History of Probability (Part 5) – Laplace (1749-1827)

Pierre-Simon Laplace was born in 1749, when the phrase "theory of probability" was gradually replacing "doctrine of chances." Assigning a number to the *probability of an event* was seen as more convenient mathematically than calculating in terms of odds. At the same time there was a growing sense that the mathematics of probability was useful and important beyond analysis of games of chance. By the time Laplace was in his twenties he was a major contributor to this shift. He remained a "giant" of theoretical probability for 50 years. His work, which we might call the **foundations of classical probability theory**, established the way we still teach probability today.

Laplace grew up during a period of tremendous discovery and invention in math and science, perhaps the greatest in modern history, often called the age of enlightenment. In France, as the needs of the country were changing, the role of the major religious groups in higher education was diminishing. Their primary charge remained the education of those who would become the leading religious leaders, but other secular institutions of learning were established, where the emphasis was on math and science rather than theology. Laplace's father wanted him to be a cleric and sent him to the University of Caen, a religiously oriented college. Fortunately, Caen had excellent math professors who taught the still fairly new calculus. When he graduated, Laplace, against his father's wishes, decided to go to Paris to become a fulltime mathematician rather than a cleric who could do math on the side.

The sense in the air in France (and in England and the United States) was that reason and science would benefit society, and that opportunity should be based on individual talent. In France the Royal Academy of Science had been established. It paid smart people to do fulltime research. That's the position Laplace wanted. He brought a letter of introduction to the eminent mathematician, d'Alembert, who was impressed and got him a job in a new Royal military college (where young Napoleon was a student). Starting then, Laplace submitted paper after paper to the Royal Academy, and finally, after about five years, he was accepted as a member and became a fulltime professional researcher, submitting papers for many years on a great variety of topics, but especially on celestial mechanics.

The big problems of the day were about the solar system. Newton's law of gravity was being used in all kinds of analyses about objects in the sky. It had been this kind of work that led Newton's friend, Edmund Halley, in 1704 to predict the arrival of a comet at the end of 1758. Its arrival had tremendously impressed Laplace who was 10 years old then. Among the important problems that Laplace worked on was what was keeping the planets in their orbits. Why didn't they crash into each other or the sun? The calculus involved was very complicated and was one of the reasons that many European mathematicians chose to use Leibniz's algebraic notation (the style we use now) rather than Newton's more geometric system. Laplace's 5-volume work ***Mécanique Céleste*** **(*Celestial Mechanics*)** summarized all that was known about the mathematics used to describe the solar system. One of his first contributions was to "prove" that the solar system was stable, that it would not fly apart. Laplace is considered one of the greatest mathematicians of all time, sometimes called "The Newton of France."

During the five years that Laplace was hoping to be admitted to the Royal Academy, he came across Abraham De Moivre's book, *Doctrine of Chances.* It struck him that he could use ideas in it to attack the problem of describing the probability of causes of events. This is the same problem that Thomas Bayes had set himself. They were both thinking about causes in a more general way, not just in the context of gambling. Essentially, then, Laplace rediscovered Bayes' Theorem, and it was Laplace's publications that became well known and that influenced its use in probability and statistics.

In 1774, when he was 25 years old, his short paper, ***Mémoire sur la probabilité des causes par les événemens*** (***Memoir on the Probability of the Causes of Events***), was published by the Royal Academy of Sciences. It was influential in getting him appointed to the Academy. You can see in the excerpt below that the level of mathematical sophistication has jumped to calculus. Critically, the rules of combinations and permutations become unmanageable with large numbers of trials (or large numbers of data points), and so smooth models such as the normal curve come into play. But this means you need to be able to calculate integrals to measure areas, and you need tools to find parameters such as means and variance using calculus.

You can see the appearance of calculus notation in the first problem of the memoir. Note that the urn contains an infinity of tickets. This assumption leads to a calculus-based approach for summing "all possible" values of a probability from 0 to 1.

> **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 $x = 1$.

Laplace continued to write about probability his whole career, with many of his contributions presented in his influential text, ***Théorie Analytique des Probabilités*** (***Analytical Theory of Probability***)**,** first published in 1812 and then followed by other editions in 1814, 1820, and 1825. A singular contribution was his ***generalization of the Central Limit Theore****m*. De Moivre had shown that the normal curve provided a good approximation to the binomial distribution of the number of successes in repeated trials. Laplace showed that the sum of a set of repeated observations of random values from almost *any* distribution can be approximated by a normal curve. This version of the Central Limit Theorem is key to explaining why so many distributions used in statistics approach the normal distribution as the sample size increases. It is absolutely fundamental to inferential statistics. In this same book he also proved that for large numbers of observations the best way to fit observational data to models is through ***the method of least squares,*** another crucial result.

Example of using calculus in probability analysis. A probability generating function.
Consider $n = 3$ trials of a random process. Each trial will result in either success or failure. Assume that for each trial the probability of success is p=¼, so the probability for failure is q = 1-p = ¾.

The **binomial probability distribution** is a function that tells you the probability associated with the total number of successes, S, you will get in the three trials.

Here are all the possible values of S and their probabilities:

Possible value of S for n = 3	Probability, $p(s)$
s=0	$p(0) = 1\left(\frac{1}{4}\right)^0\left(\frac{3}{4}\right)^3 = \frac{27}{64}$
s=1	$p(1) = 3\left(\frac{1}{4}\right)^1\left(\frac{3}{4}\right)^2 = \frac{27}{64}$
s=2	$p(2) = 3\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^1 = \frac{9}{64}$
s=3	$p(3) = 1\left(\frac{1}{4}\right)^3\left(\frac{3}{4}\right)^0 = \frac{1}{64}$

Note that the coefficients 1, 3, 3, 1 are the entries of the third row of Pascal's triangle which are also known as binomial coefficients because they are the coefficients in the expansion of the binomial expression $(a+b)^3$.

$$(a+b)^3 = 1a^3 + 3a^2b^1 + 3a^1b^2 + 1b^3$$

The coefficients tell you the number of different ways you can achieve the value of *S*. For example, there are 3 ways to get 1 success in 3 trials.

We can find the mean, or expected value, of *S*.

$$E[S] = \sum sp(s) = 0p(0) + 1p(1) + 2p(2) + 3p(3) = 0\frac{27}{64} + 1\frac{27}{64} + 2\frac{9}{64} + 3\frac{1}{64} = \frac{48}{64} = \frac{3}{4}$$

You can see that if n is a big number, the calculation of the mean would be very tedious.

De Moivre and Laplace used the concept of a **generating function** to attack this problem. A generating function, *G(x),* allows us to replace an expression involving many discrete terms by a continuous function that has nice calculus properties. One of the nice properties is that $\boldsymbol{G'(1)}$ **gives the expected value.**

A probability generating function is a polynomial of degree *n* where the coefficients are the possible probabilities.

For our problem, $G(x) = p(0)x^0 + p(1)x^1 + p(2)x^2 + p(3)x^3$

$$= 1\left(\frac{1}{4}\right)^0\left(\frac{3}{4}\right)^3x^0 + 3\left(\frac{1}{4}\right)^1\left(\frac{3}{4}\right)^2x^1 + 3\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^1x^2 + 1\left(\frac{1}{4}\right)^3\left(\frac{3}{4}\right)^0x^3$$

By putting the *x* factors together with the ¼ factors we get:

$$G(x) = 1\left(\frac{1}{4}x\right)^0\left(\frac{3}{4}\right)^3 + 3\left(\frac{1}{4}x\right)^1\left(\frac{3}{4}\right)^2 + 3\left(\frac{1}{4}x\right)^2\left(\frac{3}{4}\right)^1 + 1\left(\frac{1}{4}x\right)^3\left(\frac{3}{4}\right)^0 = \left(\frac{1}{4}x + \frac{3}{4}\right)^3$$

$G(x) = \left(\frac{1}{4}x + \frac{3}{4}\right)^3$ is a "nice" continuous function that incorporates the information about *S*.

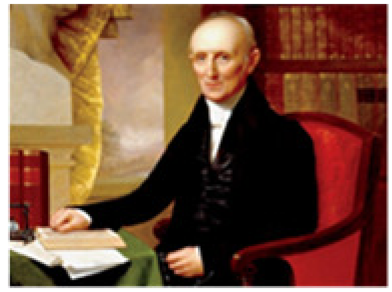
Because $G'(x) = 3\left(\frac{1}{4}x + \frac{3}{4}\right)^2\frac{1}{4}$, we immediately see that $E[S] = G'(1) = 3(1)\left(\frac{1}{4}\right) = \frac{3}{4}$, as we got earlier.

More importantly, you can also see that, in general, the generating function for any binomial distribution is $G(x) = (px+q)^n$ and that $G'(x) = n(px+q)^{n-1}p$ from which we get $G'(1) = np$. In short, we have proved that the mean of a binomial distribution is given by *np*.

Exercise

1. The variance of a random variable is given by: $G''(1) + G'(1) - G'(1)^2$. Show that the variance of a binomial random variable is *npq*.

Nathaniel Bowditch (1773-1838). American translator of Laplace



A superb annotated English translation of Laplace's monumental *Mécanique Céleste*, was published in America in 1829. The translator was Nathaniel Bowditch "… the country's most accomplished mathematician, the man Thomas Jefferson called "a meteor of the hemisphere." [Thornton,]

The translation was greeted with great acclaim. It included some corrections of errors as well as annotation to make the proofs simpler to follow. It is possible that Bowditch was the only person to read Laplace so thoroughly. Many scholars in England, and in Europe, were surprised that this level of mathematics was produced in America. A writer in the London Quarterly Review expressed his admiration in these words: "The idea of undertaking a translation of the whole *Mécanique Céleste,* accompanied throughout with a copious running commentary, is one… we should never have expected to have found originated, [on the opposite shores of the Atlantic]. The translation was hugely important for the development of astronomy in the United States.

Here is the opening of Bowditch's translation of Laplace's Volume 1.

> The object of the author, in composing this work, as stated by him in his preface, was to reduce all the known phenomena of the system of the world to the law of gravity, by strict mathematical principles; and to complete the investigations of the motions of the planets, satellites, and comets, begun by Newton in his Principia. This he has accomplished, in a manner deserving the highest praise, for its symmetry and completeness; but from the abridged manner, in which the analytical calculations have been made, it has been found difficult to be understood by many persons, who have a strong and decided taste for mathematical studies, on account of the time and labour required, to insert the intermediate steps of the demonstrations, necessary to enable them easily to follow the author in his reasoning. To remedy, in some measure, this defect, has been the chief object of the translator in the notes.

In the fourth volume of the last edition of Laplace's work, fifty-two corrections are noted as having been made by Bowditch. Laplace's widow acknowledged Bowditch's contribution by presenting him a marble bust of Laplace.

Nathaniel Bowditch is still known today for his contribution to commercial maritime navigation. His *New American Practical Navigator,* first published in 1802, has been continually updated and is now available online. To this day mariners refer to it simply as "Bowditch." In 1806 he was offered the chair of mathematics and physics at Harvard in 1806, but turned it down to maintain his position in commercial maritime insurance.

Sources

Weisberg, H. I. (2014). *Willful Ignorance: The Mismeasure of Uncertainty*: Wiley Publishing.
McGrayne, S. B. (2011). *The theory that would not die*. New Haven Conn.: Yale University Press.
Katz, V. J. (1998). *A history of mathematics* (2nd ed.). Reading, Mass.: Addison-Wesley.
Wikipedia article on Laplace
Portrait of Laplace: http://patrick.maher1.net/318/lectures/laplace.pdf
Portrait of Bowditch: *Harvard Magazine*, July 2016
Laplace 1774 memoir: http://www.york.ac.uk/depts/maths/histstat/memoir1774.pdf
Thornton, T. P. Nathaniel Bowditch, *Harvard Magazine*, July 2016
For more on generating functions: http://www.cl.cam.ac.uk/teaching/0708/Probabilty/prob06.pdf
Quotation from Quarterly Review; vol. xlvii., 1832, pp. 558, 559.