The first statement of the formula for the Normal Curve

Cantillo, Andres

University of Missouri Kansas City

18 December 2011

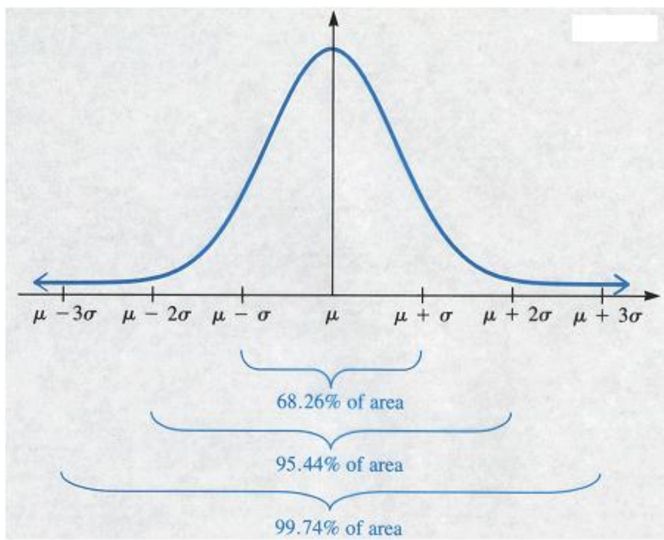
“The first statement of the formula for the Normal Curve”

Andres F. Cantillo

University of Missouri Kansas City

According to Smith (1) De Moivre's paper *"Approximatio ad Summam Terminorium Binomii"* written in 1733 and reproduced in (2) is

> "the first statement of the formula for the 'normal curve', the first method of finding the probability of the occurrence of an error of a given size when that error is expressed in terms of the variability of the distribution as a unit, and the first recognition of that value later termed the *probable error."* (1 p. 566).



De Moivre's book "The Doctrine of Chances" (2) is thorough account of what was known about probability and annuities. The proof that is the object of this paper is included in the very last pages of the book (pages 235-243). The aim of the present paper is to explicate De Moivre's first part of the proof in such a way that we can trace back the reasoning behind this creation has shaped the modern way of doing science.

Preliminaries:

Following (3) the probability of a random number of odds X in the so called Bernoulli experiment (an experiment with only two possible outcomes) follows the probability distribution function:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k},$$

$$k = 0,1, \dots, n,$$

where p is the true proportion of odds, k is the hypothetic number of odds and n is the number of experiments or trials. This distribution is called *binomial probability distribution*.

The binomial coefficient $\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{(n-k)(n-k-1)\cdots 2\cdot 1}$ counts the number of possibilities to get k odds in n drawings.

According to (3) Jacob Bernoulli tried to figure out how to obtain knowledge of the true proportions in the Bernoulli experiment by means of repeated experiments. The intuition was that by increasing the number of experiments, the proportions obtained (a-posteriori) would get close to the proportions a-priori as the size of the sample increases. According to (3), the aim of these approximations was to understand phenomena whose proportions were not clearly defined

> "But, Bernoulli asked, what about problems such as those involving disease, weather, or games of skill, where the causes are hidden and the enumeration of equally likely cases impossible? In such a situations, Bernoulli wrote, "It would be a sign of insanity to attempt to learn anything in this manner." (3 p. 65).

The key assumption was that those proportions existed but were not known a-priori; nevertheless they could be known a-posteriori.

Thus James Bernoulli's attempt (See (4)) was aimed at the estimation of the odds in the Bernoulli experiment. How large does the number of experiments need to be so that the odds of a Bernoulli experiment get close enough (Achieve moral certainty) (4) to the actual number of odds. The intuition was that there was an "increase in accuracy by increase of trials…" (4 p. 207). According to (4), although James Bernoulli stated the problem, its actual solution belongs to De Moivre and is given by his approximation.

The problem is that there was no mathematical proof about the number of trials n required in order to have a probability close enough to certainty that the proportion obtained after running the experiments was close enough to the true proportion p with certain limits (4). The binomial distribution was used in order to find the minimum number of trials necessary to achieve a probability P close to one. The problem was tackled by James Bernoulli, De Moivre and others first in the case where p (the true proportion) was known De Moivre (5) explains the problem one page before beginning his approximation with the following words:

> But suppose it should be said, that notwithstanding the reasonableness of building Conjectures upon Observations, still considering the great Power of Chance, Events might at long run fall out in a different proportion from the real Bent which they have to happen one way or the other; and that supposing for Instance that an Event might as easily happen as not happen[so we expect an equal proportion of happening and not happening], whether after three thousand Experiments it may not be possible it should have happened two

> thousand times and failed thousand; and that therefore the Odds against so great variation from Equality should be assigned, whereby the Mind would be the better disposed in the Conclusions derived from Experiments. (5 p. 242)

The calculation of the binomial distribution requires the estimation of the binomial coefficient. This procedure becomes extremely difficult when the number of experiments n is too high. For this reason De Moivre needed to find an accurate approximation of the sum terms of the binomial $(a+b)^n$. In the present paper I aim to explicate the first part of De Moivre's approximation stated in (2) and proved in (6)

Explication of De Moivre's Proof:

"A Method of approximating the Sum of the Terms of the Binomial $(\boldsymbol{a}+\boldsymbol{b})^{\boldsymbol{n}}$ expanded into a Series, from whence are deducted some practical Rules to estimate the Degree of Assent which is to be given to Experiments." (2 p. 243)

> Although the Solution of the Problems of Chance [specifically the ones that can be characterized by the binomial distribution (see Preliminaries)] often requires that several Terms of the Binomial $(a+b)^n$ be added together, nevertheless in very high Powers the thing appears so laborious, and of so great difficulty, that few people have undertaken that task; for besides *James* and *Nicolas Bernoulli,* two great Mathematicians, I know of no body that has attempted it [Se for instance James's explanation in (4)]; in which, tho' they have shewn very great skill, and have the praise which is due to their Industry, yet some things were farther required; for what

they have done is not so much an Approximation as the determining very wide limits, within which they demonstrated the Sum of the Terms was contained. Now the Method which they have followed has been briefly described in my *Miscellanea Analytica,* which the Reader may consult if he pleases, unless they rather chuse, which perhaps would be best, to consult what they themselves have writ upon that subject: for my part, what made me apply myself to that Inquiry was not out of opinion that I should excel others, in which however I might have been forgiven; but what I did was in compliance to the desire of a very worthy Gentleman [James Sterling], and good Mathematician, who encouraged me to it: I now add some new thoughts to the former; but in order to make their connexion the clearer, it is necessary for me to resume some few things that have been delivered by me a pretty while ago.

1. It is now a dozen years or more since I had found what follows [(6)]; If the Binomial 1+1 be raised to a very high Power denoted by n [$(1+1)^n$], the ratio which the middle Term [of the expansion of the binomial $(1+1)^n$] has to the Sum of all the Terms, that is, to 2^n, may be expressed by the Fraction $\frac{2A\times(n-1)^n}{n^n\times\sqrt{n-1}}$ [for n large], wherein A represents the number of which the Hyperbolic [natural, base e] Logarithm is $\frac{1}{12}-\frac{1}{360}+\frac{1}{1260}-\frac{1}{1680}$, &c.

According to the properties of the Pascal Triangle, the middle term M of the binomial $(1+1)^n$ is equal to the $\binom{n}{n/2}$ entry where n stands for the row and $\frac{n}{2}$ for the column. According to (6) De Moivre assumes n even which implies that the middle as stated exists in the Triangle. Following Pascal's triangle properties, $\binom{n}{n/2}$ is also equal to $\;_nC_{n/2}$. Thus $M=\binom{n}{n/2}=\frac{n!}{\left(\frac{n}{2}\right)!\cdot\left(n-\frac{n}{2}\right)}=\frac{n!}{\left(\frac{n}{2}\right)!^2}$.

It is also worth noting that De Moivre is assuming that the experiment's true proportions of the two possible outcomes are both $\frac{1}{2}$ which means that those outcomes are assumed equally likely. Hence $p = \frac{1}{2}$. In addition, k is assumed to be equal to $\frac{n}{2}$ so the question is how large has n to be in order to obtain $\frac{n}{2}$ close to $\frac{N}{2}$ where N stands for the number of total possible outcomes of the experiment and n is the number of successes (or failures) obtained from the experiments. By replacing this information in the binomial distribution (see Preliminaries) we obtain the following:

$P\left(X = \frac{n}{2}\right) = \binom{n}{n/2} \cdot (\frac{1}{2})^{\frac{n}{2}} \cdot (1 - \frac{1}{2})^{n-\frac{n}{2}} = \frac{\binom{n}{n/2}}{2^n}$ *" … the ratio which the middle Term has to the Sum of all the Terms, that is, to* 2^n*…"* [See De Moivre's quote above].

De Moivre's assertion means that:

$\frac{\binom{n}{n/2}}{2^n} \approx \frac{2A(n-1)^n}{n^n\sqrt{n-1}}$ for n large.

Since De Moivre does not direct the reader to a specific quotation in order to prove this statement, by following (6) in regard to De Moivre's proof we obtain:

Let $M = \binom{n}{n/2} = \frac{n!}{\left(\frac{n}{2}\right)!^2}$

Since n is even, for convenience by replacing $n = 2m$, meaning $m = \frac{n}{2}$ we obtain:

$$\binom{n}{n/2} = \frac{(2m)!}{m!^2} = \frac{(2m)(2m-1)\cdots(2m-m)(2m-m-1)\cdots(2m-m-m+1)}{m!^2}$$

$$= \frac{(m+m)(m+m-1)\cdots(m+2)(m+1)(m)(m-1)(m-2)\cdots(m-m+1)}{m!^2}$$

$$= \frac{(m+m)(m+m-1)\cdots(m+2)(m+1)m!}{m!^2} = \frac{(m+m)(m+m-1)\cdots(m+2)(m+1)}{m!} = \frac{(m+1)(m+2)\cdots\big(m+(m-1)\big)(m+m)}{(m-1)(m-2)\cdots\big(m-(m-1)\big)m}$$

Thus,

$$\ln M = \ln\frac{m+1}{m-1} + \ln\frac{m+2}{m-2} + \cdots + \ln\frac{m+(m-1)}{m-(m-1)} + \ln 2 \qquad [1]$$

According to the Taylor series expansion we have in general:

$$\ln\frac{(1+x)}{(1-x)} = 2\left(x + \frac{x^3}{3} + \frac{x^5}{5} + \frac{x^7}{7} + \cdots\right) \; for \; |x| < 1$$

In the case of the first term, $x = \frac{1}{m}$. Since $m = \frac{n}{2}$, This means $x = \frac{2}{n}$, and De Moivre assumes n large so $|x| < 1$ is achieved.

Substituting into the first term:

$$\ln\frac{m+1}{m-1} = \ln\frac{1+\frac{1}{m}}{1-\frac{1}{m}} = 2\left(\frac{1}{m} + \frac{1}{3m^3} + \frac{1}{5m^5} + \cdots\right)$$

Doing the same procedure to all the terms in $\ln M$ we obtain:

$$\ln\frac{m+2}{m-2} = \ln\frac{1+\frac{2}{m}}{1-\frac{2}{m}} = 2\left(\frac{2}{m} + \frac{2^3}{3m^3} + \frac{2^5}{5m^5} + \cdots\right)$$

$$\ln\frac{m+3}{m-3} = \ln\frac{1+\frac{3}{m}}{1-\frac{3}{m}} = 2\left(\frac{3}{m} + \frac{3^3}{3m^3} + \frac{3^5}{5m^5} + \cdots\right)$$

$$\vdots$$

Then we can find $\ln M$ by adding those term as equation [1] indicates. According to (6) De Moivre added vertically those terms in the following way where $s = m - 1$:

col. 1$= \frac{2}{m}(1 + 2 + \cdots + s)$

col. 2$= \frac{2}{3m^3}(1^3 + 2^3 + \cdots + s^3)$

col. 3$= \frac{2}{5m^5}(1^5 + 2^5 + \cdots + s^5)$

$\vdots$

The following steps were suggested by Professor Richard Delaware:

Each sum of integral powers can be calculated in closed form as polynomial in $m - 1$ using Bernoulli's Formulas. For instance, recall that $s = m - 1$:

col. 1$= \frac{2}{m}(1 + 2 + \cdots + (m-1)) = \frac{2}{m} \cdot \frac{(m-1)m}{2} = (m-1)\left[\frac{m-1}{m} + \frac{1}{m}\right] = (m-1)[y + \frac{1}{m}]$

where $y = \frac{m-1}{m}$

col. 2$= \frac{2}{m}(1^3 + 23 + \cdots + (m-1)^3) = \frac{2}{m}(\frac{(m-1)m}{2})^2 = (m-1)(\frac{y^3}{2\cdot 3} + \cdots + \square)$

col. 3$= \frac{2}{m}(1^5 + 2^5 + \cdots + (m-1)^5) = (m-1)[\frac{y^5}{3\cdot 5} + \cdots + \square)$

$\vdots$

The symbol ⬚ means that the sum of the respective column is finite. We are not interested in the exact form.

He adds the highest powers of y in these columns to get

$(m-1)[y+\frac{y^3}{2\cdot 3}+\frac{y^5}{3\cdot 5}+\cdots]$

After many simplifications this expression becomes:

$(2m-1)\ln(2m-1)-2m\ln m$

Likewise the second highest powers of y add to

$\frac{1}{2}\ln(2m-1)$

He then noticed that

$\lim_{m\to\infty}[Sum\ of\ the\ third-highest\ powers\ of\ y]=\frac{1}{12}$

$\lim_{m\to\infty}[Sum\ of\ the\ fourth\ highest\ powers\ of\ y]=-\frac{1}{360}$

$\vdots$

So he concluded that

$$\ln M=(2m-1)\ln(2m-1)-2m\ln m+\frac{1}{2}\ln(2m-1)+\frac{1}{12}-\frac{1}{360}+\cdots+\ln 2$$

Finally (6) arrives to the following expression for $\ln M$ where the final numeral series was obtained by taking a limit to infinity:

$$\ln M \approx (2m - \frac{1}{2}) \ln(2m-1) - 2m \ln m + \ln 2 + \frac{1}{12} - \frac{1}{360} + \cdots$$

Where $\frac{1}{12} - \frac{1}{360} + \cdots = \ln A$

Subtracting off $\ln 2^n = \ln 2^{2m} = 2m \ln 2$ then

$$\frac{M}{2^n} \approx \frac{2A \cdot (n-1)^n}{n^n \cdot \sqrt{n-1}}$$

for $m = \frac{n}{2}$ large, hence for n large.

QED.

Conclusions:

The question of the possibility of acquiring knowledge about the true probabilities of an experiment by means of repeated trials or observations begun with the analysis of Bernoulli's experiments. In addition, it was assumed that the true proportions of the experiment were known. The aim was to find the number of trials necessary to achieve a reasonable sense of certainty. In order to succeed in this task De Moivre had to develop mathematically the expansion of the binomial when the number of trials tends to infinity. In order to accomplish that task he had to analyze the relationship between the binomial term and the total sum of terms expressed as the approximation of the sum of infinite terms when n is assumed large. This is the core of De Moivre's proof of the approximation of the binomial to the normal distribution shown in (2).

Bibliography

1. **Smith, David.** *A source book in Mathematics.* New York : McGraw-Hill, 1959. 0-486-64690-4.

2. **DeMoivre, Abraham.** *The Doctrone of Chances.* London : Woodfall, 1738.

3. **Stigler, Stephen.** *The History of Statistics.* Cambridge, Massachusetts : Harvard University Press, 1986. 0-674-40340-1.

4. *James Bernoulli's Theorem.* **Pearson, Karl.** 3/4 (Dec), pp.201-210, s.l. : Biometrika, 1925, Vol. 17.

5. **DeMoivre, Abraham.** *The Doctrine of Chances: or, A Method of Calculating the Probabilities of Events in Play.* New York : Chelsea, 1967.

6. **Katz, Victor.** *A History of Mathematics.* Boston : Addison-Wesley, 2009.

7. **Rudas, Tomas.** *Handbook of Probability.* Los Angeles : Sage Publications, 2008. 978-1-4129-2714-7.