Statistical Science
1986, Vol. 1, No. 3, 359–378

# Laplace's 1774 Memoir on Inverse Probability

**Stephen M. Stigler**

*Abstract.* Laplace's first major article on mathematical statistics was published in 1774. It is arguably the most influential article in this field to appear before 1800, being the first widely read presentation of inverse probability and its application to both binomial and location parameter estimation. After a brief introduction, an English translation of this epochal memoir is given.

*Key words and phrases:* History, Bayesian, posterior distribution, predictive distribution, optimum estimation, nuisance parameter, double exponential distribution.

## INTRODUCTION

The history of mathematical statistics before 1800 boasts several landmark treatises, such as Jacob Bernoulli's *Ars Conjectandi* (1713) and Abraham De Moivre's *Doctrine of Chances* (1718, second edition 1738), and a smaller number of important articles in the periodical literature. Of these latter, perhaps the single most influential was Pierre Simon Laplace's "Mémoire sur la probabilité des causes par les évènemens," published in 1774. Laplace was just 25 years old when this appeared, and it was his first substantial work in mathematical statistics. At the time he began this work in 1772, he was in a period of intensely creative scientific exploration, simultaneously making major advances in mathematics and in mathematical astronomy (Stigler, 1978), and the memoir is an explosion of ideas that left an indelible imprint on statistics. In this one article, we can recognize the roots of modern decision theory, Bayesian inference with nuisance parameters, and the asymptotic approximation of posterior distributions.

A full evaluation of Laplace's contributions in this memoir in an historical context is beyond the scope of this introduction. Some aspects of the memoir have been discussed by Todhunter (1865, pages 465–473), Gillispie (1981), Sheynin (1977), and Dale (1982), and further elucidation will be found in my book (Stigler, 1986, Chapter 3). Nonetheless, and despite the fact that the memoir reads clearly and smoothly (even after two centuries it seems like a contemporary work), a brief outline in modern terminology will be useful for most readers.

After some opening references in Section I to work of his own (Laplace, 1774a, 1776) and of Lagrange (1759) on the solution of difference equations (particularly such as arise in probability theory), Laplace goes on in Section II to treat the inference problems that are the central focus of the article. He clearly announces his first goal as that of determining an unknown binomial probability, given the outcome of $p + q$ trials, of which $p$ result in white tickets, $q$ in black. He states a Principle which we would now recognize as equivalent to Bayes's theorem with all causes being a priori equally likely. If $F$ is Laplace's "event" and $\theta_1, \theta_2, \ldots, \theta_n$ the $n$ causes, then his axiomatic "Principle" is:

$$\frac{P(\theta_i \mid F)}{P(\theta_j \mid F)} = \frac{P(F \mid \theta_i)}{P(F \mid \theta_j)}$$

and

$$P(\theta_i \mid F) = \frac{P(F \mid \theta_i)}{\sum_{j=1}^{n} P(F \mid \theta_j)}.$$

Some reasons why we can be reasonably certain Laplace was unaware of Bayes's earlier work can be found in Stigler (1978).

Indeed, one of the more compelling reasons is the very different approach taken: where Bayes gives a cogent argument why an a priori uniform distribution might be acceptable (Stigler, 1982), Laplace assumes the conclusion as an intuitively obvious axiom. Laplace gives one simple application of his Principle for the case of hypergeometric sampling from one of two

*Stephen M. Stigler is a Professor in the Department of Statistics, University of Chicago, 5734 University Avenue, Chicago, Illinois 60637.*

possible urns, $A$ and $B$, evaluating

$$K = P(\text{draw } f \text{ white, } h \text{ black} \mid \text{urn } A),$$

and finding

$$P(\text{urn } A \mid \text{draw } f \text{ white, } h \text{ black}) = K/(K + K')$$

(where $K' = P(\text{draw } f \text{ white, } h \text{ black} \mid \text{urn } B)$), the correct result if $A$ and $B$ are a priori equally likely.

In Section III, Laplace applies the Principle to the case of binomial sampling, effectively assuming a uniform prior distribution for the parameter. In Problem I, he derives the beta posterior of the binomial parameter $x$ and also the predictive distribution (Laplace's $E$) for $m + n$ future trials given $p + q$ trials have occurred, first for $m = 1$ and $n = 0$, then for the general case. He applies Stirling's formula in a form he obtained from a text of Leonhard Euler's to derive a large sample approximation to the latter. He does his asymptotics in two different ways. First, he supposes $p + q$ is large (while $m + n$ is not), and he concludes that the predictive distribution is approximately equal to the one found by simply taking $x$ equal to the sample fraction, $p/(p + q)$. He then shows that a very different conclusion holds if $m + n$ is large also, say $m + n = p + q$. The remainder of the section is taken up with an analysis of the posterior distribution for large samples. He states (in a surprisingly modern $\delta$-$\varepsilon$ form) a theorem that claims the posterior consistency of the relative frequency of successes, but his proof is a tour de force that does more: it both introduces Laplace's own method for the asymptotic approximation of integrals, by expanding the integrand about its maximum at $x = p/(p + q)$, and it effectively demonstrates the asymptotic normality of the beta posterior distribution, as he approximates the posterior probability

$$E = P(|x - p/(p + q)| < w \mid p \text{ white tickets, } q \text{ black})$$

by a normal integral, and then takes the limit as $p + q \to \infty$, with $w = (p + q)^{-1/n}$, $2 < n < 3$. Since the proof includes what seems to be the earliest evaluation of the definite integral

$$2 \int_0^\infty \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{z^2}{2\sigma^2}\right) dz = 1,$$

with $\sigma^2 = pq/(p + q)^3$ here, Laplace may have been the first to integrate the normal density. In 1809, Gauss was to refer to this evaluation as "an elegant theorem of Laplace." The section ends with an intricate attempt to approximate the error made in taking $E = 1$.

Section IV is given over to Problem II, an analysis of the classical "problem of points" from a Bayesian perspective. We might now describe this as finding the expectation of the exact predictive distribution for binomial sampling, with a uniform prior distribution.

In Section V, Laplace moves on to the estimation of a location parameter for the "simple" case of three observations. He was evidently spurred on by reading a brief footnote in a 1772 review by Jean Bernoulli III: "The problem of finding the true mean among several observations, which is rarely the arithmetic mean, is of considerable interest to astronomers." Bernoulli cited works by Boscovich and Lambert that had been published, and by Lagrange and Daniel Bernoulli that were only to appear in 1776 and 1778. (The entire footnote is reproduced, with references, in Stigler (1978, page 248).) Laplace begins his treatment of his problem ("Problem III") by speculating upon the nature of the error distribution and finding the general expression for the posterior distribution of what we would now call the location parameter. Figure 1 shows the three observations at $a$, $b$, and $c$, and Figure 2 shows the error distribution. $V$ represents the true value of the location parameter, and $p$ and $q$ the gaps between $a$ and $b$ and $b$ and $c$, respectively. We might now write $a = X^{(1)}$, $b = X^{(2)}$, $c = X^{(3)}$, $p = X^{(2)} - X^{(1)}$, $q = X^{(3)} - X^{(2)}$, and note that $p$ and $q$ are ancillary statistics. Laplace lets $x$ be the distance from $V$ to $a$, and Figure 1 gives a stylized version of the posterior distribution of $x$ given $p$ and $q$, the curve $HOL$. Laplace suggests both the posterior median ("the mean of probability") and the value that minimizes the posterior expected loss ("the mean of error") as a posterior estimate of the location parameter, and proves (with the aid of Figure 3) that these are always the same. This result (characterizing the posterior median as optimal for a certain loss function) is surely one of the earliest we can recognize as truly belonging to mathematical statistics, rather than probability theory. He then returns to the question of specifying the error distribution and presents an argument for the double exponential density,

$$\phi(x) = \frac{m}{2} e^{-m|x|}$$

He finds an explicit expression for the posterior median in this case, assuming the scale parameter $m$ is known, and he shows it differs from the commonly used arithmetic mean, as Jean Bernoulli III had said.

The remainder of Section V is concerned with the case where the scale parameter $m$ is unknown but a priori uniformly distributed. His analysis here is intricate, the more so because he makes a subtle error in finding an explicit expression for the posterior median. The crucial passages are in the two consecutive paragraphs that follow his correct derivation of the posterior distribution of $m$, the two paragraphs that begin "Next, if we denote by $y$ ..." Since the nature and importance of this error seem to have escaped other commentators, it is worth detailed comment. To help clarify the argument, I shall employ the symbol $f$ as generic notation for density, so for example $f(x, m \mid p, q)$ is the conditional joint density of $x$ and $m$ given $p$ and $q$. Recall that Laplace has framed his problem in terms of the correction, $x$, to be made to the first observation, $a$, and he wishes the posterior median of $x$ (which I shall denote $x_0$). Now, earlier, when he supposed $m$ known, Laplace had correctly given the equation that would determine the posterior median $x_0$ as, essentially,

$$\int_{-\infty}^{x_0} f(x, p, q \mid m)\,dx = \frac{1}{2}\int_{-\infty}^{\infty} f(x, p, q \mid m)\,dx,$$

which, since this joint density is proportional to the conditional density, $f(x \mid p, q, m)$, with constant of proportionality $[\int f(x, p, q \mid m)\,dx]^{-1}$, is equivalent to

$$\int_{-\infty}^{x_0} f(x \mid p, q, m)\,dx = \frac{1}{2}\int_{-\infty}^{\infty} f(x \mid p, q, m)\,dx.$$

Here, when $m$ is unknown, he has derived $f(m \mid p, q)$ under the assumption that $m$ is a priori uniformly distributed on $(0, \infty)$, as

$$f(m \mid p, q) \propto m^2 e^{-m(p+q)}(1 - \tfrac{1}{3}e^{-mp} - \tfrac{1}{3}e^{-mq}).$$

Now his $y = f(x \mid p, q, m)$, and Laplace "evidently" wants to proceed as follows:

We want $x_0$ so that

$$\int_{-\infty}^{x_0} f(x \mid p, q)\,dx = \frac{1}{2}\int_{-\infty}^{\infty} f(x \mid p, q)\,dx,$$

or equivalently,

$$\int_0^{\infty}\int_{-\infty}^{x_0} f(x, m \mid p, q)\,dx\,dm = \frac{1}{2}\int_0^{\infty}\int_{-\infty}^{\infty} f(x, m \mid p, q)\,dx\,dm,$$

or equivalently,

$$\int_0^{\infty}\int_{-\infty}^{x_0} f(x \mid m, p, q) f(m \mid p, q)\,dx\,dm = \frac{1}{2}\int_0^{\infty}\int_{-\infty}^{\infty} f(x \mid m, p, q) f(m \mid p, q)\,dx\,dm.$$

Now if this last equation were equivalent to

$$\int_0^{\infty}\int_{-\infty}^{x_0} f(x, p, q \mid m) f(m \mid p, q)\,dx\,dm = \frac{1}{2}\int_0^{\infty}\int_{-\infty}^{\infty} f(x, p, q \mid m) f(m \mid p, q)\,dx\,dm,$$

we would indeed have Laplace's solution. This would be true if

$$f(x \mid m, p, q) \propto f(x, p, q \mid m)$$

and the constant of proportionality did not depend on $m$, but here it does! In fact,

$$\begin{aligned} f(x \mid m, p, q) &= f(x, p, q \mid m)/f(p, q \mid m) \\ &\propto f(x, p, q \mid m)/f(m \mid p, q), \end{aligned}$$

for Laplace's prior, so the correct solution would have been

$$\int_0^{\infty}\int_{-\infty}^{x_0} f(x, p, q \mid m)\,dx\,dm = \frac{1}{2}\int_0^{\infty}\int_{-\infty}^{\infty} f(x, p, q \mid m)\,dx\,dm,$$

an even simpler pair of integrals than Laplace considered. Because of his error, Laplace is led to finding the root of a 15th degree equation, which he does iteratively, and presents the solution in the form of a table. No wonder he only considered the case of three observations! Laplace's error is important, as it sheds light on how he conceived of conditional distributions as only defined up to proportionality, and helps explain how his axiomatic "Principle" could have such intuitive appeal to him (Stigler, 1986, Chapter 3).

Finally, Section VI explores the effect of allowing a probability to have an a priori distribution in two classical settings: coin tossing and the rolling of a die. This work departs from the earlier sections of the paper in allowing a nonuniform prior; Laplace effectively "unsharpens" the sharp null hypothesis of a fair coin by first permitting the probability to be $\tfrac{1}{2}(1 - \pi)$ rather than $\tfrac{1}{2}$, and then allowing $\pi$ a uniform distribution over $[0, q^{-1}]$. He notes that the results (the expectation of the predictive distribution, in particular) can be dramatically different for composite events from those where a fair coin is assumed. His analysis for the die is similar, although more complex. It is interesting that he calls unbalanced dice "English Dice"; presumably they had a different nickname in London. He ends with some explicit derivations for the case of a three-sided die, concluding what must be the earliest application of Bayesian ideas to a multinomial setting.

The influence of this memoir was immense. It was from here that "Bayesian" ideas first spread through the mathematical world, as Bayes's own article (Bayes, 1764) was ignored until after 1780 and played no important role in scientific debate until the twentieth century (Stigler, 1982). It was also this article of Laplace's that introduced the mathematical techniques for the asymptotic analysis of posterior distributions that are still employed today. And it was here that the earliest example of optimum estimation can be found, the derivation and characterization of an estimator that minimized a particular measure of posterior expected loss. After more than two centuries, we mathematical statisticians cannot only recognize our roots in this masterpiece of our science, we can still learn from it.

## ACKNOWLEDGMENTS

Support for this research was provided in part by National Science Foundation Grant MCS-8301459. This manuscript was prepared using computer facilities supported in part by National Science Foundation Grant MCS-8404941 to the Department of Statistics at the University of Chicago.

## REFERENCES

BAYES, T. (1764). An essay towards solving a problem in the doctrine of chances. *Philos. Trans. Roy. Soc. London* **53** 370–418. (Reprinted, with an introduction by George Barnard, in 1958 in *Biometrika* **45** 293–315, and in 1970 in *Studies in the History of Statistics and Probability* (E. S. Pearson and M. G. Kendall, eds.) 131–153. Hafner, Darien, Conn.)

DALE, A. I. (1982). Bayes or Laplace? An examination of the origin and early applications of Bayes's theorem. *Arch. History Exact Sci.* **27** 23–47.

GILLISPIE, C. C. (1981). Laplace, Pierre Simon. *Dictionary of Scientific Biography* **15,** 273–403. Scribner, New York.

LAGRANGE, J. L. (1759). Sur l'integration d'une équation differentielle à différences finies, qui contient la théorie des suites récurrentes. *Miscellanea Philos. Math. Soc. Privatae Taurinenesis* **1** 33–42.

LAPLACE, P. S. (1774a). Mémoire sur les suites récurro-récurrentes et sur leurs usages dans la théorie des hasards. *Mémoires de mathématique et de physique presentés à l'Académie royale des sciences, par divers savans, & lûs dans ses assemblées* **6** 353–371. (Reprinted in Laplace's *Oeuvres complètes* **8** 5–24.)

LAPLACE, P. S. (1774b). Mémoire sur la probabilité des causes par les évènemens. *Mémoires de mathématique et de physique presentés à l'Académie royale des sciences, par divers savans, & lûs dans ses assemblées* **6** 621–656. (Reprinted in Laplace's *Oeuvres complètes* **8** 27–65.)

LAPLACE, P. S. (1775). Mémoire, sur les solutions particulières des équations differentielles, & sur les inégalités séculaires des plànetes. *Histoire de l'Académie royale des sciences, année 1772, première partie, avec les Mémoires de mathématique et de physique, pour la même année, tirés des registres de cette académie* (published 1775) 343–377. (Reprinted in Laplace's *Oeuvres complètes* **8** 325–366.)

LAPLACE, P. S. (1776). Recherches, sur l'intégration des équations différentielles aux différence finies, & sur leur usage dans la théorie des hasards. *Mémoires de mathématique et de physique presentés à l'Académie royale des sciences, par divers savans, & lûs dans ses assemblées, 1773* (published 1776) 37–162. (Reprinted in Laplace's *Oeuvres complètes* **8** 69–197.)

LAPLACE, P. S. (1886–1912). *Oeuvres complètes de Laplace.* 14 volumes. Gauthier-Villars, Paris.

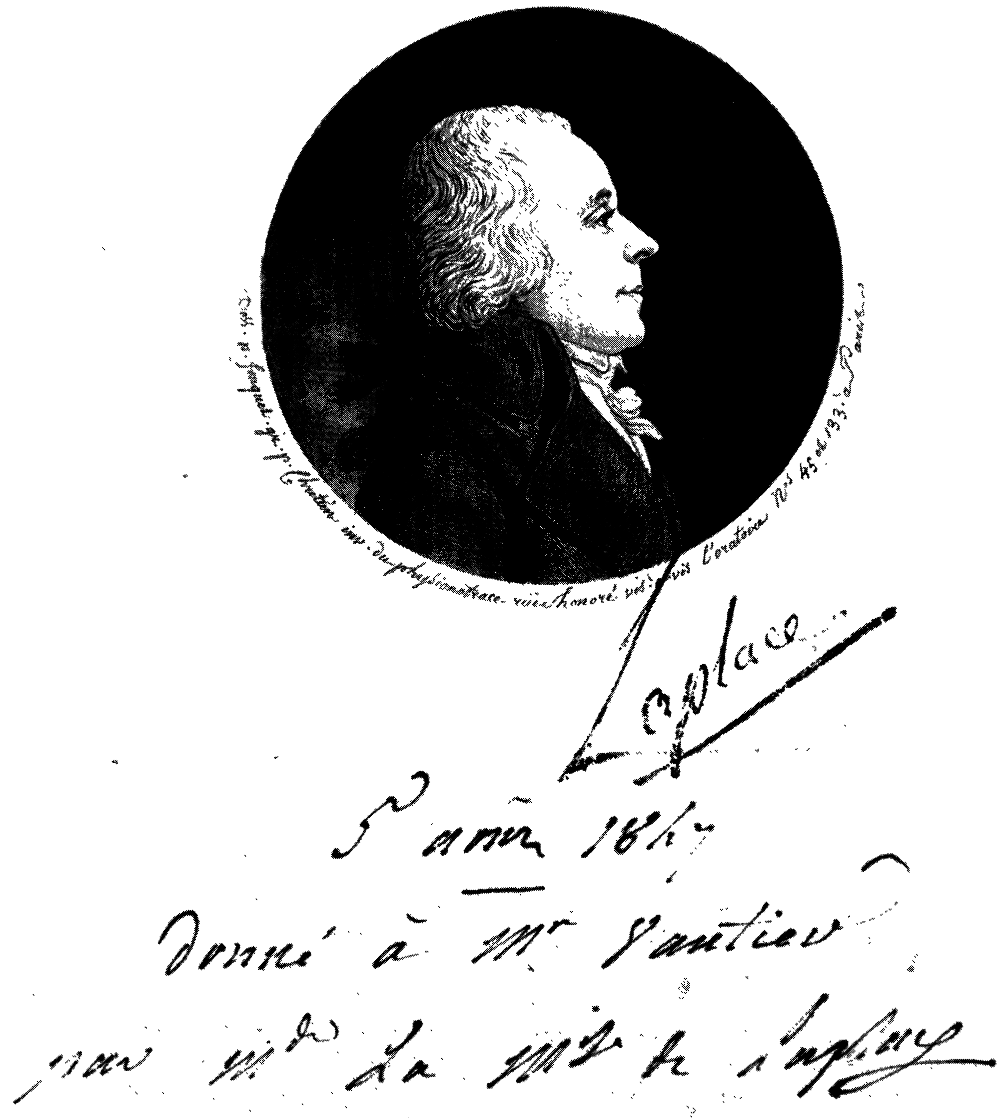
SHEYNIN, O. B. (1977). Laplace's theory of errors. *Arch. History Exact Sci.* **17** 1–61.

STIGLER, S. M. (1978). Laplace's early work: Chronology and citations. *Isis* **61** 234–254.

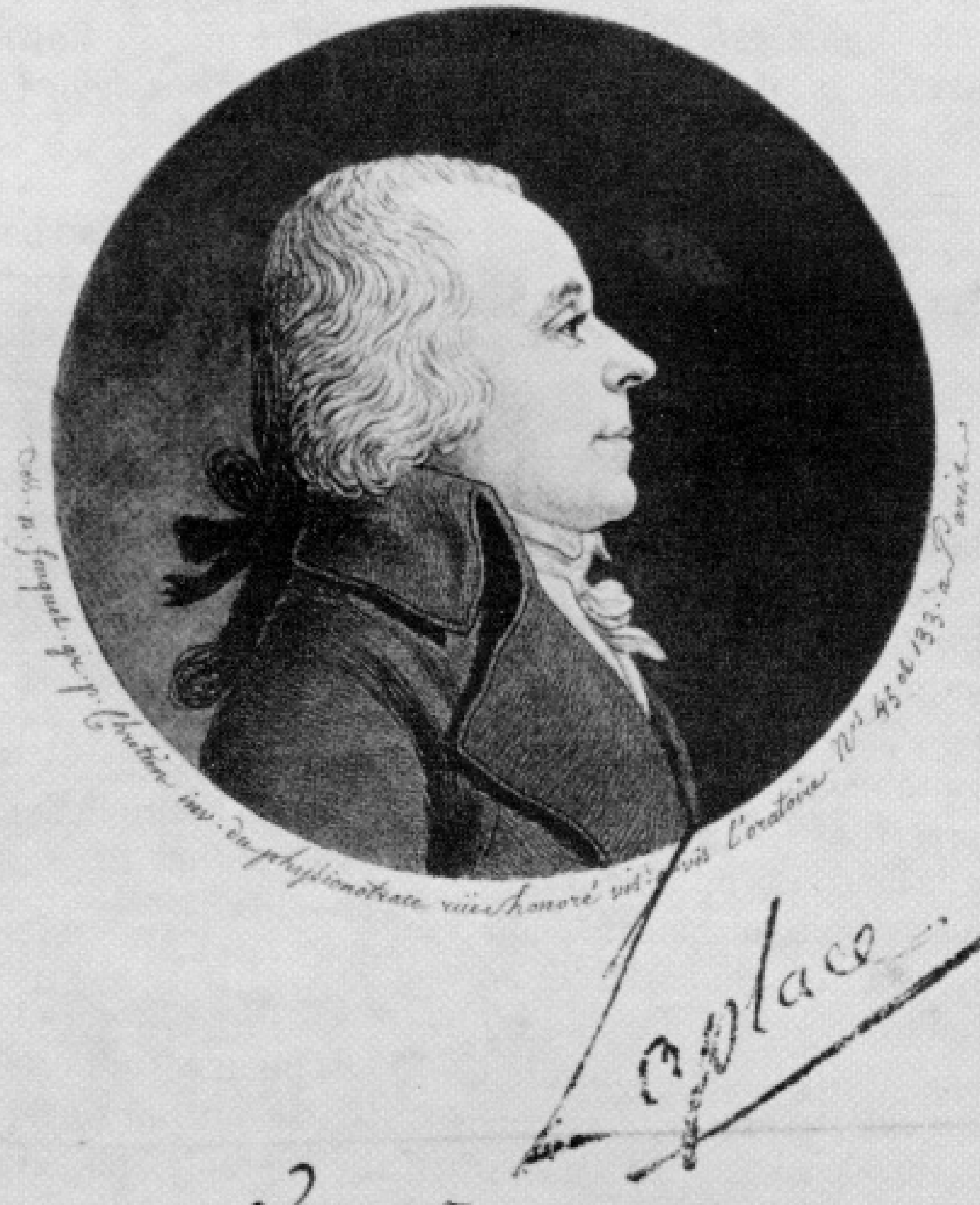
STIGLER, S. M. (1982). Thomas Bayes's Bayesian inference. *J. Roy. Statist. Soc. Ser. A* **143** 250–258.

STIGLER, S. M. (1986). *The History of Statistics: The Measurement of Uncertainty Before 1900.* Harvard Univ. Press, Cambridge, Mass.

TODHUNTER, I. (1865). *A History of the Mathematical Theory of Probability from the Time of Pascal to that of Laplace.* Macmillan, London. (Reprinted by Chelsea, New York, in 1949, 1965.)




*This lithograph of Laplace bears his printed signature and an 1847 autograph inscription by his widow. Laplace died 20 years earlier, in 1827, and the inscription ("Given to Mr. Vautier by Madame La Marquise de Laplace") suggests that it was given in response to a request for a remembrance. The face it shows is quite different from that shown in the stern formal pose that is usually reproduced, but the association of this picture with his widow suggests that it shows a likeness of Laplace recognizable to those who knew him well. The lithograph had been printed in 1799, and it thus shows Laplace at an age no more than 50.*




Laplace

5 avril 1817

donné à Mr Gautier
par Md La Mise de Laplace

# Memoir on the Probability of the Causes of Events

**Pierre Simon Laplace**

Translated from the original French by S. M. Stigler, University of Chicago. Originally published as "Mémoire sur la probabilité des causes par les évènemens," par M. de la Place, Professeur à l'École royal Militaire, in *Mémoires de Mathématique et de Physique, Presentés à l'Académie Royale des Sciences, par divers Savans & lûs dans ses Assemblées, Tome Sixième* (1774) 621–656. Reprinted in Laplace's Oeuvres complètes **8** 27–65.

## I

The theory of chances is one of the most curious and most difficult parts of analysis, due to the delicacy of the problems it produces and the difficulty of submitting them to calculation. It appears to have been treated with the most success by M. Moivre, in an excellent work entitled *Theory of Chances*. We owe to this able geometer the first research that was done on the integration of differential equations by means of finite differences. The method that he invented for that purpose is very ingenious, and he applied it quite successfully to solve many problems concerning probabilities, but one must agree that the point of view from which he considered the matter is indirect. Finite difference equations are susceptible to the same considerations as those involving infinitely small differences, and should be treated in an analogous manner; the sole difference being that, in the case of infinitely small differences, one can neglect certain quantities that it is not permissible to discard in the case of finite differences. It is this that renders the integration of the latter a thorny problem; the illustrious M. de la Grange is the first who has treated them in this analogous manner, in an elegant memoir that can be found in the first volume of those of Turin. The theory of finite difference equations is of the greatest use in the science of probabilities, and it is only by this means that one can hope for a general method to subject this science to analytical treatment.

In seeking to solve many problems concerning chances in this manner, I have frequently encountered a type of finite difference equation quite different from those previously considered. One can regard them as partial difference equations. Their importance to the analysis of chances is such that I devoted a memoir to a particular manner of treating them, *On recurro-recurrent series*, printed in this volume. In reconsidering this subject, however, it appeared to me that it was of such great utility in the science of chances that a much more general means of treating them was required than had been previously found. This consideration led to a much deeper study of the whole of the theory of the integration of finite differential equations. I have treated this in a memoir that I have read to the Academy, entitled: *Researches on the integration of finite difference equations, and on their uses in the analysis of chances*. That memoir will appear in the Academy's volume for the year in which I read it, 1773. The object of this present memoir is quite different. I propose to determine the probability of the causes of events, a question which has not been given due consideration before, but which deserves even more to be studied, for it is principally from this point of view that the science of chances can be useful in civil life.

## II

The uncertainty of human knowledge is concerned with events or with causes of events. If one is assured, for example, that an urn only contains white and black tickets in a given ratio, and one asks the probability that a ticket drawn by chance will be white, then the event is uncertain but the cause upon which the probability of its occurrence depends, the ratio of white to black tickets, is known.

In the following Problem, *an urn is supposed to contain a given number of white and black tickets in an unknown ratio; if one draws a ticket and finds it white, determine the probability that the ratio of white to black tickets is that of $p$ to $q$. The event is known and the cause is unknown.*

One can formulate all problems of the theory of chances as belonging to these two classes; we shall only discuss those of the second class here. To this end, we establish the following principle.

### Principle

If an event can be produced by a number $n$ of different causes, the probabilities of these causes given the event are to each other as the probabilities of the event given the causes, and the probability of the existence of each of these is equal to the probability of the event given that cause, divided by the sum of all the probabilities of the event given each of these causes.

The following question will clarify this principle, as well as being useful itself. I suppose that I am presented two urns, $A$ and $B$, of which the first contains $p$ white tickets and $q$ black tickets, and the second contains $p'$ white tickets and $q'$ black tickets. I take from one of these urns (I do not know which) $f + h$ tickets, of which $f$ are white and $h$ are black. We ask, what is the probability that the urn from which I've drawn the tickets is $A$, or that it is $B$?

If we suppose that this urn was $A$, the probability of drawing $f$ white tickets and $h$ black tickets is

$$\frac{(f+h)!\,(p+q-f-h)!\,p!\,q!}{f!\,h!\,(p-f)!\,(q-h)!\,(p+q)!}.$$

Let $K$ be this quantity. If we suppose now that the urn from which I have taken the tickets is $B$, the probability of drawing $f$ white tickets and $h$ black tickets can be determined by replacing $p$ and $q$ by $p'$ and $q'$ in $K$; let $K'$ be the resulting expression. Then the probbilities that the urn from which I have drawn the tickets is $A$ or $B$, are by the principle given above, as $K:K'$. The probability that the urn is $A = K/(K+K')$ and that it is $B = K'/(K+K')$.

We shall now apply this principle to the solution of several problems.

## III

### Problem I

*If an urn contains an infinity of white and black tickets in an unknown ratio, and we draw p + q tickets from it, of which p are white and q are black, then we require the probability that when we draw a new ticket from the urn, it will be white.*

SOLUTION. The ratio of the number of white tickets to the total number of tickets contained in the urn can be any fraction from 0 up to 1. Now, if we take $x$ as representing this unknown ratio, the probability of drawing $p$ white tickets and $q$ black tickets from the urn is $x^p\,(1-x)^q$. Therefore, the probability that $x$ is the true ratio of the number of white tickets to the total number of tickets is, by the principle of the preceding section,

$$= \frac{x^p(1-x)^q\,dx}{\int x^p(1-x)^q\,dx},$$

the integral being taken from $x = 0$ to $x = 1$. Now, under the supposition that $x$ is the true ratio of white tickets to the total number of tickets, the probability of drawing a white ticket from the urn is $x$; if we now multiply this quantity by the probability of the supposition, we will have for the probability of drawing a white ticket from the urn with true ratio $x$,

$$\frac{x^{p+1}\cdot dx(1-x)^q}{\int x^p\cdot dx(1-x)^q}.$$

Consequently, if we call $E$ the total probability of drawing a white ticket from the urn, we will have

$$E = \frac{\int x^{p+1}\cdot dx(1-x)^q}{\int x^p\cdot dx(1-x)^q},$$

where the integrals begin at $x = 0$ and end at $x = 1$.

It is easy, from these two expressions, to find a quite simple expression for $E$. We have

$$\begin{aligned}\int x^{p+1}\,dx(1-x)^q \\ &= \frac{q}{p+2}\int x^{p+2}\,dx(1-x)^{q-1} \\ &= \frac{q(q-1)}{(p+2)(p+3)}\int x^{p+3}\,dx(1-x)^{q-2}\end{aligned}$$

and so forth, therefore

$$\int x^{p+1}\,dx(1-x)^q = \frac{1\cdot 2\cdot 3\cdots q}{(p+2)(p+3)\cdots(p+q+2)},$$

similarly

$$\int x^p\,dx(1-x)^q = \frac{1\cdot 2\cdot 3\cdots q}{(p+1)\cdots(p+q+1)},$$

thus

$$E = \frac{p+1}{p+q+2}.$$

If we had sought the probability of drawing $m$ white tickets and $n$ black tickets from the urn, we would have found

$$E = \frac{\int x^{p+m}\,dx(1-x)^{q+n}}{\int x^p\,dx(1-x)^q},$$

from which we get $E =$

$$\frac{(q+1)(q+2)\cdots(q+n)(p+1)(p+2)\cdots(p+q+1)}{(p+m+1)(p+m+2)\cdots(p+q+m+n+1)}.$$

If $p$ and $q$ are quite large, we can simplify this expression in the following manner. I observe that we have

$$\ln 1 + \ln 2 + \ln 3 + \cdots + \ln x = \tfrac{1}{2}\ln 2\pi + (x+\tfrac{1}{2})\ln x - x + \tfrac{1}{12x} - \&c,$$

$\pi$ being the ratio of a half-circumference to a radius (see les Institutions du Calcul différentiel of M. Euler), and in the following we shall let $e$ be the number whose hyperbolic logarithm is unity. Then, supposing $p$ and $q$ are very large numbers,

$$(q+1)(q+2)\cdots(q+n) = \frac{1\cdot 2\cdot 3\cdots(q+n)}{1\cdot 2\cdot 3\cdots q} = \frac{(q+n)^{q+n+1/2}}{e^n q^{q+1/2}};$$

similarly

$$(p+1)\cdots(p+q+1) = \frac{(p+q+1)^{p+q+1/2}}{e^{q+1}p^{p+1/2}}$$

and

$$(p+m+1)\cdots(p+q+m+n+1) = \frac{(p+q+m+n+1)^{p+q+m+n+1/2}}{e^{q+n+1}\cdot(p+m)^{p+m+1/2}}.$$

Thus,

$$E = \frac{(p+q+1)^{p+q+1/2}(p+m)^{p+m+1/2}(q+n)^{q+n+1/2}}{q^{q+1/2}\cdot p^{p+1/2}\cdot(p+q+m+n+1)^{p+q+m+n+1/2}}.$$

We note here that

$$(p+q+1)^{p+q+1/2} = e(p+q)^{p+q+1/2},$$

because

$$\left(1+\frac{1}{p+q}\right)^{p+q+1/2} = e,$$

supposing $p+q$ infinitely large. Apparently then if we suppose $m$ and $n$ quite small relative to $p$ and $q$, we will have

$$(p+m)^{p+m+1/2} = e^m p^{p+m+1/2},$$

$$(q+n)^{q+n+1/2} = e^n q^{q+n+1/2},$$

and

$$(p+q+m+n+1)^{p+q+m+n+1/2} = e^{m+n+1}(p+q)^{p+q+m+n+1/2},$$

and so we will have

$$E = \frac{p^m q^n}{(p+q)^{m+n}}.$$

From this we can conclude, $p$ and $q$ being supposed quite large, that, as long as $m$ and $n$ are very much less, we can without fearing any appreciable error calculate the probability of drawing white and black tickets from the urn under the supposition that in this urn the ratio of the number of white tickets is to that of the black tickets as $p:q$. This supposition, however, becomes false when $m$ and $n$ are quite large, and it seems to me essential to note this. To make this clear, suppose $m = p$ and $n = q$; then we have

$$E = \sqrt{1/2}\,\frac{p^m q^n}{(p+q)^{n+n}} = 0.7071\,\frac{p^m q^n}{(p+q)^{m+n}}.$$

This expression, as we see, differs from the one,

$$E = \frac{p^m q^n}{(p+q)^{m+n}}$$

that we arrive at in taking $p/(p+q)$ as the ratio of the number of white tickets to the total number of tickets contained in the urn.

The solution to this problem gives a direct method for determining the probability of future events from those which have already occurred, but this is quite a broad subject, and I shall limit myself here to giving a rather singular proof of the following theorem.

*One can suppose that the numbers $p$ and $q$ are so large that it becomes as close to certainty as one wishes that the ratio of the number of white tickets to the total number of tickets contained in the urn is included between the two limits $p/(p+q) - w$ and $p/(p+q) + w$. $w$ can be supposed less than any given quantity.*

In order to prove this theorem, I observe that the probability of the ratio $x$ is, by the preceding, equal to

$$\frac{(p+1)(p+2)\cdots(p+q+1)}{1\cdot 2\cdot 3\cdots q}\cdot x^p dx(1-x)^q.$$

Let

$$x = \frac{p}{p+q} + z,$$

and we have

$$\int x^p\,dx(1-x)^q = \frac{p^p q^q}{(p+q)^{p+q}}\int dz\left(1+\frac{p+q}{p}z\right)^p\cdot\left(1-\frac{p+q}{q}z\right)^q.$$

If we integrate this quantity from $z = 0$ to $z = w$, and multiply this integral by

$$\frac{(p+1)\cdots(p+q+1)}{1\cdot 2\cdot 3\cdots q},$$

we will have the probability that the ratio of the number of white tickets to the total number of tickets is included between the limits $p/(p+q)$ and $p/(p+q)+w$.

Similarly, if we integrate

$$\frac{p^p q^q}{(p+q)^{p+q}}\int dz\left(1-\frac{p+q}{p}z\right)^p\cdot\left(1+\frac{p+q}{q}z\right)^q$$

from $z = 0$ to $z = w$, and multiply this integral by $(p+1)\cdots(p+q+1)/(1\cdot 2\cdot 3\cdots q)$, we will have the probability that the ratio of the number of white tickets to the total number of tickets is included between the limits $p/(p+q)$ and $p/(p+q) - w$. The sum of these two quantities then gives the probability that this ratio is contained within the limits $p/(p+q) - w$ and $p/(p+q) + w$. Call this probability $E$, and suppose that $p$ and $q$ are infinitely large and that $w$, the largest value of $z$, is infinitely less than $1/\sqrt[3]{p+q}$ and infinitely larger than $1/\sqrt{p+q}$, that it is equal, for example, to $1/(p+q)^{1/n}$, $n$ being larger than 2 and less than 3.

If we now set $\ln(1-((p+q)/p)z)^p = u$, we will have, expanding in a series,

$$u = -(p+q)z - \frac{(p+q)^2}{2p}zz - \frac{(p+q)^3}{3p^2}z^3 - \&\text{c.}$$

Now

$$\left(1-\frac{p+q}{p}z\right)^p = e^u$$
$$= \exp\left(-(p+q)z - \frac{(p+q)^2}{2p}zz - \&\text{c.}\right)$$

We can neglect the term $-[(p+q)^3/3p^2]z^3$ and succeeding terms, because the largest value of $z$, being by supposition equal to $1/(p+q)^{1/n}$, gives $\exp[-((p+q)^3/3p^2)z^3]$ equal to $\exp[-((p+q)^{3-3/n}/3p^2)]$. In the case where $e$ will have the largest negative exponent, as when $n$ is less than 3, this exponent is clearly infinitely small, and therefore we can suppose

$$\exp\left(-\frac{(p+q)^3}{3p^2}z^3 - \&\text{c}\right)$$

equal to unity. We will similarly have

$$\left(1+\frac{p+q}{q}z\right)^q = \exp\left((p+q)z - \frac{(p+q)^2}{2q}zz\right).$$

From this we can easily conclude

$$E = \frac{(p+1)\cdots(p+q+1)}{1\cdot 2\cdot 3\cdots q}$$
$$\cdot\frac{p^pq^q}{(p+q)^{p+q}}\int 2\,dz\cdot\exp\left(-\frac{(p+q)^3}{2pq}zz\right).$$

Since

$$\frac{(p+1)\cdots(p+q+1)}{1\cdot 2\cdot 3\cdots q} = \frac{(p+q)^{p+q+1/2}}{p^{p+1/2}q^{q+1/2}\sqrt{2\pi}},$$

$$E = \frac{(p+q)^{3/2}}{\sqrt{2\pi}\sqrt{pq}}\int 2\,dz\cdot\exp\left(-\frac{(p+q)^3}{2pq}zz\right).$$

Let $-[(p+q)^3/2pq]zz = \ln\mu$, and we will have

$$\int 2\,dz\exp\left(-\frac{(p+q)^3}{2pq}zz\right)$$
$$= -\frac{\sqrt{2qp}}{(p+q)^2}\int\frac{d\mu}{\sqrt{-\ln\mu}}.$$

The number $\mu$ can here have any value between 0 and 1, and, supposing the integral begins at $\mu = 1$, we need its value at $\mu = 0$. This may be determined using the following theorem (see M. Euler's Calcul intégral). Supposing the integral goes from $\mu = 0$ to $\mu = 1$, we have

$$\int\frac{\mu^n\,d\mu}{\sqrt{(1-\mu^{2i})}}\cdot\int\frac{\mu^{n+i}\,d\mu}{\sqrt{(1-\mu^{2i})}} = \frac{1}{i(n+1)}\cdot\frac{\pi}{2},$$

whatever be $n$ and $i$. Supposing $n = 0$ and $i$ is infinitely small, we will have $(1-\mu^{2i})/(2i) = -\ln\mu$, because the numerator and the denominator of this quantity become zero when $i = 0$, and if we differentiate them both, regarding $i$ alone as variable, we will have $(1-\mu^{2i})/(2i) = \ln\mu$, therefore $1-\mu^{2i} = -2i\ln\mu$. Under these conditions we will thus have

$$\int\frac{\mu^n\,d\mu}{\sqrt{(1-\mu^{2i})}}\int\frac{\mu^{n+i}\,d\mu}{\sqrt{(1-\mu^{2i})}}$$
$$= \int\frac{d\mu}{\sqrt{2i}\sqrt{-\ln\mu}}\int\frac{d\mu}{\sqrt{2i}\sqrt{-\ln\mu}} = \frac{1}{i}\frac{\pi}{2};$$

Therefore

$$\int\frac{d\mu}{\sqrt{-\ln\mu}} = \sqrt{\pi}$$

supposing the integral is from $\mu = 0$ to $\mu = 1$. In our case, however, the integral is from $\mu = 1$ to $\mu = 0$, and we will have

$$\int -\frac{d\mu}{\sqrt{-\ln\mu}} = \sqrt{\pi}.$$

Thus

$$\int 2\,dz\exp\left(-\frac{(p+q)^3}{2pq}zz\right) = \frac{\sqrt{pq}\sqrt{2\pi}}{(p+q)^{3/2}},$$

from which we obtain $E = 1$. We see, then, that, neglecting infinitely small quantities, we can consider it certain that the ratio of the number of white tickets to the total number of tickets is between the limits $p/(p+q)+w$ and $p/(p+q)-w$, where $w$ is equal to $1/\sqrt[n]{(p+q)}$ and $n$ is greater than 2 and less than 3, a fortiori when $n$ is greater than 3 and therefore $w$ can be supposed smaller than any given quantity.

Suppose now that we wish to determine the error made in setting $E = 1$, when we give $z$ a very small value $w$. Here is a quite simple means of obtaining it.

We need to integrate

$$\int dz\left(1+\frac{p+q}{p}z\right)^p\cdot\left(1-\frac{p+q}{q}z\right)^q$$

from $z = 0$ to $z = w$. Let $K$ be this integral from $z = 0$ to $z = q/(p+q)$, an integral that is clearly too large for our purpose, it being necessary to subtract the integral $\int dz\{1+[(p+q)/p]z\}^p\cdot\{1-[(p+q)/q]z\}^q$ from $z = w$ to $z = q/(p+q)$. Let $z = w+f$, and we have

$$\int dz\left(1+\frac{p+q}{p}z\right)^p\cdot\left(1-\frac{p+q}{q}z\right)^q$$

$$=\left(1+\frac{p+q}{p}w\right)^p\cdot\left(1-\frac{p+q}{q}w\right)^q$$

$$\cdot\int df\exp\left(-\frac{(p+q)^3wf-\&\text{c}}{pq-w(pp-qq)-w^2(p+q)^2}\right).$$

I note that $w$ can, by the preceding, be supposed infinitely larger than $1/\sqrt{p+q}$. Suppose $f$ infinitely less than $1/\sqrt{p+q}$, then we can neglect terms involving $f^2, f^3$, &c. in the exponent of $e$ and we will have

$$\int df\exp\left(-\frac{(p+q)wf}{pq-w(pp-qq)-w^2(p+q)^2}\right)$$

$$=\frac{pq-w(pp-qq)-w^2(p+q)^2}{(p+q)^3\,w}$$

$$\cdot\left[1-\exp\left(-\frac{(p+q)^3wf}{pq-w(pp-qq)-w^2(p+q)^2}\right)\right].$$

Suppose next that $wf$ is of an infinitely greater order than $1/(p+q)$, as is possible; then

$$\exp\left(-\frac{(p+q)^3wf}{pq-w(pp-qq)-w^2(p+q)^2}\right)$$

becomes negligible with respect to unity, and the preceding integral becomes, under the supposition that $w$ is very small, equal to $pq/(p+q)^3w$. We then will have

$$\int dz\left(1+\frac{p+q}{p}z\right)^p\cdot\left(1-\frac{p+q}{q}z\right)^q$$

$$=\left(1+\frac{p+q}{p}w\right)^p\cdot\left(1-\frac{p+q}{q}w\right)^q\cdot\frac{pq}{(p+q)^3w},$$

the integral being supposed to begin when $z=w$ and finish when $z=w+1/(p+q)^n$, $n$ being greater than ½. Now, the difference between this integral and the entire integral taken from $z=w$ to $z=q/(p+q)$ is infinitely less. To show this, I observe that if we denote, for short, the quantity

$$\left(1+\frac{p+q}{p}z\right)^p\cdot\left(1-\frac{p+q}{q}\right)^q$$

by $y$, we will have, when $z=w+1/(p+q)^n$, $n$ being larger than ½,

$$y=\left(1+\frac{p+q}{p}w\right)^p\cdot\left(1-\frac{p+q}{q}w\right)^q$$

$$\cdot\exp\left(-\frac{(p+q)^{3-n}w}{pq-w(pp-qq)-w^2(p+q)^2}\right).$$

If we increase $z$, $y$ becomes smaller, therefore $\int y\,dz$ from $z=w+1/(p+q)^n$ to $q/(p+q)$ is less than

$$\left(\frac{q}{p+q}-w-\frac{1}{(p+q)}\right)$$

$$\cdot\exp\left(-\frac{(p+q)^{3-n}w}{pq-w(pp-qq)-w^2(p+q)^2}\right).$$

Now, because we have supposed $w/(p+q)^n$ infinitely larger than $1/(p+q)$, the preceding quantity is infinitely less than $pq/(p+q)^3w$, since in general $e^{\infty 1/n}>\infty^m$ for any finite numbers $m$ and $n$.

We thus will have

$$\int dz\left(1+\frac{p+q}{p}z\right)^p\cdot\left(1-\frac{p+q}{q}z\right)^q$$

$$=K-\left(1+\frac{p+q}{p}w\right)^p$$

$$\cdot\left(1-\frac{p+q}{q}w\right)^q\cdot\frac{pq}{(p+q)^3w},$$

supposing the integral begins when $z=0$ and ends when $z=w$. Similarly, under the same conditions we will have

$$\int dz\left(1-\frac{p+q}{p}z\right)^p\cdot\left(1+\frac{p+q}{q}z\right)^q$$

$$=K'-\left(1-\frac{p+q}{p}w\right)^p$$

$$\cdot\left(1+\frac{p+q}{q}w\right)^q\cdot\frac{pq}{(p+q)^3w},$$

where $K'$ is obtained from $K$ by changing $p$ to $q$ and $q$ to $p$. As a consequence, we will have

$$E=\frac{(p+1)\cdots(p+q+1)}{1\cdot2\cdot3\cdots q}\cdot\frac{p^pq^q}{(p+q)^{p+q}}$$

$$\cdot\left\{\begin{array}{l}K-\left(1+\dfrac{p+q}{p}w\right)^p\cdot\left(1-\dfrac{p+q}{q}w\right)^q\cdot\dfrac{pq}{(p+q)^3w}\\[2ex] +K'-\left(1-\dfrac{p+q}{p}w\right)^p\cdot\left(1+\dfrac{p+q}{q}w\right)^q\cdot\dfrac{pq}{(p+q)^3w}.\end{array}\right.$$

But clearly,

$$\frac{(p+1)\cdots(p+q+1)}{1\cdot2\cdot3\cdots q}\cdot\frac{p^pq^q}{(p+q)^{p+q}}(K+K')=1,$$

from which we easily find

$$E = 1 - \frac{\sqrt{(pq)}}{w\sqrt{(2\pi)} \cdot (p+q)^{3/2}} \cdot \left\{ \begin{array}{l} \left(1 + \frac{p+q}{p} w\right)^p \cdot \left(1 - \frac{p+q}{q} w\right)^q \\ + \left(1 - \frac{p+q}{p} w\right)^p \cdot \left(1 + \frac{p+q}{q} w\right)^q. \end{array} \right.$$

By means of this formula, we can judge the error made in taking $E = 1$.

## IV

### Problem II

*Two players A and B, whose respective skills are unknown, play some game, for example piquet, where the first player to win a number n points receives a sum a deposited at the beginning of play. I suppose that the two players are forced to abandon play with player A lacking f points and player B lacking h points. In this situation, we ask how we should divide the sum a between the two players.*

SOLUTION. If the respective skills of the two players $A$ and $B$ were supposedly known, and they were in the ratio of $p$ to $q$, we would find, taking $p + q = 1$, that the sum that should be returned to $B$ equals

$$a \cdot q^{f+h-1} \cdot \left\{ \begin{array}{l} 1 + \frac{p}{q} \cdot (f+h-1) + \frac{p^2}{q^2} \\ \qquad \cdot \frac{(f+h-1) \cdot (f+h-2)}{1.2} + \&\text{c.} \\ \cdots + \frac{p^{f-1}}{q^{f-1}} \cdot \frac{(f+h-1) \cdots (h+1)}{1 \cdot 2 \cdot 3 \cdots (f-1)}. \end{array} \right.$$

This proposition is proved in many works. It can be quite easily deduced by the method of recurrorecurrent series, as can be seen in the Memoir cited at the beginning of this one; there the general solution to the problem of points in the case of three or more players is similarly deduced, a problem that had not previously been resolved by anyone, to my knowledge, although geometers who have worked on these matters have long desired the solution (see the second edition of the *Analyse des jeux de hasard* of M. Montmort, page 247).

Now, because the probability that $A$ will win a point is unknown, we may suppose it to be any unspecified number whatever between 0 and 1. Suppose that one of these numbers $x$ represents this probability; then the probability that of $2n - f - h$ points, $A$ wins $n - f$ and $B$, $n - h$, would be $x^{n-f}(1-x)^{n-h}$. It then follows from the principle of section II that the probability of the value we have supposed for $x$ is

$$\frac{x^{n-f}(1-x)^{n-h}\, dx}{\int x^{n-f}(1-x)^{n-h}\, dx},$$

the integral being taken beginning at $x = 0$ and ending at $x = 1$. Now, $x$ being supposed the probability that $A$ wins a point, we find that the sum that should be returned to $B$ is

$$a(1-x)^{f+h-1} \cdot \left\{ \begin{array}{l} 1 + \frac{x}{1-x}(f+h-1) \\ \quad + \frac{x^2}{(1-x)^2} \frac{(f+h-1) \cdot (f+h-2)}{1 \cdot 2} + \&\text{c.} \\ \cdots + \frac{x^{f-1}}{(1-x)^{f-1}} \cdot \frac{(f+h-1) \cdots (h+1)}{1 \cdot 2 \cdot 3 \cdots (f-1)}. \end{array} \right.$$

Thus the sum that should really be returned to player $B$ is

$$a \int x^{n-f}\, dx \cdot (1-x)^{f+h-1} \cdot \frac{\left[1 + \frac{x}{1-x}(f+h-1) \cdots \frac{x^{f-1}}{(1-x)^{f-1}} \frac{(f+h-1) \cdots (h+1)}{1 \cdot 2 \cdot 3 \cdots (f-1)}\right]}{\int x^{n-f}\, dx \cdot (1-x)^{n-h}},$$

both integrals being taken from $x = 0$ to $x = 1$. It is easy to show that

$$\int x^{n-f}\, dx \cdot (1-x)^{n-h} = \frac{1 \cdot 2 \cdot 3 \cdots (n-h)}{(n-f+1) \cdots (2n-f-h+1)}$$

similarly

$$\int x^{n-f}\, dx \cdot (1-x)^{f+h-1} = \frac{1 \cdot 2 \cdot 3 \cdots (f+h-1)}{(n-f+1) \cdots 2n},$$

and so forth. Then we will have for the sum that should be returned to $B$,

$$\frac{a(n-h+1) \cdots (h+f-1)}{(2n-f-h+2) \cdots 2n} \cdot \left\{ \begin{array}{l} 1 + \frac{(f+h-1)}{1} \cdot \frac{(n-f+1)}{(f+h-1)} \\ \quad + \frac{(f+h-1) \cdot (f+h-2)}{1 \cdot 2} \frac{(n-f+1)(n-f+2)}{(f+h-1)(f+h-2)} \\ \cdots + \frac{(f+h-1) \cdots (h+1)}{1 \cdot 2 \cdot 3 \cdots (f-1)} \frac{(n-f+1) \cdots (n-1)}{(f+h-1) \cdots (h+1)}. \end{array} \right.$$

## V

We can, by means of the preceding theory, solve the problem of determining the mean that one should take among many given observations of the same phenomenon. Two years ago I presented such a solution to the Academy, as a sequel to the Memoir "Sur les Séries récurrorécurrentes" printed in this volume, but it appeared to me to be of such little usefulness that I suppressed it before it was printed. I have since learned from Jean Bernoulli's astronomical journal that Daniel Bernoulli and Lagrange have considered the same problem in two manuscript memoirs that I have not seen. This announcement both added to the usefulness of the material and reminded me of my ideas on this topic. I have no doubt that these two illustrious geometers have treated the subject more successfully than I; however, I shall present my reflections here, persuaded as I am that through the consideration of different approaches, we may produce a less hypothetical and more certain method for determining the mean that one should take among many observations.

### Problem III

*Determine the mean that one should take among three given observations of the same phenomenon.*

SOLUTION. Let time be represented by a line $AB$ (Figure 1), and suppose that the first observation fixes the instant of the phenomenon at the point $a$, the second at the point $b$, and the third at the point $c$. Suppose further that the time unit is seconds, and that the interval from $a$ to $b$ is $p$ seconds and that from $b$ to $c$, $q$ seconds. We wish to find the point $V$ on the line $AB$ where we should fix the mean that we should take between the three observations $a$, $b$, and $c$.

For this we must observe that it is more probable that a given observation deviates from the truth by 2 seconds than by 3 seconds, by 3 seconds than by 4 seconds, &c. The law by which this likelihood diminishes as the difference between the observation and the truth increases is unknown, however. Suppose then (Figure 2) that the point $V$ is the true instant of the phenomenon, and that the probabilities that the observation differs from the truth by $VP$, $VP'$, &c. can be represented by a curve $RMM'$ which decreases according to an unspecified law. If we represent the abscissa $VP$ by $x$ and the corresponding ordinate $PM$ by $y$, then we shall write the equation of this curve as $y = \phi(x)$. This curve has the following properties.

1. It must be divided into two entirely similar parts by the line $VR$, because it is just as probable that the observation deviates from the truth to the right as to the left.

2. It must have the line $KP$ as an asymptote, because the probability that the observation differs from the truth by an infinite distance is evidently zero.

3. The entire area of this curve must be equal to one, because it is certain that the observation will fall on one of the points of the line $KP$.

Suppose now (Figure 1) that the true instant of the phenomenon is at the point $V$, at the distance $x$ from the point $a$. The probability that the three observations $a$, $b$, and $c$ deviate by the distances $Va$, $Vb$, and $Vc$ will be $\phi(x) \cdot \phi(p - x) \cdot \phi(p + q - x)$. If we suppose the true instant were at the point $V'$ and that $aV' = x'$, then this probability would be $= \phi(x') \cdot \phi(p - x') \cdot \phi(p + q - x')$. It follows then from our fundamental principle of section II that the probabilities that the true instant of the phenomenon is at the points $V$ or $V'$, are to each other as $\phi(x) \cdot \phi(p - x) \cdot \phi(p + q - x) : \phi(x') \cdot \phi(p - x') \cdot \phi(p + q - x')$. Thus if we construct a curve $HOL$ with the equation $y = \phi(x) \cdot \phi(p - x) \cdot \phi(p + q - x)$, the ordinates of this curve would represent the probabilities of the corresponding points on the abscissa.

In seeking the mean that we should choose among many observations, there are two objects we may have in mind.

The first is the instant such that it is equally probable that the true instant of the phenomenon falls before it or after it. We can call this instant the *mean of probability*.

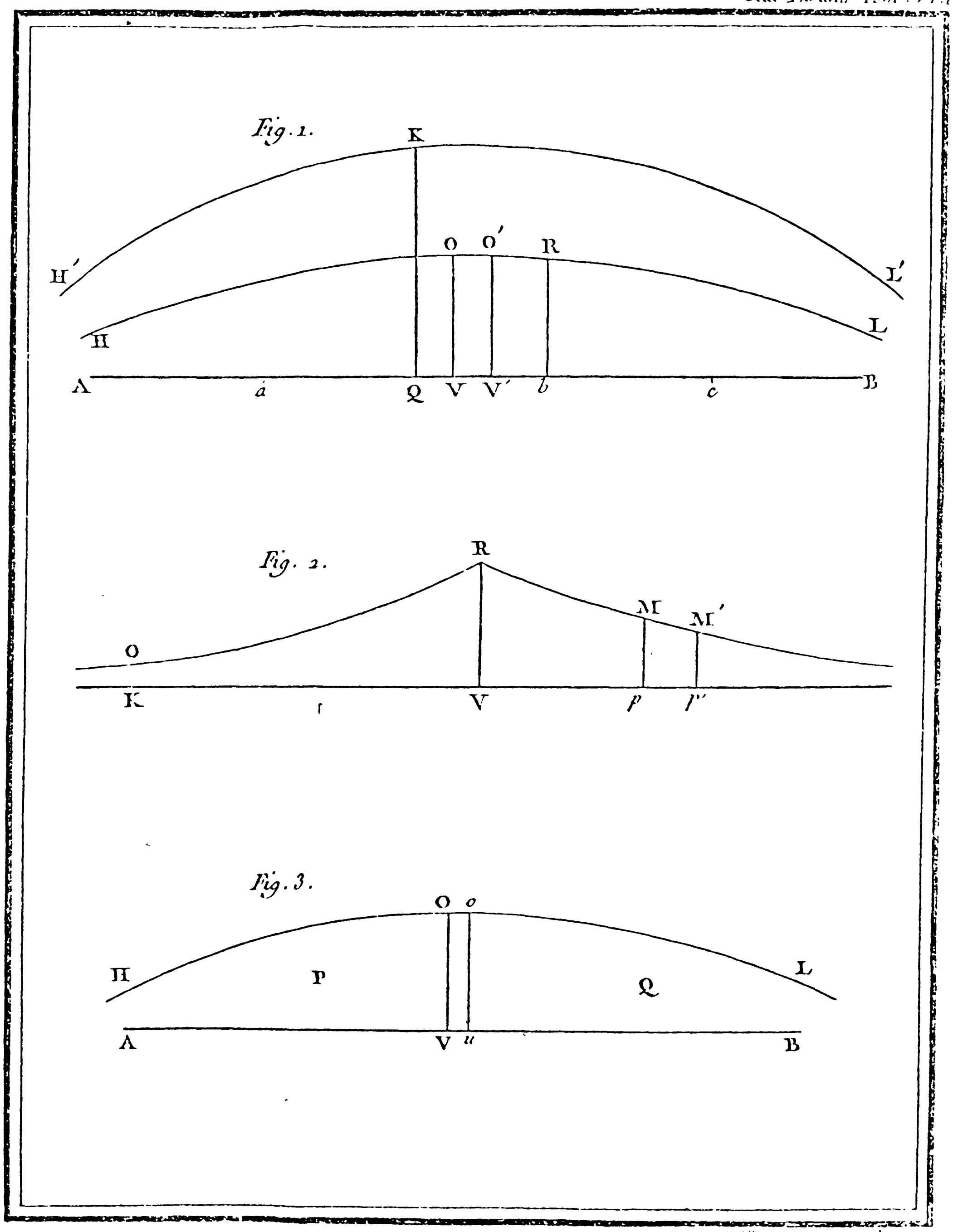
The second is the instant that *minimizes* the sum of the errors to be feared multiplied by their probabilities. We can call this the *mean of error* or *astronomical mean*, since it is that which astronomers should give preference to.

To find the first mean, it is necessary to determine the ordinate $OV$ which divides the area of the curve $HOL$ in two equal parts, since then it is clearly as probable that the true instant of the phenomenon falls to the right as to the left of the point $V$.

To find the second mean, it is necessary to choose (Figure 3) a point $V$ on the abscissa such that the sum of the ordinates of the curve $HOL$, multiplied by their distance from the point $V$, is a *minimum*. Now I claim that the second mean differs not at all from the first. To see this, introduce the ordinate $ou$ infinitely close to $OV$, and let $Vu = dx$, $OV = y$. Let $Q$ be the center of gravity of the part $uoL$ of the curve; let $M$ be the area of this part; $z$ the distance from the point $Q$ to the ordinate $OV$. Let $P$ be the center of gravity of the part $VOH$; let $N$ be the area of this part; and let $z'$ be the distance from $P$ to the ordinate $OV$. Then if we take the point $V$ as the mean, the sum of the ordinates multiplied by their distances from this point will be

$$Mz + Nz' + \tfrac{1}{2}ydx^2,$$

Sav. Etrang. Tom. VI. p. 656




Elle Haussard Sculp.

and if instead we take $u$ for the mean, the sum of the ordinates multiplied by their distance from the point $u$ will be

$$M(z - dx) + N(z' + dx) + \tfrac{1}{2}ydx^2.$$

Now we see that the difference of these two quantities will be $Ndx - Mdx$, which must be equal to zero in the case of a *minimum*. We will thus have in this case $M = N$; that is, the ordinate $OV$ will divide the area of this curve in two equal parts. We thus see that the *astronomical* mean differs not at all from that of *probability*, and that both are determined by the ordinate $OV$ that divides the area of the curve $HOL$ in two equal parts.

To find this ordinate, it is necessary to know $\phi(x)$. But of an infinite number of possible functions, which choice is to be preferred? The following considerations can determine a choice. It is true (Figure 2) that if we have no reason to suppose the point $P$ more probable than the point $P'$, we should take $\phi(x)$ to be constant, and the curve $ORM'$ will be a straight line infinitely near the axis $KP$. But this supposition must be rejected, because if we suppose there existed a very large number of observations of the phenomenon, it is presumed that they become rarer the farther they are spread from the truth. We can also easily see that this diminution cannot be constant, that it must become less as the observations deviate more from the truth. Thus not only the ordinates of the curve $RMM'$, but also the differences of these ordinates must decrease as they become further from the point $V$, which in this figure we always suppose to be the true instant of the phenomenon. Now, as we have no reason to suppose a different law for the ordinates than for their differences, it follows that we must, subject to the rules of probabilities, suppose the ratio of two infinitely small consecutive differences to be equal to that of the corresponding ordinates. We thus will have

$$\frac{d\phi(x + dx)}{d\phi(x)} = \frac{\phi(x + dx)}{\phi(x)}.$$

Therefore

$$\frac{d\phi(x)}{dx} = -m\phi(x),$$

which gives $\phi(x) = Ce^{-mx}$. Thus, this is the value that we should choose for $\phi(x)$. The constant $C$ should be determined from the supposition that the area of the curve $ORM$ equals unity, which represents certainty, which gives $C = \tfrac{1}{2}m$. Therefore $\phi(x) = (m/2)e^{-mx}$, $e$ being the number whose hyperbolic logarithm is unity.

One can object that this law is repugnant in that if $x$ is supposed extremely large, $\phi(x)$ will not be zero, but to this I reply that while $e^{-mx}$ indeed has a real value of all $x$, this value is so small for $x$ extremely large that it can be regarded as zero.

Now, accepting this law, we determine the area of the curve $HOL$ (Figure 1).

1. From $a$ to $b$, the ordinate of the curve $HOL$ is $y = (m^3/8)e^{-m(2p+q-x)}$. Therefore, the area of the curve in this interval will be $= (m^2/8) \cdot e^{-m(2p+1)}(e^{mx} - 1)$.

2. From $b$ to $c$, the ordinate of the curve will be $y = (m^3/8)e^{-m(x+q)}$, and the area of the curve in this interval will be $= (m^2/8)e^{-mq}(e^{-mp} - e^{-mx})$.

3. From $c$ to infinity, the area of the curve will be $= (m^2/3 \cdot 8)e^{-m(p+2q)}$.

4. From $a$ to infinity, on the side of $A$, the area of the curve will be $= (m^2/3 \cdot 8)e^{-m(q+2p)}$. The whole area of the curve is thus

$$=(m^2/4)e^{-m(p+q)}(1 - \tfrac{1}{3}e^{-mp} - \tfrac{1}{3}e^{-mq}).$$

We can observe that the point $V$, such that the ordinate $OV$ divides the area of the curve into two equal parts, must necessarily fall between the points $a$ and $b$, supposing $p > q$, or between $b$ and $c$, supposing $q > p$. This is because the area of the curve to the left of the ordinate $bR$ is

$$(m^2/8)e^{-m(p+q)}(1 - \tfrac{2}{3}e^{-mp}),$$

which is clearly greater or smaller than half the entire area according to whether or not $p$ is greater than or less than $q$. We suppose that $p$ is greater than $q$ in the following calculations; then to determine the distance $x$ of the point $a$ from the point $V$ where we should fix the true instant of the phenomenon, we will have the following equation.

$$m^2e^{-m(2p+q-x)} = m^2e^{-m(p+q)}(1 + \tfrac{1}{3}e^{-mp} - \tfrac{1}{3}e^{-mq}),$$

from which we find

$$x = p + (1/m)\ln(1 + \tfrac{1}{3}e^{-mp} - \tfrac{1}{3}e^{-mq}).$$

### Remark on the Method of Arithmetic Means

The method commonly in use among observers consists of taking an arithmetic mean of the three observations, which gives $x = (2p + q)/3$. Now this method follows from the preceding formulae with $m = 0$ or infinitely small, because then we have $\ln(1 + \tfrac{1}{3}e^{-mp} - \tfrac{1}{3}e^{-mq}) = \tfrac{1}{3}e^{-mp} - \tfrac{1}{3}e^{-mq}$. Now $\tfrac{1}{3}e^{-mp} = \tfrac{1}{3} - \tfrac{1}{3}mp$ and $\tfrac{1}{3}e^{-mq} = \tfrac{1}{3} - \tfrac{1}{3}mq$, so $1/m \ln(1 + \tfrac{1}{3}e^{-mp} - \tfrac{1}{3}e^{-mq}) = -\tfrac{1}{3}p + \tfrac{1}{3}q$. Therefore, $x = p + (1/m)\ln(1 + \tfrac{1}{3}e^{-mp} - \tfrac{1}{3}e^{-mq}) = (2p + q)/3$, the same value given by the method of arithmetic means.

The supposition that $m$ is infinitely small makes (Figure 2) all points on the line $KP$ equally probable, at least up to an extremely large distance, which is very unlikely both by the nature of the problem and by the result of calculation, as we shall see in a moment. This supposition may often be felt to be unnatural, and in delicate circumstances it may be necessary to make use of the following method.

If $m$ were known, it would be easy to find the value of $x$ by the above method, but as $m$ is unknown, recourse to other means of obtaining this value is required.

From the fundamental principle of section II, the probabilities of different values of $m$ are to each other as the probabilities that the three observations have the respective distances between them that they do, given the different values of $m$. Now, the probabilities that the three observations $a$, $b$, and $c$ (Figure 1) differ from one another by the distances $p$ and $q$, are to each other as the areas of the curves $HOL$ corresponding to different values of $m$, as we can easily verify. From this it follows from the principle of section II that the probability of $m$ is proportional to

$$m^2dm \cdot e^{-m(p+q)}(1 - \tfrac{1}{3}e^{-mp} - \tfrac{1}{3}e^{-mq}),$$

and we see from this that the probability that $m = 0$ or infinitely small (the supposition that leads to the method of arithmetic means) is infinitely less than that $m$ equals any finite quantity whatever.

Next, if we denote by $y$ the probability, corresponding to $m$, that the true instant of the phenomenon falls at a distance $x$ from the point $a$, the whole probability that this instant falls at this distance will be proportional to

$$\int ym^2\,dm \cdot e^{-m(p+q)} \cdot \left(1 - \frac{1}{3}e^{-mp} - \frac{1}{3}e^{-mq}\right),$$

the integral being taken from $m = 0$ to $m = \infty$. If we then construct a new curve $H'KL'$ on the axis $AB$ whose ordinates are proportional to this quantity, the ordinate $KQ$ which divides the area of this curve in two equal parts cuts the axis at the point we should take as the mean between the three observations.

The area of this new curve will evidently be proportional to the integral of the product of the area of the curve $HOL$ by

$$m^2dme^{-m(p+q)}(1 - \tfrac{1}{3}e^{-mp} - \tfrac{1}{3}e^{-mq}).$$

Then since in order to determine $x$ under a particular supposition for $m$ we have

$$m^2e^{-m(2p+q-x)} = m^2e^{-m(p+q)}(1 + \tfrac{1}{3}e^{-mp} - \tfrac{1}{3}e^{-mq}),$$

we will have

$$\int m^4\,dme^{-m(3p+2q-x)}\left(1 - \frac{1}{3}e^{-mp} - \frac{1}{3}e^{-mq}\right)$$

$$= \int m^4\,dme^{-m(2p+2q)}\left(1 + \frac{1}{3}e^{-mp} - \frac{1}{3}e^{-mq}\right) \cdot \left(1 - \frac{1}{3}e^{-mp} - \frac{1}{3}e^{-mq}\right),$$

where the integrals go from $m = 0$ to $m = \infty$.

To integrate these quantities we should observe that

$$\int m^4\,dme^{-Km}$$

$$= \frac{-1}{K}m^4e^{-Km} + \int \frac{4m^3}{K}\,dme^{-KM}$$

$$= \frac{-1}{K}m^4e^{-Km} - \frac{4m^3}{K^2}e^{-Km} + \int \frac{3 \cdot 4 \cdot m^2}{K^2}e^{-Km},$$

and so forth. Therefore

$$\int m^4\,dme^{-Km}$$

$$= C - \frac{1}{K}m^4e^{-Km} - \frac{4m^3}{K^2}e^{-Km} - \frac{3 \cdot 4}{K^3}m^2e^{-Km} - \frac{2 \cdot 3 \cdot 4}{K^4}me^{-Km} - \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5}{K^5}e^{-Km},$$

and because this integral vanishes when $m = 0$, we have

$$C = \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5}{K^5}.$$

Then, since the integral ends at $m = \infty$, we have in that case $m^4e^{-Km} = 0$, $m^3e^{-Km} = 0$, &c. Therefore

$$\int m^4\,dme^{-Km} = \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5}{K^5},$$

and in order to obtain $x$ we have the following equation,

$$(w)\quad \begin{aligned} &\frac{1}{(3p+2q-x)^5} - \frac{1}{3(4p+2q-x)^5} - \frac{1}{3(3p+3q-x)^5} \\ &= \frac{1}{(2p+2q)^5} - \frac{2}{3(2p+3q)^5} - \frac{1}{9(4p+2q)^5} + \frac{1}{9(2p+4q)^5}. \end{aligned}$$

This is a 15th degree equation and gives 15 values for $x$, but we should observe that in the case of the preceding problem, $x$ must be positive and less than $p$, which makes a great number of these values useless. If there were many that satisfy these two conditions, though, it would be impossible to determine which is preferable. Fortunately, this does not happen here, and we shall see that there is only one that satisfies them, which is esential for the usage of this method.

Suppose that one of the roots of $x$ is $p - f$, and let $K$ stand for the second term of the equation ($w$). We will have

$$\frac{1}{(2p+2q+f)^5} - \frac{1}{3(3p+2q+f)^5} - \frac{1}{3(2p+3q+f)^5} = K.$$

Suppose that $p - f - u$ is a second root of $x$, $f + u$

being positive and less than $p$. We will then have

$$\frac{1}{(2p+2q+f)^5\left(1+\dfrac{u}{2p+2q+f}\right)^5} - \frac{1}{3(3p+2q+f)^5\left(1+\dfrac{u}{3p+2q+f}\right)^5} - \frac{1}{3(2p+3q+f)^5\left(1+\dfrac{u}{2p+3q+f}\right)^5} = K.$$

Let

$$\frac{1}{(2p+2q+f)^5\left(1+\dfrac{u}{2p+2q+f}\right)^5} = \frac{1}{(2p+2q+f)^5}\left(1+\frac{1}{l}\right),$$

$$\frac{1}{(3p+2q+f)^5\left(1+\dfrac{u}{3p+2q+f}\right)^5} = \frac{1}{(3p+2q+f)^5}\left(1+\frac{1}{l'}\right),$$

$$\frac{1}{(2p+3q+f)^5\left(1+\dfrac{u}{2p+3q+f}\right)^5} = \frac{1}{(2p+3q+f)^5}\left(1+\frac{1}{l''}\right),$$

where $l$, $l'$, $l''$ may be positive or negative, according to whether $u$ is positive or negative. Further, we will have $l < l'$ and $l < l''$, and

$$\frac{1}{l(2p+2q+f)^5} - \frac{1}{3l'(3p+2q+f)^5} - \frac{1}{3l''(2p+3q+f)^5} = 0.$$

But we have

$$\frac{1}{l(2p+2q+f)^5} - \frac{1}{3l(3p+2q+f)^5} - \frac{1}{3l(2p+3q+f)^5} = \frac{K}{l}.$$

So

$$\frac{K}{l} + \frac{1}{3(3p+2q+f)^5}\left(\frac{1}{l}-\frac{1}{l'}\right) + \frac{1}{3(2p+3q+f)^5}\left(\frac{1}{l}-\frac{1}{l''}\right) = 0.$$

Now $K$ must necessarily be positive, so this equation is clearly impossible unless we suppose

$$\frac{1}{l}=0, \quad \frac{1}{l'}=0, \quad \text{and} \quad \frac{1}{l''}=0,$$

which would give $u = 0$. Thus there is only one root of $x$ which satisfies the conditions prescribed above.

The difficulty of solving the equation $(w)$ for $x$ makes the use of the preceding method extremely laborsome, but one can employ it in delicate circumstances, where the mean among several observations is required with precision. While in the preceding problem we have only considered three observations, it is clear that the solution is entirely similar for any number whatever.

To give an example of the preceding method and how to use it, suppose (Figure 1) that the observations $b$ and $c$ coincide, so $q = 0$, and setting $x = pz$, the equation $(w)$ gives

$$\frac{2}{3(3-z)^5} - \frac{1}{3(4-z)^5} = \frac{1.3229}{3 \cdot 2^5},$$

and letting $(3-z)/2 = \mu$, we will have

$$\mu = \sqrt[5]{\frac{2}{1.3229 + 1/(\tfrac{1}{2}+\mu)^5}}.$$

If in a first approximation we neglect the term $1/(\tfrac{1}{2}+\mu)^5$, we will have a first value for $\mu$ which, substituted in the equation, will give a closer second value of $\mu$, and so forth. In this manner I have found $\mu = 1.0697$, which gives $z = 0.860$. Therefore $x = p \cdot 0.860$. This is consequently the mean that one should take among three observations of which two coincide. For example, if the first gives the instant of the phenomenon at $m^h\ 30'\ 0''$ and the other two at $m^h\ 30'\ 10''$, we should suppose the true instant of the phenomenon is $m^h\ 30'\ 8''.6$. From the method used by astronomers we would find it to be $m^h\ 30'\ 6''\tfrac{2}{3}$. We thus see that the preceding method gives an instant closer to the two coincident observations, and that this conforms better with the probabilities. So we easily see that the mean that should be taken among three observations of which two coincide is not given by the method of arithmetic means.

Here now is a small table that I have constructed for the use of observers. As the value $q$ had been supposed in our calculations to be less than that of $p$, I have made it successively equal to $0 \cdot p$, $0.1 \cdot p$, $0.2 \cdot p$, $0.3 \cdot p$, &c., up to $p$. I have calculated the corresponding values of $x$. If the value of $q$ were to fall between two of these decimals, it would be easy to find $x$ by interpolation.

For the use of this table, we should observe that $x$ expresses the distance from the extreme observation that differs most from the intermediate one, to the mean that we should take among the three observations.

| | |
|---|---|
| $q=0p$ | $x=p\cdot 0.860$ |
| $q=0.1p$ | $x=p\cdot 0.894$ |
| $q=0.2p$ | $x=p\cdot 0.916$ |
| $q=0.3p$ | $x=p\cdot 0.932$ |
| $q=0.4p$ | $x=p\cdot 0.944$ |
| $q=0.5p$ | $x=p\cdot 0.955$ |
| $q=0.6p$ | $x=p\cdot 0.965$ |
| $q=0.7p$ | $x=p\cdot 0.975$ |
| $q=0.8p$ | $x=p\cdot 0.984$ |
| $q=0.9p$ | $x=p\cdot 0.992$ |
| $q=p$ | $x=p$ |

## VI

The preceding theory has led me to the following considerations which are perhaps not useless in the theory of chances, and with which I end this memoir.

I suppose that $A$ and $B$ play "heads or tails" under these conditions: if $A$ throws heads on the first toss, $B$ will give him two écus; he will give him four if $A$ doesn't throw heads until the second toss, eight if he doesn't throw it until the third, and so forth until $x$ tosses are complete. It is easy to determine $A$'s expectation, the sum that he should give to $B$ before beginning the game. For let $y_x$ be this sum, then if we suppose that the number of tosses instead of $x$ is increased by one, it is clear that the $A$'s expectation will be increased by $2^{x+1}$ écus times the probability $(1/2^{x+1})$ of obtaining this on the toss $x+1$. We will thus have $y_{x+1} - y_x = 1$, from which we find by integrating, $y_x = x + C$, $C$ being an arbitrary constant. Now for $x = 1$, $y_x = 1$, so $C = 0$. Thus $A$ should give $B$ $x$ écus.

In this solution, we assumed that the coin which was tossed in the air had no tendency to favor either heads or tails. Now, this supposition is only mathematically admissible because physically there must be an inequality. But as the two players $A$ and $B$ are ignorant of it at the beginning of the game, of which side has the greater tendency, we can believe that this uncertainty neither increases nor decreases the advantage. We shall see, however, that nothing is less founded than this supposition, that it follows that the science of chances must be used with care, and must be modified when we pass from the mathematical case to the physical.

We shall examine the consequences of supposing that the coin has a greater tendency to fall on one side or the other. Let $(1-\pi)/2$ be the probability that when the coin is tossed in the air, heads or tails (we do not know which) will occur. If $(1+\pi)/2$ is supposed the probability of heads, $A$'s expectation will be equal to

$$(1+\pi)[1+(1-\pi)+(1-\pi)^2+\cdots+(1-\pi)^{x-1}] = \frac{(1+\pi)[(1-\pi)^x-1]}{-\pi};$$

if we supposed the probability of heads is $(1-\pi)/2$, then $A$'s expectation will be equal to $((1-\pi)\cdot[(1+\pi)^x-1])/\pi$. Now since it is as natural to attribute the probability $(1+\pi)/2$ to heads as to tails, if we let $E$ be $A$'s expectation we will have

$$E = 1 + \frac{(1-\pi\pi)}{2}[(1+\pi)^{x-1}-(1-\pi)^{x-1}],$$

and if we regard $\pi$ as very small, we will have, if $x$ is not too large,

$$E = x + \left\{\frac{(x-1)(x-2)(x-3)}{1\cdot 2\cdot 3} - \frac{(x-1)}{1}\right\}\pi\pi.$$

So $A$'s expectation is less than $x$ if $x$ is between 1 and 5, and it is equal to $x$ if $x=5$. After a larger number of tosses, $A$'s expectation becomes greater than $x$, and taking $x$ infinite, it is infinitely larger.

Since the value of $\pi$ is unknown, it is hardly possible to evaluate $A$'s expectation for a number $n$ of tosses; however, if we are assured that $\pi$ cannot exceed a certain quantity, for example $1/q$, but that it is equally able to be any fraction between 0 and $1/q$, then we can calculate $A$'s expectation in the following manner.

If we conceive of the fraction $1/q$ as partitioned into an infinity of equal parts, represented by $d\pi$, it is clear that the element of $A$'s expectation will be equal to $Eqd\pi$, and that the total expectation will be

$$\begin{aligned}\int Eq\,d\pi &= \int q\,d\pi\left(1+\frac{(1-\pi\pi)}{2\pi}[(1+\pi)^{x-1}-(1-\pi)^{x-1}]\right)\\ &= n + \left\{\frac{(n-1)(n-2)(n-3)}{1\cdot 2\cdot 3} - \frac{(n-1)}{1}\right\}\frac{1}{3q^2}\\ &\quad + \left\{\frac{(n-1)\cdots(n-5)}{1\cdot 2\cdot 3\cdot 4\cdot 5} - \frac{(n-1)\cdots(n-3)}{1\cdot 2\cdot 3}\right\}\frac{1}{5q^4}\\ &\quad + \&c.\end{aligned}$$

(after integrating and adding the appropriate constant). If we suppose $q$ quite large, this quantity is reduced to its first two terms, as long as $n$ is sufficiently small, and $A$'s expectation will then be

$$n + \left\{\frac{(n-1)(n-2)(n-3)}{1\cdot 2\cdot 3} - \frac{(n-1)}{1}\right\}\frac{1}{3q^2}.$$

It is remarkable that this expectation will be less than $n$ when the number of tosses is between 1 and 5, that it will equal $n$ when $n = 5$, and finally that it will be larger when $n$ is larger than 5.

If we suppose $n = 2$ and $1/q, = 1/10$, $A$'s expectation will equal $2 - 1/300$ écus, from which we see that $A$ plays with a disadvantage if he only gives $B$ 2 écus, because he should only give him $2 - 1/300$ écus.

If we seek the probability of obtaining heads in two tosses, we find by this method that it is equal to $1/4 + 1/12qq$, consequently greater than $1/4$. It follows that we deceive ourselves in calculating these probabilities in the ordinary manner, that is, without paying attention to the inequalities that may be found between the two faces of the coin.

These considerations give rise to a new type of problem concerning chance, one quite useful in the application of the calculus of probabilities. We see that even if we are ignorant of which side of the coin has the larger probability, this uncertainty can make the lot of one of the players more advantageous than the other. It is thus of great interest to know in which different cases, which player has the greater advantage.

But it is principally in the application of the science of probabilities to the game of dice that this theory requires modification. Often a die which appears to be a perfect cube will exhibit a quite appreciable unequal tendency toward its different sides, such that in a large number of tosses one face will appear more frequently than another. This can be due to the heterogeneity of the material the die is made from, or to a lack of a perfectly cubical shape, as I have observed even with the most regular and the most homogeneous dice I could find, and more particularly with what are called *English dice*. We shall now examine the changes that these inequalities produce in the solutions of problems concerning the game of dice.

$A$ and $B$ play together with the condition that if $A$ throws a given face of a die in any one of $n$ tosses, $B$ will give him the sum $a$. We ask what sum $A$ must give to $B$.

By the theory of chances, we find that $A$'s expectation is $a - (5^n/6^n)a$, and that this is the sum that he must give to $B$. This solution supposes that all faces of the die are perfectly equal, which is only true mathematically speaking.

Let $(1 + \pi)/6$ be the probability that one of the faces of the die (we are ignorant of which) will be thrown at the first toss; let

$$\frac{1+\pi'}{6}, \frac{1+\pi''}{6}, \ldots, \frac{1+\pi^{V}}{6}$$

be the probabilities the other faces are thrown on the first toss. We will have

$$\frac{1+\pi}{6} + \frac{1+\pi'}{6} + \cdots + \frac{1+\pi^{V}}{6} = 1,$$

so $\pi + \pi' + \cdots + \pi^{V} = 0$. Now if we suppose that the given face has the probability $(1 + \pi)/6$ of being thrown on a single toss, the probability that it will not occur in $n$ tosses will be

$$\frac{(5 + \pi' + \pi'' + \cdots + \pi^{V})^n}{6^n} = \frac{(5-\pi)^n}{6^n};$$

$A$'s expectation is thus

$$a\left(1 - \frac{(5-\pi)^n}{6^n}\right).$$

Similarly, if the probability that the given face will be thrown on the first toss is $(1 + \pi')/6$, we will have $A$'s expectation

$$= a\left(1 - \frac{(5-\pi')^n}{6^n}\right),$$

and so forth. It follows that $A$'s true expectation is

$$a - a\frac{(5-\pi)^n}{6^{n+1}} - a\frac{(5-\pi')^n}{6^{n+1}} - \cdots - a\frac{(5-\pi^{V})^n}{6^{n+1}}.$$

If we suppose $\pi$, $\pi'$, $\pi''$, &c. are quite small and $n$ fairly large, we will have this expectation

$$= a - \frac{5^n}{6^n}a - \frac{n(n-1)}{1 \cdot 2} \cdot \frac{5^{n-2}}{6^{n+1}} \cdot a[\pi^2 + \pi'^2 + \pi''^2 + \cdots + \pi^{V^2}],$$

from which it follows that if $\pi$, $\pi'$, &c. are not zero, which would be physically impossible, $A$'s expectation is less than $a - (5^n/6^n)a$, unless $n = 1$. Thus, it follows that $A$, in giving to $B$ the sum $a - (5^n/6^n)a$, plays at a disadvantage.

If $n$ were a large number, we would find $A$'s expectation equal to

$$a - \frac{5^n}{6^n}a - \frac{n(n-1)}{1 \cdot 2} \cdot \frac{5^{n-2}}{6^{n+1}} a[\pi^2 + \pi'^2 + \pi''^2 \cdots + \pi^{V^2}] + \frac{n(n-1)(n-2)}{1 \cdot 2 \cdot 3} a \frac{5^{n-3}}{6^{n+1}} \cdot [\pi^3 + \pi'^3 + \cdots + \pi^{V^3}] + \&c.$$

Now since it is as natural to suppose $\pi$, $\pi'$, &c negative as positive, it is clear that we should discard terms where they are raised to an odd power, thus $A$'s expectation will be

$$a - \frac{5^n}{6^n}a - \frac{n(n-1)}{1 \cdot 2}\frac{5^{n-2}}{6^{n+1}}a[\pi^2 + \pi'^2 + \cdots + \pi^{V^2}]$$
$$- \frac{n(n-1)(n-2)(n-3)}{1 \cdot 2 \cdot 3 \cdot 4}\frac{5^{n-4}}{6^{n+1}} \cdot a[\pi^4 + \cdots + \pi^{V^4}] - \&c.$$

which is always less than $a - (5^n/6^n)a$ whatever $n$ is.

*If the quantities $\pi$, $\pi'$ $\pi''$, &c. are unknown, but we are assured that they can neither exceed $1/q$, nor be less than $-(1/q)$, we propose to find A's expectation.*

This problem presents many difficulties and requires particular consideration, because the quantitities $\pi$, $\pi'$, $\pi''$, &c. are mutually dependent, which renders the various values they may take more or less probable. To simplify calculation, instead of a die imagine a triangular prism which can only fall on the three rectangular faces. In this case, supposing $\pi$ quite small and $n$ fairly large, $A$'s expectation is

$$a - \frac{2^n}{3^n}a - \frac{n(n-1)}{1 \cdot 2}\frac{2^{n-2}}{3^{n+1}}a[\pi^2 + \pi'^2 + \pi''^2].$$

Here $\pi + \pi' + \pi'' = 0$, and we will have $\pi'' = -\pi - \pi'$; thus $A$'s expectation is

$$a - \frac{2^n}{3^n}a - \frac{n(n-1)}{1 \cdot 2}\frac{2^{n-1}}{3^{n+1}}a[\pi^2 + \pi'\pi + \pi'^2].$$

I now suppose $\pi'$ positive and constant, and I seek $A$'s expectation in this case. To this end, I multiply the preceding quantity by $d\pi$, which after integrating gives

$$a\pi - \frac{2^n}{3^n}a\pi - \frac{n(n-1)}{1 \cdot 2} \cdot \frac{2^{n-1}}{3^{n+1}}a\left[\frac{1}{3}\pi^3 + \frac{\pi'\pi^2}{2} + \pi'^2\pi\right] + C.$$

Now the largest positive value $\pi$ can have is $1/q - \pi'$, so, supposing the integral vanishes when $\pi = 0$, we will have $C = 0$ and the integral corresponding to positive $\pi$ is

$$\left(a - \frac{2^n}{3^n}a\right)\left(\frac{1}{q} - \pi'\right) - \frac{n(n-1)}{1 \cdot 2} \cdot \frac{2^{n-1}}{3^{n+1}}a\left\{\frac{1}{3}\left(\frac{1}{q} - \pi'\right)^3 + \frac{\pi'}{2}\left(\frac{1}{q} - \pi'\right)^2 + \pi'^2\left(\frac{1}{q} - \pi'\right)\right\}.$$

For the integral corresponding to negative $\pi$, I put negative $\pi$ in the above expression for $A$'s expectation, which then becomes

$$a - \frac{2^n}{3^n}a - \frac{n(n-1)}{1 \cdot 2} \cdot \frac{2^{n-1}}{3^{n+1}}a(\pi^2 - \pi'\pi + \pi'^2).$$

If we multiply this quantity by $d\pi$ and integrate, we will have

$$\left(a - \frac{2^n}{3^n}a\right)\pi - \frac{n(n-1)}{1 \cdot 2} \cdot \frac{2^{n-1}}{3^{n+1}}a\left(\frac{1}{3}\pi^3 - \frac{1}{2}\pi'\pi^2 + \pi'^2\pi\right).$$

Now the largest value $\pi$ can have in this case is $1/q$. We thus will have for the total integral corresponding to negative $\pi$,

$$\left(a - \frac{2^n}{3^n}a\right)\frac{1}{q} - \frac{n(n-1)}{1 \cdot 2} \cdot \frac{2^{n-1}}{3^{n+1}}a\left(\frac{1}{3q^3} - \frac{1}{2}\pi'\frac{1}{q^2} + \pi'^2\frac{1}{q}\right).$$

If we add this integral to the preceding, it is clear that their sum expresses the sum of all of $A$'s expectations which correspond to this value of $\pi'$, and consequently to all possible $\pi$ from $-(1/q)$ to $1/q - \pi'$. This sum will be

$$\left(a - \frac{2^n}{3^n}a\right)\left(\frac{2}{q} - \pi'\right) - \frac{n(n-1)}{1 \cdot 2} \cdot \frac{2^{n-1}}{3^{n+1}} \cdot a\left[\frac{1}{3}\left(\frac{1}{q} - \pi'\right)^3 + \frac{1}{3q^3} + \frac{\pi'^2}{2}\left(\frac{2}{q} - \pi'\right)\right].$$

If we multiply this quantity by $d\pi'$ and integrate it, we will have

$$\left(a - \frac{2^n}{3^n}a\right)\left(\frac{2}{q}\pi' - \frac{1}{2}\pi'^2\right) - \frac{n(n-1)}{1 \cdot 2} \cdot \frac{2^{n-1}}{3^{n+1}} \cdot a\left\{\frac{1}{12q^4} - \frac{1}{12}\left(\frac{1}{q} - \pi'\right)^4 + \frac{\pi'^3}{3q} - \frac{1}{8}\pi'^4 + \frac{\pi'}{3q'}\right\} + C.$$

Evaluating this from $\pi' = 0$ to $\pi' = 1/q$, we get

$$\left(a - \frac{2^n}{3^n}a\right)\frac{3}{2qq} - \frac{n(n-1)}{1 \cdot 2}a\frac{5 \cdot 2^{n-4}}{3^{n+1}q^4}.$$

This expresses the sum total of $A$'s expectation corresponding to all possible positive variations of $\pi'$; to have the resulting expectation for $A$, it is clearly necessary to divide this sum by the total number of possible positive variations of $\pi'$. Now, the number of corresponding variations of $\pi'$, is, by the preceding, $2/q - \pi'$; multiplying by $d\pi'$ and integrating, we find $3/2qq$ as the divisor for the preceding quantity. Thus $A$'s expectation, corresponding to positive $\pi'$, is

$$a - \frac{2^n}{3^n}a - \frac{n(n-1)}{1 \cdot 2} \cdot \frac{2^{n-3}}{3^{n+2}} \cdot \frac{5a}{q^2}.$$

Now, the expectation which corresponds to negative $\pi'$ is clearly the same, as we would as soon wager that $\pi'$ is negative as positive: the total of $A$'s expectation is thus

$$a - \frac{2^n}{3^n}\,a - \frac{n(n-1)}{1\cdot 2}\cdot\frac{2^{n-3}}{2^{n+2}}\cdot\frac{5a}{q^2}.$$

Following the same process, we could obtain the solution to the preceding problem in the cases where the solid would have 4, 5, 6, &c. faces. The only additional difficulty is the length of the calculation.

These examples are sufficient to show the precaution needed in applying the mathematical considerations of the calculus of probabilities to physical objects. We suppose in the theory that the different ways in which an event can occur are equally probable, or where they are not, that their probabilities are in a given ratio. When we wish then to make use of this theory, we regard two events as equally probable when we see no reason that makes one more probable than the other, because if they were unequally possible, since we are ignorant of which side is the greater, this uncertainty makes us regard them as equally probable.

When it is only a question of simple probabilities, it would appear that this inequality of probabilities would not diminish the correctness of applying this calculus to physical objects. If $B$, for example, agrees to give two écus to $A$ if a head occurs on the first toss, then by the theory, that is, supposing head and tail equally possible, $A$ should give $B$ one écu before beginning the game. It is the same, as we can easily assure ourselves, if we were to suppose an unequal probability for heads and for tails, where we were ignorant of which side is the greater. But when composite probabilities are in question, it appears to me that our application of this theory to physical events requires modification. For example, if at the game of heads and tails $B$ wagers with $A$ that a head will not occur on either of two tosses, the probability that $B$ will win is clearly composite, since it is the probability that a head will not occur on the first toss, and that it will not occur on the second toss, multiplied together. Now, in this case the probability for $B$ from the ordinary theory is $1/4$, while if instead we suppose heads and tails unequally possible, this probability is larger than $1/4$.

This aberration in the ordinary theory, which has not to my knowledge been previously noted by anyone, appears to me to merit the attention of geometers. It seems to me that it is essential to consider it whenever we apply the calculus of probabilities to the different objects of civil life.[1]

[1] Section VII of the *Memoir*, which announced four theorems concerning differential equations which were totally unrelated to Sections I–VI, is omitted.