IEOR 4701: Stochastic Models in FE

Summer 2007, Professor Whitt

Class Lecture Notes: Wednesday, July 11.

The Central Limit Theorem and Stock Prices

1. **The Central Limit Theorem (CLT)**

See Section 2.7 of Ross.

(a) Time on My Hands:

Suppose that I have a lot of time on my hands, e.g., because I am on a New Jersey Transit train in the tunnel under the Hudson river waiting for a disabled Amtrak train ahead of me to be removed. Fortunately, I have a coin in my pocket. And now I decide that this is an ideal time to see if heads will come up half the time in a large number of coin tosses. Specifically, I decide to see what happens if I toss a coin many times. Indeed, I toss my coin $1,000,000$ times. **Below are various possible outcomes**, i.e., various possible numbers of heads that I might report having observed:

1. 500,000
2. 500,312
3. 501,013
4. 511,062
5. 598,372

What do you think of these reported outcomes? How believable are these outcomes? How likely are these outcomes?

We rule out outcome 5; there are clearly too many heads. We rule out outcome 1; it is "too perfect." Even though $500,000$ is the most likely single outcome, it itself is extremely unlikely. But how do we think about the remaining three?

The other possibilities require more thinking. We can answer the question by doing a **normal approximation**; see Section 2.7 of Ross, especially pages 79-83.

We introduce a probability model. We assume that successive coin tosses are independent and identically distributed (commonly denoted by IID) with probability of 1/2 of coming out heads. Let S_n denote the number of heads in n coin tosses. The random variable S_n is approximately normally distributed with mean $np = 500,000$ and variance $np(1-p) = 250,000$. Thus S_n has standard deviation $SD(S_n) = \sqrt{Var(S_n)} = 500$. Case 2 looks likely because it is less than 1 standard deviation from the mean; case 3 is not too likely, but not extremely unlikely, because it is just over 2 standard deviations from the mean. On the other hand, Case 4 is extremely unlikely, because it is over 20 standard deviations from the mean. See the Table on page 81 of the text.

(b) The Power of the CLT

The normal approximation for the binomial distribution with parameters (n, p) when n is not too small and the normal approximation for the Poisson with mean λ when λ is not too small are both special cases of the **central limit theorem** (**CLT**). The CLT states that a properly normalized sum of random variables *converges in distribution* to the normal distribution.

Of course there are conditions. We give a formal statement; see Theorem 2.2 on p. 79 of Ross. For that purpose, let $N(m, \sigma^2)$ denote a random variable having a normal distribution with mean m and variance σ^2. Let $\Rightarrow$ denote convergence in distribution.

Theorem 0.1 (central limit theorem (CLT)) *Suppose that* $\{X_n : n \geq 1\}$ *is a sequence of independent and identically distributed* (*IID*) *random variables, each distributed as* X. *Form the partial sums*

$$S_n \equiv X_1 + \cdots + X_n \quad \textit{for} \quad n \geq 1 \ .$$

If $E[X^2] < \infty$ *or, equivalently, if* $\sigma^2 \equiv Var(X) < \infty$ *(which implies that the mean is finite), then*

$$\frac{S_n - E[S_n]}{\sqrt{Var(S_n)}} \Rightarrow N(0,1) \quad \textit{as} \quad n \to \infty \ ,$$

i.e.,

$$P\left(\frac{S_n - E[S_n]}{\sqrt{Var(S_n)}} \leq x\right) \to P(N(0,1) \leq x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2} \, dy$$

as $n \to \infty$ *for each* x.

Where does the sum appear in our application? A random variable that has a binomial distribution with parameters (n, p) can be regarded as the sum of n IID random variables with a Bernoulli distribution having parameter p; each of these random variables X_i assumes the value 1 with probability p and assumes the value 0 otherwise. A random variable having a Poisson distribution with mean λ can be regarded as the sum of n IID random variables, each with a Poisson distribution with mean λ/n (for any n).

And what about the normalization? We simply subtract the mean of S_n and divide by the standard deviation of S_n to make the normalized sum have mean 0 and variance 1. Note that

$$\frac{S_n - E[S_n]}{\sqrt{Var(S_n)}} = \frac{S_n - n\mu}{\sqrt{n\sigma^2}} \tag{1}$$

has mean 0 and variance 1 whenever

$$S_n \equiv X_1 + \cdots + X_n \ ,$$

where $\{X_n : n \geq 1\}$ is a sequence of IID random variables with mean μ and variance σ^2. (It is crucial that the mean and variance be finite.)

The CLT applies much more generally; it has remarkably force. **The random variables being added do not have to be Bernoulli or Poisson; they can have any distribution.** We only require that the distribution have finite mean μ and variance σ^2. The statement of a basic CLT is given in Theorem 2.2 on p. 79 of Ross. The conclusion actually holds under even weaker conditions. The random variables being added do not actually have to be independent;

it suffices for them to be "weakly dependent;" and the random variables do not have to be identically distributed; it suffices for no single random variable to be large compared to the sum. But the statement then need adjusting: the first expression in (1) remains valid, but the second does not.

What does the CLT say? The precise mathematical statement is a **limit** as $n \to \infty$. It says that, as $n \to \infty$, the normalized sum in (1) **converges in distribution** to $N(0,1)$, a random variable that has a normal distribution with mean 0 and variance 1, whose distribution is given in the table on page 81 of our textbook. (Let $N(a,b)$ denote a normal distribution with mean a and variance b.) What does convergence in distribution mean? It means that the cumulative distribution functions (cdf's) converge to the cdf of the normal limit, denoted by

$$\frac{S_n - E[S_n]}{\sqrt{Var(S_n)}} \Rightarrow N(0,1) \ ,$$

which means that

$$P\left(\frac{S_n - E[S_n]}{\sqrt{Var(S_n)}} \le x\right) \to P\left(N(0,1) \le x\right) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2}\, dy$$

for all x. Note that convergence in distribution means convergence of cdf's, which means convergence of functions.

How do we apply the CLT? We *approximate* the distribution of the normalized sum in (1) by the distribution of $N(0,1)$. The standard normal (with mean 0 and variance 1) has no parameters at all; its distribution is given in the Table on page 81. By scaling, we can reduce other normal distributions to this one. The approximation is

$$\frac{S_n - E[S_n]}{\sqrt{Var(S_n)}} \approx N(0,1) \ ,$$

which, upon undoing the normalization becomes

$$S_n \approx E[S_n] + \sqrt{Var(S_n)} N(0,1) \stackrel{\mathrm{d}}{=} N(E[S_n], Var(S_n)) \ .$$

As a consequence of the CLT, we conclude that S_n is approximately normally distributed with its true mean and variance. The CLT states that the distribution is approximately normal, regardless of the distribution of the underlying random variables X_i. The CLT helps explain why the normal distribution arises so often.

2. An Application of the CLT: Modelling Stock Prices

Given the generality of the CLT, it is nice to consider an application where the random variables being added in the CLT are not Bernoulli or Poisson, as in many applications. Hence we consider such an application now.

(a) An Additive Random Walk Model for Stock Prices

We start by introducing a random-walk (RW) model for a stock price. Let S_n denote the price of some stock at the end of day n. We then can write

$$S_n = S_0 + X_1 + \cdots + X_n \ , \tag{2}$$

where X_i is the *change* in stock price between day $i-1$ and day i (over day i) and S_0 is the initial stock price, presumably known (if we start at current time and contemplate the evolution of the stock price into the uncertain future. We are letting the index n count days, but we could have a different time unit.

We now make a probability model. We do so by assuming that the successive changes come from a sequence $\{X_n : n \geq 1\}$ of IID random variables, each with mean μ and variance σ^2. This is roughly reasonable. Moreover, we do not expect the distribution to be Bernoulli or Poisson. The stochastic process $\{S_n : n \geq 0\}$ is a **random walk** with steps X_n, but a general random walk. If the steps are Bernoulli random variables, then we have a simple random walk, as discussed in Chapter 4, in particular, in Example 4.5 on page 183 and Example 4.15. But here the steps can have an arbitrary distribution.

We now can apply the CLT to deduce that the model implies that we can approximate the stock price on day n by a normal distribution. In particular,

$$P(S_n \leq x) \approx P(N(S_0 + n\mu, n\sigma^2) \leq x) = P(N(0,1) \leq (x - S_0 - n\mu)/\sigma x) \ .$$

How do we do that last step? Just re-scale: subtract the mean from both sides and then divide by the standard deviation for both sides, inside the probabilities. The normal variable is then transformed into $N(0,1)$. We can clearly estimate the distribution of X_n by looking at data. We can investigate if the stock prices are indeed normally distributed.

(b) A Multiplicative Model for Stock Prices

Actually, many people do not like the previous model, because they believe that the change in a stock price should be somehow proportional to the price. (There is much much more hard-nosed empirical evidence, not just idle speculation.) That leads to introducing an alternative multiplicative model of stock prices. Instead of (2) above, we assume that

$$S_n = S_0 \times X_1 \times \cdots \times X_n \ , \tag{3}$$

where the random variables are again IID, but now they are random daily multipliers. Clearly, the random variable X_n will have a different distribution if it is regarded as a multiplier instead of an additive increment.

But, even with this modification, we can apply the CLT. We obtain an additive model again if we simply take logarithms (using any base, but think of standard base $e = 2.71828....$). Note that

$$\log(S_n) = \log(S_0) + \log(X_1) + \cdots + \log(X_n) \ , \tag{4}$$

so that, by virtue of the CLT above,

$$\log(S_n) \approx N(\log(S_0) + n\mu, n\sigma^2) \ , \tag{5}$$

where now (with this new interpretation of X_n)

$$\mu \equiv E[\log(X_1)] \quad \text{and} \quad \sigma^2 \equiv Var(\log(X_1) \ . \tag{6}$$

As a consequence, we can now take exponentials of both sides of (5) to deduce that

$$S_n \approx e^{(N(\log(S_0)+n\mu, n\sigma^2)} \ . \tag{7}$$

That says that S_n has a **lognormal distribution**. Some discussion of this model appears on page 608 of our textbook. It underlies *geometric Brownian motion*, one of the fundamental stochastic models in finance.

3. Advanced (Optional) Topic: Stochastic-Process Limits

There exists a generalization of the CLT that explains Brownian motion and geometric Brownian motion. With appropriate scaling of time and space, the entire random walk (additive model) converges to Brownian motion, while an appropriate sequence of the corresponding multiplicative models converges to geometric Brownian motion.

To briefly explain, consider the simple case in which the IID random variables X_i have mean 0 and variance 1. Then the CLT says that $S_n/\sqrt{n} \Rightarrow N(0,1)$ as $n \to \infty$. But there is a more general result, which implies the CLT as a special case. The whole random walk is the sequence $\{S_k : k \geq 0\}$. With appropriate scaling of time and space, this random walk converges to Brownian motion $\{B(t) : t \geq 0\}$, discussed in Chapter 10 of Ross. The scaling creates a new scaled random walk for each n. The sequence of scaled random walks, generated from the one initial random walk, converges to Brownian motion as $n \to \infty$. Brownian motion is a continuous-time stochastic process having continuous sample paths and independent increments, with $B(t)$ distributed as $N(0,t)$ for each t. (See Chapter 10 in Ross.) Thus $B(1)$ is distributed as $N(0,1)$.

Given the entire random walk, $\{S_k : k \geq 0\}$ we can get convergence to Brownian motion by considering a sequence of stochastic processes with scaling depending upon n. For each n, we scale time by n and space by $\sqrt{n}$. To do so, consider one fixed n. For n given, plot $S_k/\sqrt{n}$ at time k/n for all $k \geq 0$. That compresses time by the factor n, but scales space by dividing by $\sqrt{n}$. In other words, consider the continuous-time stochastic process $\{S_{\lfloor nt \rfloor}, t \geq 0\}$, where $\lfloor x \rfloor$ is the greatest integer less than or equal to x; i.e., $\lfloor nt \rfloor = k/n$ for $k/n \leq t < (k+1)/n$. The generalized "functional" CLT concludes that the entire stochastic process $\{S_{\lfloor nt \rfloor}/\sqrt{n}, t \geq 0\}$ converges in distribution to Brownian motion $\{B(t), t \geq 0\}$ as $n \to \infty$:

$$\{S_{\lfloor nt \rfloor}/\sqrt{n}, t \geq 0\} \Rightarrow \{B(t), t \geq 0\} \quad \text{as} \quad n \to \infty \ .$$

Mathematically, there is a question about what the convergence $\Rightarrow$ means in this more general context. It is interpreted as convergence in distribution, but the objects should be interpreted as random functions. These more general limits are stochastic-process limits; see Chapter 1 of my book, **Stochastic-Process Limits**, available online at: http://www.columbia.edu/∼ww2040/book.htm

As a corollary to the stochastic-process limit, by considering what happens at one time point t, we get

$$S_{\lfloor nt \rfloor}/\sqrt{n} \Rightarrow B(t)$$

for each t. If we consider the single time point $t = 1$, then we get the ordinary CLT:

$$S_n/\sqrt{n} \Rightarrow B(1) \ .$$

That implies the CLT because $B(1)$ is distributed as $N(0,1)$.

Paralleling this generalization of the CLT, there is a limit for sequences of multiplicative models in which the limit process is the exponential of Brownian motion, called geometric Brownian motion. Thus the CLT explains the prevalence of geometric Brownian motion in finance models. The convergence of the sequence of multiplicative models to geometric Brownian motion is discussed in Section 2.1 in Chapter 2 of my book. We will come back to this topic in the second half of the course.