CHAPTER XIII

A Historical Note on Brownian Motion

Historically, the mathematical roots of Brownian motion lie in the central limit theorem (CLT). The first CLT seems to have been obtained in 1733 by DeMoivre[1] for the normal approximation to the binomial distribution (i.e., sum of i.i.d. Bernoulli 0 or 1-valued random variables). In his 1812 treatise Laplace[2] obtained the far reaching generalization to sums of arbitrary independent and identically distributed random variables having finite moments of all orders. Although by the standards of rigor of present day mathematics Laplace's derivation would not be considered complete, the essential ideas behind this remarkable result may be found in his work. The first rigorous proof[3] of the CLT was given by Lyapounov almost 100 years later using characteristic functions under the Lyapounov condition for sums of independent, but not necessarily identically distributed, random variables having finite $(2+\delta)$th moments for some $\delta > 0$. This moment condition was relaxed in 1922 by Lindeberg[4] to prove the more general CLT, and in 1935, Feller[5] showed that the conditions are necessary (as well as sufficient), under uniform asymptotic negligibility of summands. The most

---

[1]DeMoivre's normal approximation to the binomial first appeared in a pamphlet "Approximatio ad summam terminorum binomii" in 1733. It appeared in book form in the 1756 edition of the Doctrine of Chance, London.

[2]Laplace, P.-S. (1812), "Théorie Analytique des Probabilités", Paris.

[3]Lyapunov, A.M. (1901). Nouvelle forme du théorème sur la limite de probabilités. *Mem. Acad. Imp. Sci. St.-Petersberg* **12** (5), 1–24.

[4]Lindeberg, J.W. (1922). Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung. *Math. Zeitschr.* **15**, 211–225.

[5]Feller, W. (1935). Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung. *Math. Zeitschr.* **40**, 521–559. Also, ibid (1937), **42**, 301–312.

popular form of the CLT is that for i.i.d. summands with finite second moments due to Paul Lévy.[6]

There are not many results in mathematics that have had such a profound impact as the CLT, not only on probability and statistics but also on many other branches of mathematics, as well as the natural and physical sciences and engineering as a whole. The idea of a stochastic process $\{B_t : t \geq 0\}$ that has independent Gaussian increments also derives from it. One may consider an infinite i.i.d. sequence $\{X_m : m \geq 1\}$ with finite second moments as in the CLT, and consider sums $S_n, S_{2n} - S_n, S_{3n} - S_{2n}, \ldots$, over consecutive disjoint blocks of n of these random variables X_m having mean μ and variance σ^2. The block sums are independent, each approximately Gaussian with mean $n\mu$ and variance $n\sigma^2$. If one scales the sums as $\frac{S_n - n\mu}{\sigma\sqrt{n}}, \frac{S_{2n} - S_n - n\mu}{\sigma\sqrt{n}} \ldots$, then in the limit one should get a process with independent Gaussian increments. If time is scaled so that one unit of time in the new *macroscopic* scale is equal to n units of time in the old scale, the $B_1, B_2 - B_1, B_3 - B_2, \ldots$ are independent Gaussian $\Phi_{0,1}$. Brownian motion is precisely such a process, but constructed for all times $t \geq 0$ and having continuous sample paths. The conception of such a process was previously introduced in a1900 PhD thesis by Bachelier[7] as a model for the movements of stock prices.

Brownian motion is named after the nineteenth-century botanist Robert Brown, who observed under the microscope perpetual irregular motions exhibited by small grains or particles of the size of colloidal molecules immersed in a fluid. Brown[8] himself credited earlier scientists for having made similar observations. After some initial speculation that the movements are those of living organisms was discounted, the movements were attributed to inherent molecular motions. Independently of this debate and unaware of the massive experimental observations that had been made concerning this matter, Einstein[9] published a paper in 1905 in which he derived the *diffusion equation*

$$\frac{\partial C(t,x)}{\partial t} = D\left(\frac{\partial^2 C(t,x)}{\partial x_1^2} + \frac{\partial^2 C(t,x)}{\partial x_2^2} + \frac{\partial^2 C(t,x)}{\partial x_3^2}\right), \quad x = (x_1, x_2, x_3), \qquad (13.1)$$

for the concentration $C(t,x)$ of large solute molecules of uniform size and spherical shape in a stationary liquid at a point x at time t. The argument (at least implicit in the above article) is that a solute molecule is randomly displaced frequently by collisions with the molecules of the surrounding liquid. Regarding the successive

---

[6]Lévy, P. (1925).

[7]Bachelier, L. (1900). Théorie de las spéculation. *Ann. Sci. École Norm. Sup.* **17**, 21–86. (In: *The Random Character of Stock Market Prices*, Paul H. Cootner, ed. MIT Press, 1964).

[8]Brown, R. (1828). A brief account of microscopical observations made in the months of June, July, and August, 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *Philos. Magazine N.S.* **14**, 161–173.

[9]Einstein, A. (1905). On the movement of small particles suspended in a stationary liquid demanded by the molecular–kinetic theory of heat. *Ann. der Physik* **17**, 549.

displacements as independent (and identically distributed) with mean vector zero and dispersion matrix $\mathrm{Diag}(d, d, d)$, one deduces a Gaussian distribution of the position of the solute molecule at time t with mean vector zero and a dispersion matrix $2t$ $\mathrm{Diag}(D, D, D)$, where $2D = fd$ with f as the average number of collisions, or displacements, per unit time. The law of large numbers (assuming that the different solute molecules move independently) then provides a Gaussian concentration law that is easily seen to satisfy the equation (13.1), away from the boundary. It is not clear that Laplace was aware of the profound fact that the operator $\Delta = \sum_1^3 \partial^2/\partial x_i^2$ in (13.1) bearing his name is intimately related to the central limit theorem he had derived.

Apprised of the experimental evidence concerning the so-called Brownian movement, Einstein titled his next article[10] on the subject, "On the theory of the Brownian movement." In addition to deriving the form of the equation (13.1), Einstein used classical thermodynamics, namely the Maxwell–Boltzmann steady-state (Gaussian) velocity distribution and Stokes' law of hydrodynamics (for the frictional force on a spherical particle immersed in a liquid) to express the *diffusion coefficient* D as $D = kT/3\pi\eta a$, where a is the radius of the spherical solute molecule, η is the coefficient of viscosity, T is the temperature, and k is the Boltzmann constant. In particular, the *physical parameters* are embodied in a *statistical parameter*. Based on this derivation, Jean Baptiste Perrin[11] estimated k or, equivalently, Avogadro's number, for which he was awarded the Nobel Prize in 1926. Meanwhile, in 1923, Wiener[12] proved that one may take Brownian paths to be continuous almost surely. That is, he constructed the probability measure Q, the so-called *Wiener measure* on $C[0, \infty)$, extending the normal distribution to infinitely many dimensions in the sense that the coordinate process $X_t(\omega) := \omega(t)$, $\omega \in C[0, \infty)$, $t \geq 0$, has independent Gaussian increments, namely, $X_{t+s} - X_t$ has the normal distribution $\Phi_{0,s} \equiv N(0, s)$, $\forall\ 0 \leq t < \infty$, $s > 0$, and $\{X_{t_{i+1}} - X_{t_i} : i = 1, 2, \ldots, m-1\}$ are independent $\forall\ 0 \leq t_1 < t_2 < \cdots < t_m$ $(\forall\ m > 1)$. This was a delicate result, especially since the Brownian paths turned out to have very little smoothness beyond continuity. Indeed, in 1933 it was shown by Paley, Wiener, and Zygmund[13] that with probability one, a Brownian path is continuous but nowhere differentiable. This says that a Brownian particle has no velocity, confirming some remarkable empirical observations in the early physics of Brownian motion. In his monograph "Atoms", Perrin exclaims: "The trajectories are confused and complicated so often and so rapidly that it is impossible to follow them; the trajectory actually measured is very much simpler and shorter than the real one. Similarly, the apparent mean speed of a grain during a given time varies *in the wildest way* in magnitude and direction, and does not tend to a limit as the time taken for an observation decreases, as may be easily shown by noting, in the camera lucida, the

[10] Einstein, A. (1906). On the theory of the Brownian movement. *Ann. der Physik* **19**, 371–381. English translation in *Investigations on the Theory of the Brownian Movement* (R. Fürth, ed.), Dover, 1954.

[11] Jean Perrin, *Atoms*, Ox Bow Press, 1990 (French original, 1913).

[12] Wiener, N. (1923). Differential space. *J. Math. Phys.* **2**, 131–174.

[13] Paley, R.E.A.C., Wiener, N. and Zygmund, A. (1933). Notes on random functions. *Math. Zietschr.* **37**, 647–668.

positions occupied by a grain from minute to minute, and then every five seconds, or, better still, by photographing them every twentieth of a second, as has been done by Victor Henri Comandon, and de Broglie when kinematographing the movement. It is impossible to fix a tangent, even approximately, at any point on a trajectory, and we are thus reminded of the continuous underived functions of the mathematicians."

A more dynamical theory of Brownian (particle) motion was given by Ornstein and Uhlenbeck,[14] following the turn-of-the-century work of Langevin.[15]

The so-called Langevin equation used by Ornstein and Uhlenbeck is a *stochastic differential equation* given (in one dimension) by

$$dv(t) = -\beta v(t)dt + \sigma dB(t), \tag{13.2}$$

where $v(t)$ is the velocity of a Brownian molecule of mass m, $-m\beta v$ is the frictional force on it, and $\sigma^2 = 2\beta^2 D$ (D as above). By integrating $v(t)$ one gets a differentiable model of the Brownian molecule. If β and $\sigma^2 \to \infty$ such that $s^2/2\beta^2 = D$ remains a constant, then the position process converges to Einstein's model of Brownian motion (with variance parameter $2D$), providing a scale range for which the models approximately agree.[16] Within the framework of stochastic differential equations one sees that the *steady state velocity distribution* for the Langevin equation is a Gaussian distribution. On physical grounds this can be equated with the Maxwell–Boltzmann velocity distribution known from statistical mechanics and thermodynamics. In this way one may obtain Einstein's fundamental relationship between the physical parameters and statistical parameters mentioned above.

Brownian motion is a central notion throughout the theoretical development of stochastic processes and its applications. This rich history and its remarkable consequences are brought to life under several different guises in major portions of the theory of stochastic processes.

---

[14] Uhlenbeck, G.E. and Ornstein, L.S. (1930). On the theory of Brownian motion. *Phys. Rev.* **36**, 823–841. Reprinted in *Selected Papers on Noise and Stochastic Processes* (1954). (N. Wax, ed.), Dover. Also see Chandrasekhar, S. (1943). Stochastic problems in physics and astronomy. *Rev. Modern Physics* **15**, 2–91. Reprinted in *Selected Papers on Noise and Stochastic Processes* (1954) (N. Wax, ed.), Dover.

[15] Langevin, P. (1908). Sur La théorie du movement brownien. *C.R. Acad. Sci. Paris* **146**, 530–533.

[16] For a complete dynamical description see Nelson, E. (1967). *Dynamical Theories of Brownian Motion.* Princeton Univ. Press, Princeton, N.J.