1 Asymptotic Normality

1.1 Delta Method

Definition 1. The functional f is called *Hadamard differentiable* at F if there exists a linear functional $L_F : \mathcal{D} \to \mathbb{R}^k$ such that for any $\varepsilon_n \to 0$ and $\{D, D_1, ...\} \subset \mathcal{D} : \sup_x |D_n(x) - D(x)| \to 0, n \to \infty$, $F_n + \varepsilon n D_n \in \mathcal{F}$

$$\lim_{n\to\infty} \left(\frac{f(F + \varepsilon_n D_n) - f(F)}{\varepsilon_n} - L_F(D_n) \right) = 0$$

We'll consider $\mathcal{D} = \{F - G,\ F, G \in \mathcal{F}\}$.

Definition 2. *The influence function* of functional f is defined by

$$I_F(x) := L_F(\delta_x - F) = \lim_{\varepsilon\to 0} \frac{f((1-\varepsilon)F + \varepsilon\delta_x) - f(F)}{\varepsilon},$$

where $\delta_x(u) = I_{u\geq x}$ is the c.d.f. of x.

Theorem 1. *Let f be Hadamard differentiable functional on $\mathcal{F}$. Then $f(\widehat{F_n})$ is an asymptotical normal estimator for $f(F)$ with $\sigma^2(F) = \int_\infty^\infty I_F(x)^2 dF(x)$.*

We'll prove the theorem later.

Lemm 1. *Particularly, let $g : \mathbb{R}^n \to \mathbb{R}$ be a smooth function, f_1,...,f_n be Hadamard differentiable functionals. Then*

$$L_F(D; g(f_1, ..., f_n)) = \left.\frac{dg(f_1(F + xD), ..., f_n(F + xD))}{dx}\right|_{x=0} = \sum_{i=1}^{n} \left.\frac{dg}{du_i}\right|_{\vec{u}=(f_1(F),...,f_n(F))} L_F(D; f_i). \quad (1)$$

Доказательство. By definition,

$$L_F(D; f) = \left.\frac{df(F + xD)}{dx}\right|_{x=0}.$$

Then (1) is a consequence of chain rule. □

2 Some applications

2.1 Covariance estimation

Now consider some applications:

Example 1. Let (X_i, Y_i) be i.i.d. random vectors with d.f. $F(x, y)$ and let

$$f(F) = \operatorname{cov}(X, Y) = \int_{\mathbb{R}^2} xy dF(x, y) - \int_{\mathbb{R}^2} x dF(x, y) \int_{\mathbb{R}^2} y dF(x, y),$$

where $\mathcal{F} = \{F(x, y) : \int_{\mathbb{R}^2} x^2 dF(x, y) < \infty, \int_{\mathbb{R}^2} y^2 dF(x, y) < \infty\}$. Consider $f_1(F) = \mathbf{E}XY$. Then

$$\frac{f_1(F + \varepsilon_n D_n) - f_1(F)}{\varepsilon_n} = L_F(D_n),$$

where

$$L_F(D_n) = \int_{\mathbb{R}^2} uv dD_n(u, v).$$

Therefore, $f(F)$ is Hadamard differentiable with

$$I_F(x, y; f_1) = L_F(\delta_{x,y} - F) = \int_{\mathbb{R}^2} uv d(\delta_{x,y}(u, v) - F(u, v)) = xy - \int_{\mathbb{R}} uv dF(u, v).$$

Similarly,

$$I_F(x,y;f_2) = x - \int_{\mathbb{R}^2} u dF(u,v), \quad I_F(x,y;f_3) = y - \int_{\mathbb{R}^2} v dF(u,v),$$

where $f_2(F) = \mathbf{E}X$, $f_3(F) = \mathbf{E}Y$. Let $g(x,y,z) = x - yz$. Then due to (1) $f(F) = \mathrm{cov}(X,Y) = g(f_1(F), f_2(F), f_3(F))$ is Hadamard differentiable functional with

$$I_F(x;f) = L_F(\delta_x, f_1) - f_3(F)L_F(\delta_x, f_2) - f_2(F)L_F(\delta_x, f_3) = xy - \mathbf{E}_F XY - (x - \mathbf{E}_F X)\mathbf{E}_F Y - (y - \mathbf{E}_F Y)\mathbf{E}_F X = (x - \mathbf{E}_F X)(y - \mathbf{E}_F Y) - \mathrm{cov}(X,Y).$$

Therefore, $f(\widehat{F}_n) = \overline{XY} - \overline{X}\,\overline{Y}$ is asymptotically normal estimator for $\mathrm{cov}(X,Y)$ with

$$\sigma^2(F) = \mathbf{E}_F(x - \mathbf{E}_F X)^2(y - \mathbf{E}_F Y)^2 - (\mathrm{cov}(X,Y))^2.$$

Example 2. Let (X_i, Y_i) be i.i.d. random vectors with d.f. $F(x,y)$ and let

$$f(F) = \frac{\mathrm{cov}(X,Y)}{\sqrt{\mathbf{D}_F X \mathbf{D}_F Y}}.$$

Denote

$$f_1(F) = \mathrm{cov}(X,Y), \quad f_2(F) = \mathbf{D}_F(X), \quad f_3(F) = \mathbf{D}_F(Y).$$

Then

$$L_F(D;f) = \frac{L_F(D;f_1)}{\sqrt{\mathbf{D}_F X \mathbf{D}_F Y}} - \frac{f_1(F)L_F(D;f_2)}{2\sqrt{(\mathbf{D}_F X)^3 \mathbf{D}_F Y}} + \frac{f_1(F)L_F(D;f_3)}{2\sqrt{(\mathbf{D}_F Y)^3 \mathbf{D}_F X}}.$$

Here

$$I_F(\delta_x;f_1) = (x - \mathbf{E}_F X)(y - \mathbf{E}_F Y) - \mathrm{cov}(X,Y), \quad I_F(\delta_x;f_2) = (x - \mathbf{E}_F X)^2 - \mathbf{D}_F X, \quad I_F(\delta_x;f_3) = (y - \mathbf{E}_F Y)^2 - \mathbf{D}_F Y.$$

Therefore,

$$I_F(x;f) = \widetilde{x}\widetilde{y} - \frac{1}{2}\mathrm{corr}(X,Y)(\widetilde{x}^2 + \widetilde{y}^2),$$

where

$$\widetilde{x} = \frac{x - \mathbf{E}_F X}{\sqrt{D_F X}}, \quad \widetilde{y} = \frac{y - \mathbf{E}_F Y}{\sqrt{D_F Y}}.$$

Sample quantile

Example 3. Suppose that $\mathcal{F}$ is a set of absolutely continious distributions with p.d.f. $p(x)$, such that $p(x_\alpha) > 0$, $p(x)$ is continuous in some neighborhood of x_α, where x_α — α-quantile of F. Consider $f(F) = F^{-1}(\alpha)$, $f(F) = x_\alpha$ as $F \in \mathcal{F}$. Let's prove that the sample quantile

$$f(\widehat{F}_n) = \widehat{F}_n^{-1}(\alpha) = X_{([\alpha n])}$$

is an asymptotical normal estimator for x_α. Really,

$$F_{\varepsilon,x}(x_\alpha) := (1-\varepsilon)F(u) + \varepsilon\delta_x(u) = \begin{cases} (1-\varepsilon)\alpha + \varepsilon, & x_\alpha > x, \\ (1-\varepsilon)\alpha, & x_\alpha \le x. \end{cases}$$

Then for every x and ε small enough

$$F_{\varepsilon,x}^{-1}(\alpha) = \begin{cases} F^{-1}\left(\frac{\alpha}{1-\varepsilon}\right), & x \ge x_\alpha, \\ F^{-1}\left(\frac{\alpha-\varepsilon}{1-\varepsilon}\right), & x < x_\alpha. \end{cases}$$

So

$$L_F(x) = \lim_{\varepsilon\to 0}\frac{F_{\varepsilon,x}^{-1}(\alpha) - F^{-1}(\alpha)}{2\varepsilon} = \begin{cases} \lim_{\varepsilon\to 0}\frac{F^{-1}\left(\frac{\alpha}{1-\varepsilon}\right) - F^{-1}(\alpha)}{\varepsilon}, & x \ge x_{1/2}, \\ \lim_{\varepsilon\to 0}\frac{F^{-1}\left(\frac{\alpha-\varepsilon}{1-\varepsilon}\right) - F^{-1}(\alpha)}{\varepsilon}, & x < x_{1/2}, \\ 0, & x = x_{1/2}, \end{cases}$$

We have

$$\lim_{\varepsilon\to 0}\frac{F^{-1}\left(\frac{\alpha}{1-\varepsilon}\right)-F^{-1}(\alpha)}{\varepsilon}=\left(F^{-1}\left(\frac{\alpha}{1-u}\right)\right)'\Bigg|_{u=0}=\frac{1}{F'\left(F^{-1}\left(\frac{\alpha}{1-u}\right)\right)}\frac{\alpha}{(1-u)^2}\Bigg|_{u=0}=\frac{\alpha}{p(x_\alpha)},$$

where p is the probability density function of F. Similarly,

$$\lim_{\varepsilon\to 0}\frac{F^{-1}\left(\frac{\alpha-\varepsilon}{1-\varepsilon}\right)-F^{-1}(1/2)}{\varepsilon}=\frac{\alpha-1}{p(x_\alpha)}.$$

Therefore,

$$\sigma^2(F)=\int_{x>x_\alpha}\frac{\alpha^2}{p(x_\alpha)^2}dF(x)+\int_{x<x_\alpha}\frac{(\alpha-1)^2}{p(x_\alpha)^2}dF(x)=\frac{(1-\alpha)\alpha}{p(x_\alpha)^2}.$$

So, the sample quantile is an asymptotically normal estimator of $x_{1/2}$ with asymptotic variance $\sigma^2(F)$. The proof of Hadamard differentiability can be found in the Appendix.