

IEOR 3106: Introduction to OR: Stochastic Models

Fall 2013, Professor Whitt

Class Lecture Notes: Tuesday, September 10.

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 subway travelling the full length of the subway system. 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 each of these possible outcomes? How likely are these outcomes? (Assume that I would report only one of these outcomes after doing the experiment. Assume that I would actually do such an experiment.)

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 three possibilities (alternatives 2 – 4) require more thinking. 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. Observe that S_n has *exactly* (according to the model) a *binomial distribution*.

Carefully examining the too-perfect case. First, here in these notes we quantify just how unlikely is the “too perfect” outcome of exactly the mean value 500,000 in 10^6 tosses. If the probability of heads in one toss were p , then we can exploit properties of the binomial distribution to conclude that the probability of k heads in n tosses would be

$$P(S_n = k) = b(k; n, p) = \frac{n!}{k!(n-k)!} p^k (1-p)^{(n-k)}$$

Now we are interested in the case and $p = 1/2$ and $k = n/2$ (for n even), i.e.,

$$P(S_n = n/2) = b(n/2; n, 1/2) = \frac{n!}{(n/2)!(n/2)!} (1/2)^n.$$

It is good to be able to roughly estimate these probabilities. To do so, we can use Stirling's formula (see p. 146):

$$n! \approx \sqrt{2\pi n} (n/e)^n.$$

We thus see that

$$P(S_n = n/2) = b(n/2; n, 1/2) \approx \frac{\sqrt{2\pi n} (n/e)^n}{(\pi n) (n/2e)^n} (1/2)^n = \sqrt{2/\pi n} \approx \frac{0.8}{\sqrt{n}} < \frac{1}{\sqrt{n}}.$$

Hence, the probability of getting the outcome is approximately 0.8/1000, less than 1/1000. Of course, this special outcome is the most likely single outcome, and it could of course occur, but a probability less than 1/1000 is quite unlikely.

a normal approximation. We now consider alternatives 2 – 4 above. We can answer the question by doing a **normal approximation**; see Section 2.7 of Ross, especially pages 79-83. A key property of the normal distribution is that it has only two parameters: its mean and variance. Thus all we should need to know are the mean and variance.

The random variable S_n is approximately normally distributed with mean $np = 500,000$ and variance $np(1-p) = 250,000$, where $p = 1/2$. 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.

When we consider case 3, we do not look at the probability $P(S_n = 501,013)$, because the probability of each value is necessarily small, as we have just seen. Instead, we want to look at $P(S_n \geq 501,013)$. We want to see the probability of getting that large value or any larger value.

In more detail we would do the following computation for case 3:

$$\begin{aligned} P(S_n \geq 501,013) &= P(S_n - E[S_n] \geq 501,013 - E[S_n]) \\ &= P\left(\frac{S_n - E[S_n]}{\sqrt{Var(S_n)}} \geq \frac{501,013 - E[S_n]}{\sqrt{Var(S_n)}}\right) \\ &\approx P\left(N(0,1) \geq \frac{501,013 - E[S_n]}{\sqrt{Var(S_n)}}\right) \\ &= P\left(N(0,1) \geq \frac{501,013 - 500,000}{500}\right) \\ &= P(N(0,1) \geq 2.006) \approx 0.0225, \end{aligned}$$

where the first two lines follow from simple arithmetic, doing the same on both sides, while the third line is the normal approximation justified by the CLT and the final numerical value is obtained from the Table 2.3 on page 82 of the class textbook. Such a large result is unlikely but not extremely unlikely.

Case 4 is worth additional discussion. An outcome of 20 standard deviations above the mean is *extremely* unlikely. However, there actually are two possible causes. On the one

hand, the report may be inaccurate. On the other hand the model may be inaccurate. Given that the model is accurate, it would be extremely unlikely that the reported outcome actually occurred. However, there is another possibility. It is possible that the model is not accurate. If the probability of heads coming up on each toss were actually around 0.51, then Case 4 would be reasonable, and Case 2 and 3 would not be reasonable. If Case 4 were the outcome of a proper experiment, we could conclude that the probability of heads must not actually be exactly 0.500 or anything less than that.

(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*

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

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

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

i.e.,

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

as $n \rightarrow \infty$ 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 $n\lambda$ can be regarded as the sum of n IID random variables, each with a Poisson distribution with mean λ (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

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

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.)

Please note that the normalization is not a significant part of the CLT statement. For any random variable Z , the associated normalized random variable $(Z - E[Z])/SD(Z)$ has mean 0 and variance 1. Since the normalized sums above have mean 0 and variance 1 for all n , there is some hope that there might be a limiting distribution, which we expect to have mean 0 and variance 1. But, for an arbitrary random variable Z , the associated normalized random variable $(Z - E[Z])/SD(Z)$ does *not* need to be normally distributed. Indeed, it is not unless Z itself is normally distributed. The amazing part of the CLT is that the distribution of the normalized sum $(S_n - E[S_n])/\sqrt{Var(S_n)}$ does approach the normal distribution as n gets large.

Moreover, 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 $n \rightarrow \infty$. It says that, as $n \rightarrow \infty$, 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

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

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

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

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.

We apply this normal approximation to approximate the distribution of S_n . As we did for case 3 above, we write

$$\begin{aligned} P(S_n \leq c) &= P(S_n - E[S_n] \leq c - E[S_n]) \\ &= P\left(\frac{S_n - E[S_n]}{\sqrt{\text{Var}(S_n)}} \leq \frac{c - E[S_n]}{\sqrt{\text{Var}(S_n)}}\right) \\ &\approx P(N(0, 1) \leq \frac{c - E[S_n]}{\sqrt{\text{Var}(S_n)}}), \end{aligned}$$

and then calculate $b = (c - E[S_n])/\sqrt{\text{Var}(S_n)}$ and look up the value $P(N(0, 1) \leq b)$ in Table 2.3 of the normal distribution (or use a program for that purpose). We can use \leq or \geq , but we have to be careful that we are consistent. We use $P(N(0, 1) \geq b) = 1 - P(N(0, 1) \leq b)$.

2. An Application of the CLT: Modeling 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

$$S_n = S_0 \times X_1 \times \cdots \times X_n , \quad (3)$$

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 $e = 2.71828\dots$). Note that

$$\log(S_n) = \log(S_0) + \log(X_1) + \cdots + \log(X_n) , \quad (4)$$

so that, by virtue of the CLT above,

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

where now (with this new interpretation of X_n)

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

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

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

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.