Lecture 5: Edgeworth Expansion and Models

Xuejun Ma

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 adifferentiable 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 generlized 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 μ,when $y - \mu = 0(1)$,there is

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

Introduce it into the term of interest,

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

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 $\boldsymbol{Y}_n = (Y_{n1}, Y_{n2}, \cdots, Y_{nk})$ is a sequence of random vectors such that $\sqrt{n}(\boldsymbol{Y}_n - \boldsymbol{\mu}) \xrightarrow{d} N(\boldsymbol{0}, \boldsymbol{\Sigma})$.Let $g : R^k \to R$ and let $g : \mathcal{R}^k \to \mathcal{R}^m$ be once diőerentiable at $\boldsymbol{\mu}$ with the gradient matrix $\Delta_g(\boldsymbol{\mu})$, then

$$\sqrt{n}(g(\boldsymbol{Y}_n) - g(\boldsymbol{\mu})) \xrightarrow{d} N_m(\boldsymbol{0}, \Delta_g^\top(\boldsymbol{\mu})\boldsymbol{\Sigma}\Delta_g(\boldsymbol{\mu}))$$

provided $\Delta_g^\top(\boldsymbol{\mu})\boldsymbol{\Sigma}\Delta_g(\boldsymbol{\mu})$ is positive definite.

Example 1: The Delta Method

Let $X_1, \cdots, X_n$ be i.i.d with finite mean μ and finite variance σ^2.By the CLT,

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

. Let $W_n = e^{\hat{X}_n}$.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 $X_1, \cdots, X_n$ be i.i.d with finite mean μ and finite covariance matrix Σ.The according to the multivariate CLT, we have

$$\sqrt{n}\Big(\begin{bmatrix} \bar{X}_{n1} \\ \bar{X}_{n2} \end{bmatrix} - \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \Big) \xrightarrow{d} N_2(\mathbf{0}, \mathbf{\Sigma}$$

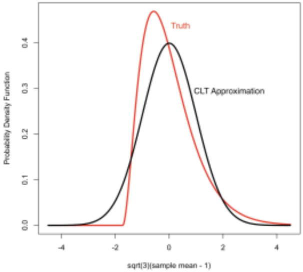
Let $g(, t) = s^2 + t^2$,then can apply the delta method and we have that

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

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

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 $X_i \widetilde{i.i.d} Gamma(1,1), i = 1,2,3$,We know that $EX_1 = VarX_1 = 1$.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

$$\Phi_{Z_n}(t) = \Phi_Y^n(\frac{t}{\sqrt{n}})$$

where $Z_n = \frac{1}{\sigma}\sqrt{n}[\frac{1}{n}\sum_{i=1}^{n}(X_i - EX_i)]$ 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 k_j be jth cumulant of Y,where

$$k_1 = EY, k_2 = Var(Y), k_3 = E(Y_E Y)^3, k_4 = E(Y - EY)^4 - 3VarY$$

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

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

- Since $Y = (X_1 - EX_1)/\sigma$,we have $k_1 = EY = 0, k_2 = VarY = 1$,so

$$\Phi_Y(t) = exp(-\frac{t^2}{2} + \sum_{j=3}^{\infty} \frac{(it)^j}{j!}) = exp(-\frac{t^2}{2} + \frac{k_3^3(it)}{3!} + \cdots + \frac{k_j^j}{j!} + \cdots)$$

- Introduce it into the equation for

$$\Phi_{Z_n}(t) = e^{-t^2/2}[1 + n^{-1/2} r_i(it) + n^{-1} r_2(it) + o(n^{-1})]$$

where $r_1(it) = k_3^3(it)/6$,and $r_2(it) = \frac{1}{24} k_4^4(it) + \frac{1}{72} k_3^{12}(it)$

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

Therefore, we have the Edgeworth expansion:

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

where $p_1(z) = -k_3(z^2 - 1)/6$ and

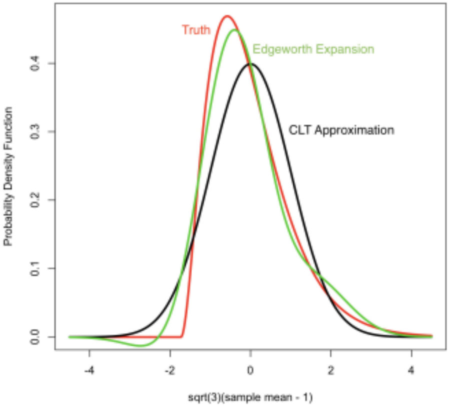
$$p_2(z) = -z[k_4(z^2 - 3)/24 + k_3^2(z^4 - 10z^3 + 15)/72]$$

.

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

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

where $p_1(z) = -k_3(z^2 - 1)/6$ and $k_3 = 2$.



Homework

1. Let $X_1, \cdots, X_n, \cdots$ be a sequence of independent random variables such that
$$E[X_i] = \mu, \qquad Var(X_i) < \sigma^2, \qquad n = 1, 2, \cdots$$
With Chebychev's inequality,prove that $\bar{X}_n \xrightarrow{P} \mu$.
2. Let $U_1, U_2, \cdots$ be independent random variables having the uniform distribution on [0,1] and $Y_n = (\prod_{i=1}^{n} U_i)^{-1/n})$. Show that $\sqrt{n}(Y_n - e) \xrightarrow{d} N(0, e^2)$.
3. Let $X_1, \cdots, X_n$ be i.i.d. random variables following Uniform[0,1].Let $Y_n = \min(X_1, \ldots, X_n)$.
(i) Show that $Y_n \xrightarrow{a.s} 0$ as $n \to \infty$.
(ii) Show that $nY_n \xrightarrow{d} \exp(1)$,where Exp(1) is the exponential distribution with density $f(x) = e^{-x}$ for $x > 0$.