Lecture 11: Strong Central Limit Theorems

Bev Thurber

Department of Mathematics, MIT

March 15, 2001

1 Central Limit Theorems

Throughout these lecture notes ΔX_i are **independent** random variables, $X_n = \Sigma_{i=1}^n \Delta x_i$, $< \Delta x_i >= 0$ and $< \Delta x_i^2 >= \sigma^2 < \infty$. In this section we also assume that the variables ΔX_i are identical.

Recall from the previous lecture two central limit theorems:

1. Chebyshev: If all the moments exist, then the cumulative distribution function (CDF) $\Phi_n(\zeta)$ of $\frac{X_n}{\sigma\sqrt{n}}$ tends to a standard normal or Gaussian CDF (pointwise). A Gram-Charlier expansion for $\Phi_n(\zeta)$ can also be derived as in lecture #10. A similar expansion for the probability density function (PDF) $\phi_n(\zeta) = \Phi'_n(\zeta)$ can also be derived:

$$\phi_n(\zeta) = \phi_0(\zeta)(1 + \frac{f_1(\zeta)}{\sqrt{n}} + \frac{f_2(\zeta)}{n} + ...) \tag{1}$$

as long as sense can be made of the derivative.

2. Berry-Eséen Theorem: If the (absolute) third moment is finite, $< |\Delta x_i|^3 >< \infty$, then the CDF $\Phi_n(\zeta)$ of $\frac{X_n}{\sigma\sqrt{n}}$ obeys the following inequality:

$$|\Phi_n(\zeta) - \Phi(\zeta)| \leq \frac{3r}{\sigma^3\sqrt{n}} \tag{2}$$

This shows that the higher order terms don't matter as long as the third one is finite. Note that this theorem is much stronger than Chebyshev's because moments higher than the third need not exist. It also provides a uniform bound on the deviation from the asymptotic normal distribution which is valid for all $n \geq 1$ and $|\zeta| < \infty$.

Often probability texts present the Central Limit Theorem in this way:

Central Limit Theorem: If $< \Delta x_i >= 0$ and $< \Delta x_i^2 >= \sigma^2 < \infty$, then the CDF of $\zeta_n = \frac{X_n}{\sigma\sqrt{n}}$ tends to a standard normal CDF.

This implies that only the second moment must be finite and nonzero to have precise convergence to anormal distribution for large enough n. Finite variance is the key.

The proof of this theorem rests on the fact that $\log \hat{p}_k = \frac{-k^2\sigma^2}{2} + o(k^2)$ as k goes to 0.

2 Strong Central Limit Theorems

Let $U(x) = \int_{-x}^{x} y^2 p(y) dy$. How bad is this when the variance diverges? If $U(x) \sim L(xz)$ as $x \to \infty$, where L is slowly varying (that is, $\lim_{s\to\infty} \frac{L(sx)}{L(s)} = 1$ when x is fixed), then $\zeta_n = \frac{X_n}{f(n)}$ goes to a Gaussian. So it is really not very bad at all.

2.1 Example

Here is an example problem where the variance is barely infinite.

$$p(x) = \frac{A}{1+|x|^3} \tag{3}$$

Then $U(x) \sim \log x = L(x)$. $L(x)$ is slowly varying:

$$\lim_{s\to\infty} \frac{\log sx}{\log s} = \lim_{s\to\infty} 1 + \frac{\log x}{\log s} \to 1 \tag{4}$$

for fixed x. The convergence is not uniform, but then it doesn't have to be. Then $\zeta_n = \frac{X_n}{\sqrt{n \log n}}$ goes to a normal, but the scaling is no longer like $\sqrt{n}$. The $\log n$ term signals that this is a borderline case of anomalous diffusion. Here it is super diffusion because it goes faster than $\sqrt{n}$.

3 Non-Identical Independent Steps

In this case the steps that the random walker takes are not necessarily of the same size, for example in the case on the homework, where the step size is a^n.

Now there are m structure functions $\hat{p}_m(k)$ and m transition probabilities $p_m(k)$, one for each step. Independence of the steps implies that

$$\log \hat{P}_n(k) = \sum_{m=1}^{n} \log \hat{P}_m(k) = \sum_{m=1}^{n} (\sum_{l=2}^{\infty} \frac{(-ik)^l}{l!} c_{m,l}) \tag{5}$$

where $\hat{P}_n(k) = \hat{p_1}(k) * \hat{p_2}(k) * ... * \hat{p_n}(k)$ and $c_{m,l}$ is the lth cumulant of the mth step. The order of the summation can be switched so that it becomes

$$\sum_{l=2}^{\infty} \frac{(-ik)^l}{l!} C_{n,l} \tag{6}$$

where $C_{n,l} = \sum_{m=1}^{n} c_{m,l}$ is lth cumulant of the position after n steps. This is because cumulants add when the steps are independent.

3.1 Weakly Non-identical Steps

When the steps are weakly non-identical, it turns out that the cumulants of the earlier analysis can be replaced with the average cumulants ($\bar{c}_l$). The steps are weakly non-identical when the average cumulant exists, that is when the limit of equation 7 exists.

$$\bar{c}_l = \lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} c_{m,l} < \infty \tag{7}$$

If $\bar{c}_l$ is positive and finite, then $C_{n,l} \sim n\bar{c}_l$ as n goes to ∞. The variations tend to settle down for a large number of steps.

Remember that $log\hat{P}_n(k) = \frac{-k^2\bar{c}_2 n}{2} + \frac{ik^3\bar{c}_3 n}{3!} + \frac{k^4\bar{c}_4 n}{4!} - \ldots$ when the cumulants are replaced with the average cumulants. Then use the change of variables $\zeta_n = \frac{x}{\bar{\sigma}\sqrt{n}}$ and $\omega_n = k\bar{\sigma}\sqrt{n}$. Then

$$\log\hat{\phi}_n(\omega) \sim \frac{-\omega^2}{2} + \frac{i\bar{\lambda}_3}{3!}\frac{\omega^3}{\sqrt{n}} + \frac{\bar{\lambda}_4}{4!}\frac{\omega^4}{n} + \ldots \tag{8}$$

where $\bar{\lambda}_l = \frac{\bar{c}_l}{\bar{\sigma}_l}$. Here $\bar{\sigma}$ sets the scale. Don't forget that $\bar{c}_2 = \bar{\sigma}^2$ is the average variance. This equation is the same as one for the earlier analysis, but with c_l replaced by $\bar{c}_l$. In this case the Gaussian appears and $< x_n^2 >^{\frac{1}{2}} \sim \sqrt{n}$ if $\log\hat{p}_n(k) \sim \frac{-k^2}{2}\sum_{m=1}^{n} c_{m,2} + o(k^2)$, where $c_{m,2}$ is finite.

3.2 Berry-Eséen Theorem for (Weakly) Non-identical Variables

Given $< \Delta x_m >= 0$, $< \Delta x_m^2 >= \sigma_m^2 < \infty$, $< |\Delta x_m|^3 >= r_m < \infty$, the CDF of $\zeta_n = X_n/\sqrt{\sum_{m=1}^{n}\sigma_m^2}$ converges to a standard normal CDF, and furthermore satisfies the following uniform bound:

$$|\Phi_n(\zeta) - \Phi(\zeta)| \le \frac{6\sum_{m=1}^{n} r_m}{(\sum_{m=1}^{n}\sigma_m^2)^{\frac{3}{2}}} \sim \frac{6\bar{r}}{\bar{\sigma}^{\frac{3}{2}}\sqrt{n}} \tag{9}$$

Here, once again, the limit is the same as for the non-identical variable case, but with the variances replaced with their average. The extra fluctuations due to the non-identical variables turned the "3" of the first Berry-Eséen theorem simply turned into a "6'.

3.3 Central Limit Theorem for the Non-identical Variables

If $\bar{\sigma}^2 < \infty$, then $\frac{X_n}{\bar{\sigma}\sqrt{n}}$ goes to a normal distribution.

But what if $\bar{\sigma}^2$ isv not finite, or something else goes wrong? When does the central limit theorem hold?

Keep ζ defined as it was in section 3.2, and define $\omega_n = kC_{n,2}^{\frac{1}{2}}$. Then $\zeta = \frac{x}{C_{n,2}^{\frac{1}{2}}}$. Now

$$\log\hat{\phi}_n(\omega) = \frac{-\omega^2}{2} + \frac{i\omega^3}{3!}\frac{c_{n,3}}{c_{n,2}^{\frac{3}{2}}} + \ldots \tag{10}$$

so the central limit theorem applies if $\frac{c_{n,l}}{c_{n,2}^{\frac{l}{2}}}$ goes to zero for $l \ge 3$.

3.4 Examples

1. Problem set 2 problem $c_{m,l} = a^{ml} c_l$: geometrically decaying steps. Here the higher terms don't go to zero, so it is not a Gaussian. When $0 < a < 1$, exact solutions for $\phi_\infty(\zeta)$ are possible for $a = \frac{1}{2^{\frac{1}{m}}}$, with m an integer. The case of geomtrically growing steps $a > 0$ can also be addresses.

2. An example for next time $c_{m,l} = m^{\alpha l} c_l$: Here the scale for the mth step is m^α. What happens for various values of α? When α is 0, the central limit theorem applies, with $\sqrt{n}$ scaling. But as α is reduced (negative), the $\sqrt{n}$ is lost, at first with a Gaussian distribution, and eventually even the Gaussian gets lost. When α is positive, the Gaussian shape holds, but with different scaling.

4 Additional Reading

Two books that might be interesting to read are Feller's *Introduction to Probability* (especially volume 2) and Gnedenko and Kolmorgov's *Limit Theorems for Sums of Random Variables.*