

# Lecture 5: Edgeworth Expansion and Models

Ma Xuejun

School of Mathematical Sciences

Soochow University

<https://xuejunma.github.io>



# Outline

- Review
- The Delta Method
- The Edgeworth Expansion

## Terminology

- The Delta Method; The multivariate Delta method
- The Edgeworth Expansion; the expression of the expansion, including the notations related to the expression.



# Review

- The Weak Law of Large Numbers: converge in probability/distribution
- The Strong Law of Large Numbers: almost sure convergence
- Use of LLN: Monte Carlo method
- The Central Limit Theorem

# The Delta Method

- CLT is for  $\bar{X}_n$
- Can we generalize it to  $g(\bar{X})$ ? So the application of CLT is extended.
- Example: what is the limiting dist. for  $e^{\bar{X}_n}$  ?

## Theorem 5.5.24: The Delta Method

Suppose that  $\frac{\sqrt{n}(Y_n - \mu)}{\sigma} \xrightarrow{d} N(0, 1)$ , and that  $g(\cdot)$  is a **differentiable** function such that  $g'(\mu) \neq 0$ , then

$$\frac{\sqrt{n}(g(Y_n) - g(\mu))}{|g'(\mu)|\sigma} \xrightarrow{d} N(0, 1)$$

Remark.

- $Y_n = \bar{X}_n$  for the CLT. Here,  $Y_n$  is a generalized case. For any  $Y_n$  satisfying the convergence rule, we have the delta method.
- In other words,  $Y_n \xrightarrow{d} N(\mu, \sigma^2/n)$  implies that  $g(Y_n) \xrightarrow{d} N(g(\mu), (g'(\mu))^2 \sigma^2/n)$

# Proof: The Delta Method

Intuition of proof: We are interested in the term  $\frac{\sqrt{n}(g(Y_n) - g(\mu))}{|g'(\mu)|\sigma}$ . Since  $\frac{\sqrt{n}(Y_n - \mu)}{\sigma}$  converges to normal distribution, we have that  $Y_n - \mu = O_p(n^{-0.5})$ . Since  $g$  is differentiable at  $\mu$ , when  $y - \mu = o(1)$ , there is

$$g(y) = g(\mu) + g'(\mu)(y - \mu) + o(1)(y - \mu)$$

Introduce it into the term of interest,

$$\begin{aligned}\frac{\sqrt{n}(g(Y_n) - g(\mu))}{|g'(\mu)|\sigma} &= \frac{\sqrt{n}(g(\mu) + g'(\mu)(Y_n - \mu) + rem - g(\mu))}{|g'(\mu)|\sigma} \\ &= \frac{\sqrt{n}(g'(\mu)(Y_n - \mu) + rem)}{|g'(\mu)|\sigma} \\ &= \sqrt{n} \left[ \frac{(Y_n - \mu)}{\sigma} \right] + \sqrt{n}|g'(\mu)|\sigma \times rem\end{aligned}$$

The first term converges to  $N(0, 1)$  in distribution, and the second term converges to 0. So the summation would converge to  $N(0, 1)$  in distribution. Note. The

point here is that  $g(y) = g(\mu) + g'(\mu)(y - \mu) + o(1)(y - \mu)$  when  $y$  changes with a small quantity. Since  $\delta$  is always used to denote small quantity, so the method is called "The Delta Method".

## The Multivariate Delta Method

Suppose that  $Y_n = (Y_{n1}, Y_{n2}, \dots, Y_{nk})$  is a sequence of random vectors such that  $\sqrt{n}(Y_n - \mu) \xrightarrow{d} N(0, \Sigma)$ . Let  $g : R^k \rightarrow R$  and let

$$\begin{pmatrix} \frac{\partial g}{\partial y_1} \\ \vdots \\ \frac{\partial g}{\partial y_k} \end{pmatrix}$$

Let  $\nabla_\mu$  denote  $\nabla g(y)$  evaluated at  $y = \mu$  and assume that the elements of  $\nabla_\mu$  are nonzero. Then

$$\sqrt{n}(g(Y_n) - g(\mu)) \xrightarrow{d} N(0, \nabla_\mu^\top \Sigma \nabla_\mu)$$

- The generalisation of multivariate CLT.

## Example 1: The Delta Method

Let  $X_1, \dots, X_n$  be i.i.d with finite mean  $\mu$  and finite variance  $\sigma^2$ . By the CLT,

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{d} N(0, 1).$$

Let  $W_n = e^{\bar{X}_n}$ . Thus,  $W_n = g(\bar{X}_n)$  where  $g(t) = e^t$ . Since  $g'(t) = e^t$ , the delta method implies that

$$\frac{\sqrt{n}(W_n - e^\mu)}{\sigma e^\mu} \xrightarrow{d} N(0, 1)$$

## Example 2: The Delta Method

Let  $X_1, \dots, X_n$  be i.i.d with finite mean  $\mu$  and finite covariance matrix  $\Sigma$ . Then according to the multivariate CLT, we have

$$\sqrt{n} \left( \begin{bmatrix} \bar{X}_{n1} \\ \bar{X}_{n2} \end{bmatrix} - \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \right) \xrightarrow{d} N(0, \Sigma)$$

Let  $g(s, t) = s^2 + t^2$ , then can apply the delta method and we have that

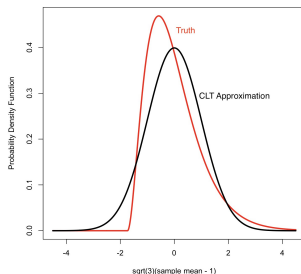
$$\sqrt{n}(\bar{X}_{n1}^2 + \bar{X}_{n2}^2 - \mu_1^2 - \mu_2^2) \xrightarrow{d} 2\mu_1 Z_1 + 2\mu_2 Z_2$$

where  $Z = (Z_1, Z_2)^\top \sim N(0, \Sigma)$ .



# The Edgeworth Expansion

- Note that CLT is a good estimate for **large  $n$**
- When  $n$  is small, it might be away from the truth
- Example: Consider  $X_i \stackrel{i.i.d.}{\sim} \text{Gamma}(1, 1), i = 1, 2, 3$ , We know that  $\mathbf{E}X_1 = \mathbf{Var}X_1 = 1$ . Let  $Z = \sqrt{3}(\bar{X}_3 - 1)$ , then according to CLT,  $Z \sim N(0, 1)$  approximately.



- Since  $n$  is small, there is large difference! How to get a good estimate?

# The Edgeworth Expansion

Recall that in the proof of CLT, we figured out the characteristic function as

$$\phi_{Z_n}(t) = \phi_Y^n\left(\frac{t}{\sqrt{n}}\right)$$

where  $Z_n = \frac{1}{\sigma}\sqrt{n}\left[\frac{1}{n}\sum_{i=1}^n(X_i - \mathbf{E}X_1)\right]$  and  $Y = (X_1 - EX_1)/\sigma$ . When we take it as  $\phi_Y(t/\sqrt{n}) = 1 - \sigma^2 t^2 / (2\sigma^2 n) + o(1/n)$ , then we have CLT.

What if we consider higher order moments?

- Let  $\kappa_j$  be  $j$ -th cumulant of  $Y$ , where

$$\kappa_1 = \mathbf{E}Y, \quad \kappa_2 = \mathbf{Var}(Y), \quad \kappa_3 = \mathbf{E}(Y - \mathbf{E}Y)^3, \quad \kappa_4 = \mathbf{E}(Y - \mathbf{E}Y)^4 - 3\mathbf{Var}^2 Y$$

- Consider  $K_Y = \log(\phi_Y(t))$  and the Taylor series expansion of  $K_Y(t)$  at  $t = 0$ ,

$$K_Y(t) = \frac{1}{\phi_Y(0)} \frac{d\phi_Y(t)}{dt} \Big|_{t=0} + \frac{1}{\phi_Y(0)} \frac{d^2\phi_Y(t)}{dt^2} \Big|_{t=0} + \cdots + = \sum_{j=1}^{\infty} \kappa_j \frac{(it)^j}{j!}$$

- Since  $Y = (X_1 - \mathbf{E}X_1)/\sigma$ , we have  $\kappa_1 = \mathbf{E}Y = 0$ ,  $\kappa_2 = \mathbf{Var}Y = 1$ , so

$$\phi_Y(t) = \exp\left(-\frac{t^2}{2} + \sum_{j=3}^{\infty} \frac{(it)^j}{j!}\right) = \exp\left(-\frac{t^2}{2} + \frac{\kappa_3^3(it)}{3!} + \dots + \frac{\kappa_j^j}{j!} + \dots\right)$$

- Introduce it into the equation for

$$\phi_{Z_n}(t) = e^{-t^2/2} \left[ 1 + n^{-1/2} r_1(it) + n^{-1} r_2(it) + o(n^{-1}) \right]$$

where  $r_1(it) = \kappa_3^3(it)/6$ , and  $r_2(it) = \frac{1}{24}\kappa_4^4(it) + \frac{1}{72}\kappa_3^{12}(it)$

- Here, we apply the higher order moments of  $X$  to depict the characteristic function more clearly

Therefore, we have the Edgeworth expansion:

$$F_{Z_n}(z) = \Phi(z) + \frac{1}{\sqrt{n}}p_1(z)\phi(z) + \frac{1}{n}p_2(z)\phi(z) + o\left(\frac{1}{n}\right)$$

where  $p_1(z) = -\kappa_3(z^2 - 1)/6$  and

$$p_2(z) = -z \left[ \kappa_4(z^2 - 3)/24 + \kappa_3^2(z^4 - 10z^2 + 15)/72 \right].$$

Returning to our example. Consider the first order Edgeworth expansion, we have

$$P(Z_e \leq z) \approx \Phi(z) + \frac{1}{\sqrt{3}}\phi(z)p_1(z)$$

where  $p_1(z) = -\kappa_3(z^2 - 1)/6$  and  $\kappa_3 = 2$ .

