Measure-Theoretic Probability

Shiwei Chen

May 2024

Contents

1 Introduction

More than two and a half centuries ago, Probability Theory was first conceived by enterprising mathematicians helping gamblers to analyze games of chance. Since then the field has developed into a broad discipline with deep connections to other branches of mathematics such as physics, statistics, and economics. Today, the most rigorous studies in probability utilize measure theory, with the fundamental building block of modern probability being measure spaces $(X, \mathbb{M}, \mu)$ such that $\mu(X) = 1$.

2 Terminology

Probability theory has its own vocabulary whose terms were created before the connection with measure theory was made explicit. Below is a list of comparisons between the Analysts' and Probabilists' Terms for various Measure Theory concepts.

List of Terms	
Analysis Terms	Probability Terms
Measure Space $(X, \mathbb{M}, \mu)(\mu(X) = 1)$	Sample Space $(\Omega, \mathbb{B}, P)$
$(\sigma-)$ algebra	$(\sigma-)$ field
Measurable set	Event
Measurable real-valued function f	Random variable X
Integral of f	Expectation of mean of X, $E(X)$
Convergence in measure	Convergence in probability
Almost everywhere	Almost surely
Borel Probability Measure on $\mathbb{R}$	Distribution
Characteristic function	Indicator function

Typically, probabilists have an aversion to displaying the arguments of random variables. Measurable sets such as $P(\{w : X(w) > \alpha\})$ are displayed as $P(X > \alpha))$. In addition, we define the **variance** $\sigma^2(X)$ and **standard deviation** $\sigma(X)$ by the following formulas:

$$\sigma^2(X) = \inf_{a \in \mathbb{R}} E[(X - a)^2]$$

$$\sigma(X) = \sqrt{\sigma^2(X)}$$

Theorem 0.1: If $X \in L^2$, then $E[(X - \alpha)^2] = E[X^2 - 2\alpha X + \alpha^2] = E(X^2) - 2\alpha E(X) + \alpha^2$ is a quadratic function of α whose minimum occurs when $a = E(X)$. Hence,

$$\sigma^2(X) = E[(X - E(X))^2] = E[X^2 - E(X)^2] = E(X^2) - E(X)^2$$

Corollary 0.1: Given a Borel Probability measure λ on $\mathbb{R}$, let the **mean** be denoted as $\bar{\lambda}$. We obtain the following equivalences by Theorem 0.1:

$$\bar{\lambda} = \int t d\lambda(t), \sigma^2 = \int (t - \bar{\lambda})^2 d\lambda(t).$$

Distribution: If X is a random variable on Ω then P_X is a probability measure on $\mathbb{R}$ called the **distribution** of X, and the function

$$F(t) = P_X((-\infty, t]) = P(X \leq t)$$

is called the **distribution function** of X. If $\{X_\alpha\}$ is a family of random variables such that $P_{X_\alpha} = P_{X_\beta}$ for all $\alpha, \beta \in A$, the X_α are said to be **identically distributed.** More generally, for any finite sequence of $X_1, ..., X_n$ of random variables, we can consider $(X_1, ..., X_n)$ as a map from Ω to $\mathbb{R}^n$. and the measure $P_{(X_1,...,X_n)}$ on $\mathbb{R}^n$ is called the **joint distribution** of $X_1, ..., X_n$.

As a general principle, all properties of random variables that are relevant to probability theory can be expressed in terms of their joint distributions. For example, take **Proposition 0.1**:

$$E(X) = \int t dP_x(t), \sigma^2(X) = \int (t - E(X))^2 dP_X(t)$$

$$E(X + Y) = \int (t + s) dP_{(X,Y)}(t, s).$$

Stochastic Independence: Consider a probability space $(\Omega, \mathbb{B}, P)$ and an event E such that $P(E) > 0$. Then the set function $P_E(F) = P(E \cap F)/P(E)$ represents the probability of an event F given that E occurs. If $P_E(F) = P(F)$, that is, the probability of F is the same whether or not we restrict to E, then F is said to be **independent** of E. Thus, F is independent of E if and only if $P(E \cap F) = P(E)P(F)$.

Definition: A collection $\{X_\alpha\}_{\alpha \in A}$ of events in Ω to be **independent** if $P(E_{\alpha_1} \cap ... \cap E_{\alpha_n}) = \prod_1^n P(E_{\alpha_i})$ for all $n \in \mathbb{N}$ and all distinct $\alpha_1, ...\alpha_n \in A$.

3 Independent Random Variables

Many results exist for independent random variables, as one of the central focuses of Measure-Theoretic Probability.

Theorem 0.2: Let $\{X_{nj} : 1 \leq j \leq J(n), 1 \leq n \leq N.\}$ be a two dimensional list of random independent variables, and where $J(n)$ is some function. Let $f_n : \mathbb{R}^{J(n)} \to \mathbb{R}$ be Borel measurable for $1 \leq n \leq N$. Then the random variables $Y_n = f_n(X_{n1}, ..., X_{nJ(n)})$ are independent for $1 \leq n \leq N$.

Proof. Let $X_n = (X_{n1}, ..., X_{nJ(n)})$. If $B_1, ..., B_N$ are Borel subsets of $\mathbb{R}$, we have $Y_n^{-1}(B_n) = X_n^{-1}(f_n^{-1}(B_n))$ and hence,

$$\begin{aligned}(Y_1, ..., Y_N)^{-1}(B_1 \times ... \times B_N) &= \bigcap_1^N Y_n^{-1}(B_n) \\ &= (X_1, ..., X_N)^{-1}(f_1^{-1}(B_1) \times ... \times f_N^{-1}(B_N)).\end{aligned} \tag{1}$$

Therefore, by the independence of the X_{nj} and Fubini's theorem,

$$\begin{aligned}P_{(Y_1,...,Y_n)}(B_1 \times ... \times B_n) &= P_{(X_1,...,X_N)}(f_1^{-1}(B_1) \times ... \times f_N^{-1}(B_N) \\ &= \left(\prod_{n=1}^{N}\prod_{j=1}^{J(n)} P_{X_{n_j}}\right)(f_1^{-1}(B_1) \times ... \times f_N^{-1}(B_N)) \\ &= \prod_{n=1}^{N} P_{X_n}(f_n^{-1}(B_n)) \\ &= \prod_{n=1}^{N} P_{Y_n}(B_n).\end{aligned} \tag{2}$$

□

Definition: Let μ and v be measures on a Borel σ algebra on $\mathbb{R}^n$. The **convolution** of μ and v, denoted $\mu * v$ is the measure defined by

$$\mu * v(B) := \int \chi_B(x + y)d\mu(x)dv(y)$$

where χ_B is the indicator (characteristic) function of B. Using this definition of convolution, it can be shown that if $\lambda_1, ..., \lambda_n \in M(\mathbb{R})$, then $\lambda_1 * ... * \lambda_n$ is given by the equation:

$$\lambda_1 * ... * \lambda_n(E) = \int ... \int \chi_E(t_1 + ... + t_n)d\lambda_1(t_1)...d\lambda_n(t_n). \tag{3}$$

Theorem 0.3: If $\{X_j\}_1^n$ are independent random variables, then

$$P_{X_1+...+X_n} = P_{X_1} * ... * P_{X_n}$$

Proof. Let $A(t_1, ..., t_n) = \sum_1^n t_j$. This implies $X_1 + ... + X_n = A(X_1, ..., X_n)$, so

$$P_{X_1+...+X_n} = (P_{(X_1,...,X_n)})_A = \left(\prod_1^N P_{X_j}\right)_A,$$

and by equation (3), the last expression equals $P_1 * ... * P_n$. This proves the first assertion, and once the first assertion is known, the same argument, with the absolute values removed, proves the second one. □

Theorem 0.4: Suppose that $\{X_j\}_1^n$ are independent random variables. If $X_j \in L^1$ for all j, then $\prod_1^n |X_j| \in L^1$ and $E(\prod_1^n X_j) = \prod_1^n E(X_j)$.

Proof. We have $\prod_1^n |X_j| = f(X_1, ..., X_n)$ where $f(t_1, ..., t_n) = \prod_1^n |t_j|$. Hence,

$$\begin{aligned} E\left(\prod_1^n |X_j|\right) &= \int f dP_{(X_1,...,X_n)} = \int f d\left(\prod_1^n P_{X_j}\right) \\ &= \prod_1^n \int |t_j| dP_{X_j}(t_j) = \prod_1^n E(|X_j|) \end{aligned} \tag{4}$$

□

Theorem 0.5: If $\{X_j\}_1^n$ are independent and in L^2 then $\sigma^2(X_1+...+X_n) = \sum_1^n \sigma^2(X_j)$.

Proof. Let $Y_j = X_j - E(X_j)$. Then $\{Y_j\}_1^n$ are independent and have mean zero, so

$$E(Y_j Y_k) = E(Y_j)E(Y_k) = 0, (j \neq k).$$

Therefore,

$$\begin{aligned} \sigma^2(X_1 + ... + X_n) = E((Y_1 + ... + Y_n)^2) &= \sum_{j,k} E(Y_j Y_k) \\ = \sum_j E(Y_j^2) &= \sum_j \sigma^2(X_j). \end{aligned} \tag{5}$$

□

4 The Law of Large Numbers

If one plays a gambling game many times, one's average winnings or losses per game should be roughly the expected winnings or losses in each individual game. In symbols, if $\{X_j\}_1^\infty$ is a sequence of independent random variables and $E(X_j) = \mu_j$, then the average $n^{-1}\sum_1^n X_j$ should be close to the constant $n^{-1}\sum_1^n \mu_j$ when n is large. The Law of Large numbers comes in several versions, depending on the hypotheses one wishes to make.

The Weak Law of Large Numbers: Let $\{X_j\}_1^\infty$ be a sequence of independent L^2 random variables with means $\{\mu_j\}$ and variances $\{\sigma^2\}$. If $n^{-2}\sum_1^n \sigma_j^2 \to 0$ as $n \to \infty$, then $n^{-1}\sum_1^n (X_j - \mu_j) \to 0$ in probablity as $n \to \infty$.

Proof. $n^{-1}\sum_1^n (X_j \mu_j)$ has mean 0 and by Theorem 0.5, has variance $n^{-2}\sum_1^n \sigma_j^2$. By Chebyshev's inequality, for any $\epsilon > 0$ we have

$$P\left(|n^{-1}\sum_1^n (X_j - \mu_j)| > \epsilon\right) \leq (n\epsilon)^{-2} \sum_1^n \sigma_j^2 \to 0 \text{ as } n \to \infty.$$

□

Under stronger conditions, we can prove that $n^{-1}\sum_1^n(x_j-\mu_j)\to 0$ almost surely. First, we need another theorem. **Kolmogorov's Inequality:** Let $X_1,...,X_n$ be independent random variables with mean 0 and variances $\sigma_1^2,...,\sigma_n^2$ and let $S_k=X_1+...+X_k$. For any $\epsilon>0$,

$$P\left(\max_{1\le k\le n}|S_k|\ge\epsilon\right)\le\epsilon^{-2}\sum_1^n\sigma_k^2.$$

Proof. Let A_k be the set where $|S_j|<\epsilon$ for $j<k$ and $|S_k|\ge\epsilon$. Then the A_k's are disjoint and their union is the set where $\max|S_k|\ge\epsilon$, so

$$P(\max|S_k|\ge\epsilon)=\sum_1^n P(A_k)\le\epsilon^{-2}\sum_1^n E(\chi_{A_k}S_k^2)$$

because $S_k^2\ge\epsilon^2$ on A_k. On the other hand,

$$\begin{aligned}E(S_n^2)&\ge\sum_1^n E(\chi_{A_k}S_n^2)\\&=\sum_1^n E\left(\chi_{A_k}[S_k^2+2S_k(S_n-S_k)+(S_n-S_k)^2]\right)\\&\ge\sum_1^n E(\chi_{A_k}S_k^2)+2\sum_1^n E(\chi_{A_k}S_k(S_n-S_k)).\end{aligned}\tag{6}$$

It suffices to show that $E(\chi_{A_k}S_k(S_n-S_k))=0$ for all k, for then we have

$$P(\max|S_k|\ge\epsilon)\le\epsilon^{-2}E(S_n^2)=\epsilon^{-2}\sum_1^n\sigma_k^2$$

by Theorem 0.5, since the X_k's have mean zero. But χ_{A_k} is a measurable function of $S_1,...,S_k$ and hence of $X_1,...,X_k$, wheras S_n-S_k is a measurable function of $X_{k+1},...,X_n$. Moreover, $E(S_k)=\sum_1^k E(X_j)=0$ for all k. Therefore, by Theorems 0.2 and 0.4,

$$E(\chi_{A_k}S_k(S_n-S_k))=E(\chi_{A_k}S_k)E(S_n-S_k)=E(\chi_{A_k}S_k)\cdot 0=0.$$

□

Kolmogorov's Strong Law of Large Numbers: If $\{X_n\}_1^\infty$ is a sequence of independent L^2 random variables with means $\{\mu_n\}$ and variances $\{\sigma_n^2\}$ such that $\sum_1^\infty n^{-2}\sigma_n^2<\infty$, then $n^{-1}\sum_1^n(X_j-\mu_j)\to 0$ almost surely as $n\to\infty$.

Proof. Let $S_n=\sum_1^n(X_j-\mu_j)$. Given $\epsilon>0$ for $k\in\mathbb{N}$ let A_k be the set where $n^{-1}|S_n|\ge\epsilon$ for some n such that $2^{k-1}\le n<2^k$. Then on A_k we have $|S_n|\ge\epsilon 2^{k-1}$ for some $n<2^k$. By Kolmogorov's Inequality,

$$P(A_k) \leq (\epsilon 2^{k-1})^{-2} \sum_1^{2k} \sigma_n^2.$$

Therefore,

$$\sum_1^\infty P(A_k) \leq \frac{4}{\epsilon^2} \sum_{k=1}^\infty \sum_{n=1}^{2^k} 2^{-2k} \sigma_n^2 = \frac{4}{\epsilon^2} \sum_{n=1}^\infty \left(\sum_{k \geq \log_2 n} 2^{-2k} \right) \sigma_n^2 \leq \frac{8}{\epsilon^2} \sum_{n=1}^\infty \frac{\sigma_n^2}{n^2} < \infty,$$

so $P(\limsup A_k) = 0$ by the Borel-Cantelli lemma. But $\limsup A_k$ is precisely the set where $n^{-1}|S_n| \geq \epsilon$ for infinitely many n, so

$$P(\limsup n^{-1}|S_n| < \epsilon) = 1.$$

Letting $\epsilon \to 0$ through a countable sequence of values, we conclude that $n^{-1}S_n \to 0$ almost surely. □

5 Central Limit Theorem

Proposition 2: If $a > 0$,

$$\int_{\mathbb{R}^n} \exp(-a|x|^2) dx = \left(\frac{\pi}{a}\right)^{n/2}.$$

Suppose $\mu \in \mathbb{R}$ and $\sigma > 0$. Taking proposition 2 and using some elementary calculus, we can show that the measure $v_\mu^{\sigma^2}$ on $\mathbb{R}$ defined by

$$dv_\mu^{\sigma^2}(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{(t-\mu)^2/2\sigma} dt$$

is a probability measure that satisfies

$$\int t dv_\mu^{\delta^2}(t) = \mu, \int (t-\mu)^2 dv_\mu^{\sigma^2}(t) = \sigma^2.$$

$dv_\mu^{\sigma^2}$ is called the **normal** or **Gaussian distribution** with mean μ and variance σ^2. The special case v_0^1 is called the **standard normal distribution.** Normal and approximately normal distributions are extremely common in applied probability and statistics, and the theoretical explanation for this phenomenon is the central limit theorem.

Definition: A sequence $\mu_n\}$ is said to **converge vaguely** to μ if there exists a dense subset $D \subset \mathbb{R}$ such that

$$\forall a, b \in D, a < b, \mu_n(a, b] \to \mu(a, b].$$

Lemma 0.1 Let λ be a Borel probability measure on $\mathbb{R}$ such that

$$\int t^2 d\lambda(t) = 1, \int t d\lambda(t) = 0.$$

For $n \in \mathbb{N}$ let $\lambda^{*n} = \lambda * ... * \lambda$ (n factors) and define the measure λ_n by $\lambda_n(E) = \lambda^{*n}(\sqrt{n}E)$ where $\sqrt{n}E = \{\sqrt{n}t : t \in E\}$. Then $\lambda_n \to v_0^1$ vaguely as $n \to \infty$.

Proof. The proof involves theorems and terminology from Fourier Analysis outside of the scope of this project. The general idea is to take the fact that the Fourier transform of λ and then to use Taylor's Theorem to get

$$\widehat{\lambda(\zeta)} = 1 - 2\pi^2\zeta^2 o(\zeta^2).$$

By making an equivalence between convolutions and multiplication, we have $(\widehat{\lambda^{*n}}) = (\widehat{\lambda})^n$. Making a change of variable, we can obtain

$$-\pi^2\zeta^2 + n \cdot o(\frac{\zeta^2}{n})$$

which tends to $-2\pi^2\zeta^2$ as $n \to \infty$. Then, applying yet more Fourier Analysis theorems, we get that $\lambda_n \to v_0^1$ vaguely as $n \to \infty$. $\square$

Lemma 0.2: If $\mu_n \to \mu$ vaguely and μ_n is positive, $\forall n \in \mathbb{N}$, then $F_n(x) \to F(x)$ at every x at which F is continuous.

Proof. $\mu_n > 0$ implies $\mu \geq 0$ and thus that F is continuous at $x = a$. If $f \in C_c(\mathbb{R})$ is the function that is 1 on $[-N, a]$, 0 on $(-\infty, -N - \epsilon)$ and $[a + \epsilon, \infty]$ and linear in betweem, we have

$$\begin{aligned} F_n(a) - F_n(-N) = \mu_n((-N - a]) \leq \int f d\mu_n \to \int f d\mu \\ \leq F(a + \epsilon) - F(-N - \epsilon). \end{aligned} \tag{7}$$

As $N \to \infty, F_n(-N)$ and $F(-N - \epsilon)$ tend to zero, we have

$$\lim_{n\to\infty} \sup F_n(a) \leq F(a + \epsilon).$$

Similarly, by considering the function that 1 is on $[-N + \epsilon, a + \epsilon]$, 0 on $(-\infty, N]$ and $[a, \infty)$ and linear in between, we see that $\lim_{n\to\infty} \inf F_n(a) \to F(a)$ as desired. $\square$

Central Limit Theorem. Let $\{X_j\}$ be a sequence of independent identically distributed L^2 random variables with mean μ and variance σ^2. As $n \to \infty$, the distribution of $(\sigma\sqrt{n})^{-1}\sum_1^n(X_j - \mu)$ converges vaguely to the standard normal distribution v_0^1 and for all $a \in \mathbb{R}$,

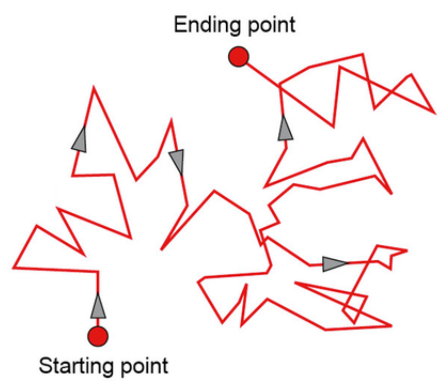
$$\lim_{n\to\infty} P\left(\frac{1}{\sigma\sqrt{n}}\sum_1^n(X_n - \mu) \leq a\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^a e^{-t^2/2}dt.$$

Proof. Replacing X_j by $\sigma^{-1}(X_j - \mu)$ we may assume that $\mu = 0$ and $\sigma = 1$. If λ is the common distribution of the X_js then λ satisfies the hypothesis of Lemma 1, and in the notation used there, λ_n is the distribution of $n^{-1/2}\sum_1^n X_j$. This therefore implies that $(\sigma\sqrt{n})^{-1}\sum_1^n(X_j - \mu)$ converges vaguely to the standard normal distribution v_0^1. To prove convergence for all $a \in \mathbb{R}$, we use Lemma 0.2 to make it equivalent to the first assertion. $\square$

6 Bonus: Brownian Motion

It was observed by biologist Robert Brown in 1827 that small particles suspended in a fluid such as water or air undergo an irregular motion now named **Brownian motion**. Nearly a century later in 1905 Albert Einstein described Brownian Motion using a probabilstic model, observing that if the kinetic energy of fluids was right, the molecules of water moved as Brown had described it, if given a random bombardment by the molecules in the fluid.

The mathematical models used to describe these random movements heavily rely upon measure theory as a foundation, and will be explored in this section. Specifically, we will explore the limiting case where the motion is assumed to result from an infinite number of collisions with molecules of infinitesimal size. Aplications of Brownian motion have become extremely diverse, reaching nearly every field related to mathematics from physics to statistics and economics.



One can consider Brownian motion in any number of space dimensions, but we will consider its theory in one dimension first. First, as a matter of normalization, we assume that the particle starts at origin time $t = 0$ and also that $X_0 = 0$ almost surely **(Condition 1)**.

Second, since any given collision affects the particle only by an infinitesimal amount, it has no long-term effect, so the motion of the particle after time t should depend on its position X_t at that time but not on its previous history. Thus we assume If $0 \leq t_0 < t_1 < ... < t_n$, then the random variables $X_{t_j} - X_{t_{j-1}}, (1 \leq j \leq n)$ are independent **(Condition 2).**

Since the physical processes underlying Brownian motion are homogeneous in time, we can further postulate that the distribution of $X_t - X_s$ depends only on $t - s$. If we divide the interval $[s, t]$ into n equal subintervals $[t_0, t_1], ..., [t_{n-1}, t_n](t_0 = s, t_n = t)$ and write $X_t - X_s = \sum_1^n (X_{t_j} - X_{t_{j-1}})$ it then follows from Condition 2 that $X_t - X_s$ is the sum of n independent identically distributed random variables. Since n can be taken arbitrarily large, the central limit theorem suggests that the distribution of $X_t - X_s$ should be normal. Finally, since the particle is as likely to move to the left as to the right, the mean of $X_t - X_s$ should be 0. Putting this all together, we are led to the third assumption: There is a constant $C > 0$ such that for $0 < s < t, X_t - X_s$ has the normal distribution $v_0^{C(t-s)}$ with mean 0 and variance $C(t-s)$ **(Condition 3).**

Definition: A Stochastic Process is a parameterized collection of random variables $\{X_t\}_{t \in T}$ defined on a probability space and assuming values in $\mathbb{R}^n$.

An **n-dimensional abstract Wiener process** is a stochastic process $\{X_t\}$ where $X_t = (X_t^1, ..., X_t^n)$ such that each $\{X_t^j\}$ fulfills conditions 1 through 3 and (ii): if Y_j is any function of the variables $\{X_t^j\}_{t>0}$ for $j = 1, ..., n$ then $Y_1, ..., Y_n$ are independent. We call the one dimensional case an **abstract Wiener process.**

Most real-world Wiener/Brownian motion processes exhibit the property that over time, the mean of the process slowly (and constantly) shifts either upwards or downwards. This property is known as the **drift parameter** and is represented by μ. By adding drift into our Brownian Motion, we get

$$Y(t) = \mu t + cX(t)$$

where $X(t)$ is a standard Brownian motion process defined earlier. The most prominent applications of Brownian motion today involve its relevance to stock prices and stock options.

Definition: A Stochastic Process is said to follow a **Geometric Brownian Motion** if it satisfies a stochastic differential equation:

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

where W_t is a Brownian Motion, μ is the constant drift, and σ is the constant percentage volatility (standard deviation). By definition, a Geometric Brownian motion process $Z(t)$ is a stochastic process such that $W(t) = \log Z(t)$ is a Brownian Motion process with variance paramter c^2 and drift parameter $\mu = \alpha - \frac{1}{2}c^2$ where α is the drift parameter for $Z(t)$. Using this idea, we can write any geometric Brownian motion process $Z(t)$ with initial value $Z(0) = z$ as a function of a standard Brownian motion process $X(t)$ in the following way:

$$Z(t) = ze^{W(t)} = ze^{(\alpha - 1/2c^2)t + cX(t)}.$$

Definition: Define T_{ab} to be the time at which the Brownian Motion process exits the interval $[a, b]$ or more precisely,

$$T_{ab} = \min[t \geq 0; Y(t) = a \text{ or } Y(t) = b].$$

Because we want to make money, we are interested in the probability of T_{ab} bringing the process to the high side of the interval, namely b. According to

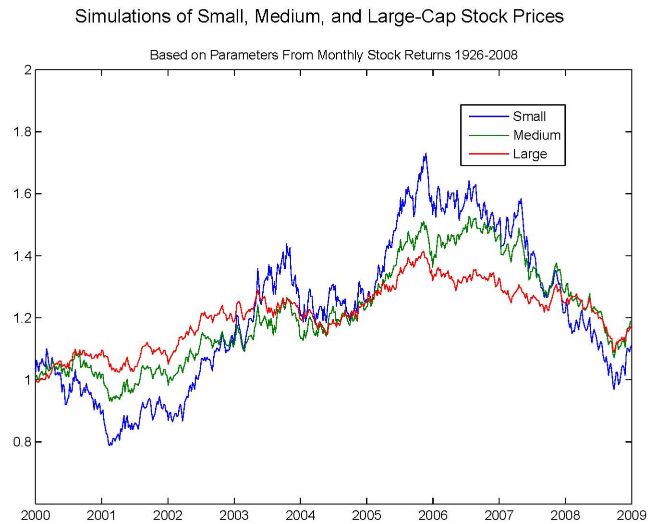


Figure 1: Simulated geometric Brownian motions with parameters from market data

Taylor and Karlin, this probability is

$$P[Y(T_{ab})|Y(0) = x] = \frac{e^{-2x\mu/c^2} - e^{-2\alpha\mu/c^2}}{e^{-2b\mu/c^2} - e^{-2\alpha\mu/c^2}}$$

(See: proof on Taylor page 509). Combining this formula with the Brownian Motion $W(t)$ defined above, we obtain a surprisingly good model for stock market behavior. Firstly, $Z(t)$ can never be negative, which is important if one wants to model the behavior of a stock or other market entity. In addition, $Z(t)$ follows a long term exponential decay or growth trajectory, due to the presence of the e, which more accurately describes many situations in trading.

As a simple example, suppose Gamestop (GME) is expecting a growth of ten percent, and on any given week of trading, the fluctuation of the stock is given by $c^2 = 5$. Suppose a shareholder buys ten shares of Gamestop at a price of 50 dollars and plans on selling them if the prices increases to 80 dollars or drops to 35 dollars, what is the probability that the shareholder sells his shares at a profit?

Let's convert Gamestop's growth to a weekly growth, since that is the unit of time being dealt with in this situation. We would therefore have $\mu = .10/52 = .002$ and $2\mu/c^2 = 2(0.002)/5 = 0.0008$. Plugging in this information with

$x = 50, a = 35, b = 80$:

$$P(\text{profit}) = \frac{e^{-(50)(0.0008)} - e^{-(35)(0.0008)}}{e^{-(80)(0.0008)} - e^{-(35)(0.0008)}} = 0.337.$$

In other words, our investor is more likely to lose money than to make a profit. What would happen if the shareholder decided to sell sooner? Suppose he decides to sell at $b = 55$. We see that

$$P(\text{profit}) = \frac{e^{-(50)(0.0008)} - e^{-(35)(0.0008)}}{e^{-(55)(0.0008)} - e^{-(35)(0.0008)}} = 0.750.$$

This new likelihood is indeed much more favorable to the investor, and shows that even when a Brownian motion process is drifting upward, the short-term fluctuations can affect the process enough to temporarily nullify this growth, given the stockholder decides to sell under a specific price.

7 References

Duncan, Thomas. "Brownian Motion: A Study of Its Theory and Applications." eScholarship, Boston College, 2007, dlib.bc.edu/islandora/object/bc-ir:102098/datastream/PDF/view.

Ermogenous, Angeliki. "Brownian Motion and Its Applications in the Stock Market." Ecommons UDayton, University of Dayton, 2006. ecommons.udayton.edu/cgi/viewcontent.cgi?article=1010&context=mth_epumd.

FOLLAND, Gerald B. Real Analysis: Modern Techniques and Their Applications. Second Edition ed., vol. 1 1, John Wiley and Sons, 1994.

"Convolution (Measure Theory)." ProofWiki, https://proofwiki.org/wiki/Definition:Convolution_of_Measures. Accessed 17 May 2024.

Taylor, Howard M. and Samuel Karlin. An Introduction to Stochastic Modeling. San Diego: Academic Press, 1998.