Appendix and historical notes on normal distributions

Appendix. Lindeberg's Theorem. Suppose for $i = 1, 2, ...,$ we have, for each i, independent random variables X_{ij} for $j = 1, ..., n_i$, with $EX_{ij} = 0$ for all i and j, and $\sum_{j=1}^{n_i} \mathrm{Var}(X_{ij} = 1$ for each i. For each $\varepsilon > 0$, let $A_{ij\varepsilon}$ be the event that $|X_{ij}| > \varepsilon$ and $I_{ij\varepsilon}$ its indicator function. Assume that for each $\varepsilon > 0$,

$$\lim_{i\to\infty} \sum_{j=1}^{n_i} E(X_{ij}^2 I_{ij\varepsilon}) = 0.$$

Then as $i \to \infty$, $S_i := \sum_{j=1}^{n_i} X_{ij}$ converges in distribution to $N(0, 1)$.

The theorem is proved in graduate probability textbooks (as in 18.175), for example, R. Dudley, *Real Analysis and Probability* (2d ed., Cambridge University Press, 2002), Theorem 9.6.1 pp. 316–318.

Historical Notes. Gauss, in 1809, published a work on celestial mechanics, in which he modeled measurement errors as being what would now be called normally distributed. He was given credit for discovering the normal distribution because his work appeared before publication in 1812 of an important book on probability by Laplace. But Laplace had published a paper defining normal densities in 1774. The first discovery was actually by Abraham de Moivre (1667–1754) in 1733, who gave the first central limit theorem, for binomial distributions. A history of probability by I. Todhunter erroneously said De Moivre treated only the special case where the binomial probability $p = 1/2$. Although De Moivre emphasized this case as an example, in fact he treated all p, $0 < p < 1$. He incorporated the theorem into the second edition, in 1738, of his book *Doctrine of Chances.* (A third edition was published posthumously in 1758.) His description of a normal density can be hard to decipher for modern readers because the notations for the number e and the function e^x were invented by Euler in 1727 and de Moivre did not use them. Rather he writes "the number which answers to the hyperbolic logarithm" x. In the 17th century and by some in the early 18th, "natural logarithm" and "hyperbolic logarithm" were both used, synonymously. The statistician Karl Pearson in 1924 noticed de Moivre's work and proposed the term "normal" distribution, now generally adopted by statisticians. By that time, the central limit theorem had been proved much more generally and the convenience of using normally distributed variables had been noticed. Further details of the history are in R. Dudley, op. cit., pp. 330–331.