posthumously published essay in 1763 introduced a method for reasoning from observed outcomes back to underlying causes. While Bayes' original work had limited impact at the time, it quietly laid the foundation for what would later become known as Bayesian inference. It was Laplace, however, who recognized the broader potential of this idea and turned it into a general method of statistical reasoning.

Beginning in the 1770s, Laplace developed and extended Bayes' framework to a wide range of applications, including astronomy, demography, and social statistics. His central idea was to use posterior probabilities to estimate unknown parameters—combining a prior belief (often taken as uniform) with observed data. In his 1774 memoir and several later works, Laplace derived mathematical formulations for estimating the probability of a cause given observed effects (Stigler, 1986, pp. 66–69). This was the essence of inverse probability: moving from data to hypotheses about the underlying mechanism.

One of Laplace's most important contributions was to apply this reasoning to real scientific problems, such as estimating the mass of celestial bodies or the likelihood of a population trend. He justified his use of uniform priors by assuming no reason to favor one value over another—a principle that would later become controversial. While this approach made the mathematics tractable, it also opened Laplace to later critiques about subjectivity and the lack of rigor in prior selection (Stigler, 1986, pp. 71–73).

Still, Laplace's version of inverse probability was a major step forward. It introduced a structured way to incorporate uncertainty into estimation, allowing scientists to express degrees of belief in different outcomes based on both data and prior assumptions. As Stigler notes, this work marked a shift from viewing probability as a tool for gambling or games of chance to treating it as a general framework for scientific inference.

Laplace's ideas, while mathematically sophisticated, were also remarkably practical. He saw probability as a "branch of the most important of sciences," and believed it could bring order to problems of judgment and decision-making. His early application of inverse probability to areas like voting patterns and population trends shows how he viewed data and uncertainty as central to public reasoning, not just natural science.

III. The Central Limit Theorem and the Curve of Errors

As Laplace continued developing his theory of inverse probability, he also turned his attention to the mathematical behavior of errors in repeated measurements. In doing so, he made one of his most lasting contributions to statistical theory: the formalization of the Central Limit Theorem (CLT). This result provided a crucial mathematical foundation for inference by showing that the sum (or average) of a large number of independent, small errors tends toward a normal distribution, regardless of the shape of the original error distribution (Stigler, 1986, pp. 78–80).

Laplace’s use of the CLT allowed him to make probabilistic claims about estimation errors, which he then connected to his broader inverse probability framework. By modeling the behavior of aggregate error, he showed that posterior distributions of estimated parameters would often be approximately normal in shape, especially as the number of observations increased. This insight gave further legitimacy to using the normal distribution—already known for its mathematical convenience—as a natural model for uncertainty in measurement.

In practical terms, this meant that confidence in parameter estimates could be quantified. Scientists could now calculate how likely it was for a true value to fall within a certain range of the observed average. This interpretation linked the frequency of observed outcomes with the uncertainty of inferences, deepening the connection between probability theory and statistical estimation.

To make the CLT usable, Laplace had to develop tools for approximating complex probability expressions. He introduced what is now known as the Laplace approximation, a method for simplifying integrals involving exponential functions by focusing on their peak value. This technique became a workhorse of applied probability and Bayesian analysis, reinforcing Laplace’s practical influence on the growing field of statistics (Stigler, 1986, pp. 81–82).

Laplace also attempted to model the curve of observational errors, trying to determine its exact form. While he initially favored a double exponential distribution (also known as the Laplace distribution), he recognized the central role of the normal curve, especially in the context of the CLT. This paved the way for the normal distribution to become the dominant model for error in both theory and practice.

Through the CLT and his efforts to model errors, Laplace made it possible to treat measurement uncertainty mathematically, even when the underlying error structure was unknown. These developments were essential in supporting the method of least squares, which would soon be independently justified by Gauss. Laplace's work showed that estimates based on many small, independent observations could be both stable and predictable, even in the face of noise.

IV. Gauss’s Justification of Least Squares

While Laplace was developing a framework for statistical inference through inverse probability and the Central Limit Theorem, Carl Friedrich Gauss was independently building a different route to the same destination. In 1809, Gauss offered a probabilistic justification for the method of least squares, basing it not on prior distributions, but on the assumption that errors in observation follow a normal distribution. This marked a key moment in the emergence of what we now recognize as frequentist inference (Stigler, 1986, pp. 95–97).

Gauss's reasoning began with the idea that observational errors were random and symmetric, with small errors being more likely than large ones. By assuming a specific mathematical form for the error distribution—the normal curve—he was able to derive that the most probable value of a parameter (such as a planetary position) is the one that minimizes the sum of squared deviations between the observed and predicted values. This matched the practical technique already in use by astronomers and geodesists, but now it had a firm theoretical foundation.

This approach aligned with what we now call the likelihood principle, although that term did not exist at the time. Gauss considered the probability of the observed data as a function of the unknown parameters and selected the values that made the data most probable. This logic underpinned his defense of least squares as a statistically optimal method for parameter estimation under the normal error model.

In contrast to Laplace, Gauss made no use of prior distributions. His analysis assumed that the unknown quantity was fixed but unknown, and that inference should be based entirely on the structure of the observed data and its associated error distribution. This difference reflected a broader philosophical divide: Laplace's method relied on inverse probability and degrees of belief, while Gauss's reasoning treated probability as a way to model repeated measurement outcomes without introducing subjective assumptions.

Nonetheless, Gauss's justification is extremely influential. His work helped establish the normal distribution as a universal model for observational error, and least squares as the standard tool for extracting signal from noisy measurements. Gauss applied this approach extensively in astronomy, especially in orbit calculations, and its success contributed to the rapid adoption of least squares across the sciences (Stigler, 1986, pp. 97–99).

Gauss's contributions thus brought a new level of mathematical rigor and conceptual clarity to data analysis. While Laplace built a more general framework grounded in prior belief and posterior probability, Gauss offered a method that could be justified on purely observational grounds—as long as one assumed normality of error. Together, these approaches were not contradictory, but complementary, and would soon be seen as converging on a shared framework of inference.

V. The Gauss–Laplace Synthesis

By the early 19th century, the separate contributions of Laplace and Gauss had begun to converge, resulting in what Stigler (1986) refers to as the Gauss–Laplace synthesis—a foundational moment in which probability theory, estimation, and observational data were united into a general framework of statistical inference (Stigler, 1986, p. 104).

Both Laplace and Gauss had justified the method of least squares, but from very different starting points. Laplace approached the problem using inverse probability—deriving posterior distributions based on prior assumptions and observed data. His view emphasized degrees of belief and uncertainty about the unknown quantity, treating it as a random variable with a distribution that reflected the data and prior knowledge. Gauss, on the other hand, took a more frequentist route. He assumed the unknown parameter was fixed, and that observed variability arose from normally distributed random error, leading to the same least squares solution through maximum likelihood reasoning (Stigler, 1986, pp. 95–99, 104–105).

Despite their philosophical differences, the mathematical outcome was the same: under the assumption of normally distributed errors, the best estimate of an unknown parameter is the value that minimizes the sum of squared residuals. This convergence marked the first time a statistical method was both practically useful and theoretically justified from multiple perspectives. It allowed scientists to confidently apply least squares not only in astronomy and geodesy, but across a growing range of empirical sciences.

Importantly, this synthesis helped establish inference as a general principle, rather than a collection of domain-specific tricks. Whether grounded in subjective priors (Laplace) or objective likelihood (Gauss), the logic of using data to estimate unknown quantities—and to quantify uncertainty about those estimates—was now widely accepted. The method could be extended to multiple variables, complex models, and new areas of research.

Stigler emphasizes that this synthesis was not the end of the story, but the beginning of modern statistical thinking. It set the stage for future debates—between frequentist and Bayesian approaches—and for further developments like sampling theory, hypothesis testing, and regression modeling. But it was this early 19th-century moment, with Laplace and Gauss reaching the same practical conclusions through different reasoning, that gave statistics its first unified identity as a science of inference (Stigler, 1986, pp. 104–106).

VI. Implications and Legacy

The synthesis achieved by Laplace and Gauss in the early 19th century had lasting consequences for the development of statistics. By arriving at the same estimation method—least squares—through different conceptual routes, they demonstrated that probability theory could serve as a bridge between theory and data. This convergence marked a turning point: statistics was no longer a loose collection of techniques or a branch of probability for games and gambling. It has become a general framework for scientific reasoning under uncertainty.

The dual foundations laid by Laplace and Gauss continue to shape modern statistical practice. The Bayesian tradition, rooted in Laplace's use of prior and posterior probabilities, treats unknown parameters as random variables and emphasizes updating beliefs in light of new data. The frequentist tradition, following Gauss's likelihood-based reasoning, treats parameters as fixed and focuses on long-run properties of estimators. Though these schools of thought often appear in contrast, they both stem from the same core achievement: integrating probability with data-based estimation.

This historical moment also helped elevate the normal distribution to a central place in statistics. Whether through Laplace's Central Limit Theorem or Gauss's assumption about error, the idea that variability in measurements could be described by a mathematical curve became foundational. It legitimized the use of probabilistic models in real science—from astronomy and physics to social science and medicine.

Finally, the Gauss–Laplace synthesis did more than solve technical problems. It represented a shift in how scientists thought about knowledge itself: accepting that data is noisy, that truth is uncertain, and that inference must be grounded in probability. As Stigler (1986) makes clear, this synthesis was the beginning of modern statistical inference, and its legacy is still felt every time we fit a model, estimate a parameter, or draw a conclusion from incomplete information.

What I Learned

This week's readings helped me see how statistical inference started to come together in a more structured way. I was especially suprised by how Laplace and Gauss, despite having different views on probability, both arrived at the method of least squares as the best way to estimate unknown values from noisy data. Before this, I hadn't realized that such a fundamental method had two very different philosophical justifications behind it—one based on belief and prior probability, and the other based on error modeling and likelihood.

What stood out most to me was how Laplace used inverse probability to try to understand the causes behind observed effects. His use of posterior distributions and assumptions about prior knowledge felt surprisingly modern, even though they were developed over two centuries ago. At the same time, Gauss's approach made sense in a different way—by assuming normally distributed errors and focusing on the most likely estimate, he gave least squares a kind of objectivity that matched the needs of astronomers and scientists at the time.

I also learned how the Central Limit Theorem helped connect these ideas. It showed why the normal distribution kept showing up in these problems, and helped explain why methods like least squares work so well in practice. I often read and learned CLT in classrooms and on textbooks, but now I see how it played a much bigger role in building the foundations of statistics.

Overall, this week helped me understand how modern inference wasn't just invented—it was built piece by piece by people like Laplace and Gauss, each working from different motivations. It also made me realize that ideas like Bayesian vs. frequentist statistics go all the way back to this early period, and that both ways of thinking are deeply rooted in the history of the field.

References

Barnouw, J. (1979). Review of *The Emergence of Probability: A Philosophical Study of Early Ideas about Probability, Induction and Statistical Inference*, by I. Hacking. *Eighteenth-Century Studies, 12*(3), 438–443. https://www.jstor.org/stable/2738528

Fienberg, S. E. (1992). A brief history of statistics in three and one-half chapters. *Statistical Science, 7*(2), 208–225. https://doi.org/10.1214/ss/1177011363

Hacking, I. (2006). *The emergence of probability: A philosophical study of early ideas about probability, induction and statistical inference* (2nd ed., Introduction). Cambridge University Press. (Original work published 1975)

Kendall, M. G. (1960). Where shall the history of statistics begin? *Biometrika, 47*(3/4), 447–449. https://doi.org/10.1093/biomet/47.3-4.447

Stigler, S. M. (1986). *The history of statistics: The measurement of uncertainty before 1900* (Chaps. 1–4). Belknap Press of Harvard University Press.