Berry-Esseen Bounds

Mladen Savov

August 2018

1 Introduction

The central limit theorem (CLT) plays a fundamental role in classical probability and modern statistics. Recall that if $(\xi_j)_{j=1}^{\infty}$ is a sequence of independent, equally distributed random variables with $\mathbf{E}[\xi_1] = 0$ and

$$\infty > Var(\xi_1) = \mathbf{E}[\xi_1^2] = \sigma^2 > 0$$

then with

$$S_n = \sum_{j=1}^{n} \xi_j$$

we have that

$$\lim_{n\to\infty} \frac{S_n}{\sigma\sqrt{n}} \stackrel{d}{=} Z \sim N(0,1), \tag{1}$$

where Z is a standard normal distribution and convergence in distribution ***in this case, but generally not***, is equivalent to

$$\lim_{n\to\infty} \mathbf{P}(S_n \leq x\sigma\sqrt{n}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{y^2}{2}} dy = \mathbf{P}(Z \leq x), \tag{2}$$

for any real x.

However, (2) would hardly be of any practical use, unless we could evaluate/estimate/bound the speed of convergence

$$Err(n,x) := \left|\mathbf{P}(S_n \leq x\sigma\sqrt{n}) - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{y^2}{2}} dy\right|.$$

Imagine that $Err(n,x)$ was of order of magnitude of $\ln\ln(n)$. How large a sample one would need to ensure some proximity of the two probabilities? This raises the question why is it a rule of thumb in statistics to approximate with normal law Z the random variable $\frac{S_n}{\sigma\sqrt{n}}$ provided $n \geq 30$?

2 Berry-Esseen bounds

The Berry-Esseen bounds give general rates of the speed of convergence provided $\mathbf{E}[|\xi_1^3|] = \rho < \infty$, where we keep the notation of the previous section.

2.1 Uniform Berry-Esseen bounds

It can be established[1] that

$$\begin{aligned} \sup_{x\in\mathbf{R}} Err(n,x) &= \sup_{x\in\mathbf{R}} \left|\mathbf{P}(S_n \le x\sigma\sqrt{n}) - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{y^2}{2}}\,dy\right| \\ &\le 0.4748\frac{\rho}{\sigma^3\sqrt{n}}. \end{aligned} \tag{3}$$

We see that ***uniformly*** the speed of convergence is of $n^{-1/2}$ and depends on the ratio ρ/σ^3. From the Holder's inequality it can be verified that

$$\rho \ge \sigma^3$$

and therefore the ratio cannot be smaller than 1.

Let us consider $n = 30$. Then the best uniform bound we can get

$$\begin{aligned} \sup_{x\in\mathbf{R}} Err(30,x) &= \sup_{x\in\mathbf{R}} \left|\mathbf{P}(S_{30} \le x\sigma\sqrt{30}) - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{y^2}{2}}\,dy\right| \\ &\le 0.086\frac{\rho}{\sigma^3}, \end{aligned} \tag{4}$$

which seems satisfactory in a rather crude manner as longs as for some x

$$0.086 \le 0.086\frac{\rho}{\sigma^3} << \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{y^2}{2}}\,dy. \tag{5}$$

Things will badly fail when x is very small since

$$\lim_{x\to-\infty} \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{y^2}{2}}\,dy = 0$$

and we require that it exceeds 0.086 in (5).

Is the uniform estimate (3) too much to ask for when x is very small? Can (3) not depend on x?

2.2 Non-Uniform Berry-Esseen bounds

It can be established[2] that

$$\begin{aligned} Err(n,x) &= \left|\mathbf{P}(S_n \le x\sigma\sqrt{n}) - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{y^2}{2}}\,dy\right| \\ &\le 23\frac{\rho}{\sigma^3\sqrt{n}(1+|x|^3)}. \end{aligned} \tag{6}$$

These are non-uniform bounds as the dependence on x is clear. Then what would happen if $n = 30$?

$$\begin{aligned} Err(30,x) &= \left|\mathbf{P}(S_{30} \le x\sigma\sqrt{30}) - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{y^2}{2}}\,dy\right| \\ &\le 23\frac{\rho}{\sqrt{30}\sigma^3(1+|x|^3)}, \end{aligned} \tag{7}$$

[1] https://en.wikipedia.org/wiki/Berry–Esseen_theorem

[2] https://arxiv.org/pdf/1301.2828.pdf

and this to be a reasonable estimate it suffices that

$$\frac{23}{\sqrt{30}} \frac{\rho}{\sigma^3(1+|x|^3)} << \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{y^2}{2}} dy. \tag{8}$$

However, as x goes to $-\infty$

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{y^2}{2}} dy$$

is of order $1/|x|e^{-x^2/2}$ and the right-hand side in (8) decays faster to 0 than the left-hand side. Therefore, no much benefit is on offer when x is too small. Since in general (8) depends on n we have that

$$\frac{23}{\sqrt{n}} \frac{\rho}{\sigma^3(1+|x|^3)} << \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{y^2}{2}} dy \sim \frac{1}{\sqrt{2\pi}|x|} e^{-\frac{x^2}{2}}, \tag{9}$$

as $x \to -\infty$. We can see what order of magnitude we need to ensure so that (9) can be satisfied. It is rather unwieldy. However, when x is large then the lower bound is improved and the approximation is excellent.

3 Conclusions

- The recipe $n \geq 30$ for approximation of $S_n/\sigma\sqrt{n}$ with normal law is mostly a rule of thumb indeed. Perhaps it works well for specific rather symmetric distributions but cannot ensure theoretical error below 0.086

- The non-uniform Berry-Esseen bounds are of not significant help when we need to approximate

  $$q(x) := \mathbf{P}(S_n \leq x\sigma\sqrt{n})$$

  and $q(x)$ is of order less than

  $$\frac{23}{\sqrt{n}} \frac{\rho}{\sigma^3(1+|x|^3)}$$

  which happens for x not very far from zero.

- The Berry-Esseen bounds (both uniform and non-uniform) ensure good estimates of

  $$q(x) = \mathbf{P}(S_n \leq x\sigma\sqrt{n})$$

  provided $q(x)$ relatively large. The theoretical error never exceeds

  $$\frac{23}{\sqrt{n}} \frac{\rho}{\sigma^3(1+|x|^3)}$$

  which is rather small for n and x large.