Lection 7. Symmetrical distributions

Symmetrical Distributions

Let's consider other estimators for θ in our model.

It follows from the symmetry that

$$\theta = f(F) = \frac{1}{1-2\alpha}\int_{\alpha}^{1-\alpha} F^{-1}(x)dx$$

for every $\alpha \in [0, 1/2)$. Therefore, it's natural to use the estimator

$$\widehat{\theta} = \frac{1}{1-2\alpha}\int_{\alpha}^{1-\alpha} \hat{F}_n^{-1}(x)dx = \frac{X_{([\alpha n])}([n\alpha]+1-n\alpha) + X_{([\alpha n]+1)} + ... + X_{(n-[\alpha n])} + X_{(n-[\alpha n]+1)}(n\alpha - [n\alpha])}{(1-2\alpha)n}$$

Definition 1. The estimator

$$\overline{X}_{\alpha} = \frac{X_{([\alpha n]+1)} + ... + X_{(n-[\alpha n])}}{n - 2[\alpha n]},$$

is called the α-*trimmed mean.*

The α-trimmed mean is similar to $\hat{\theta}$ above and has the same asymptotical distribution.

Example 1. Let's find the asymptotic breakdown point for $\hat{\theta}$. Obviously, $b(1) = \infty$. Moreover, $b(\alpha+\varepsilon) = \infty$ for every $\varepsilon > 0$ since

$$\frac{1}{1-2\alpha}\int_{\alpha}^{1-\alpha} F_n^{-1}(x)dx = \frac{1}{1-2\alpha}\int_{\alpha+\varepsilon}^{1-\alpha} F^{-1}(x)dx - n\varepsilon \to \infty,\ n \to \infty,$$

for $F_n(x) = F_0(x)I_{x \geq F_0^{-1}(\alpha+\varepsilon)} + (\alpha+\varepsilon)I_{x\in[-n, F_0^{-1}(\alpha+\varepsilon))}$. On the other hand, $b(\alpha - \varepsilon) < \infty$, since $\forall\varepsilon > 0$ we have

$$F^{-1}(x) \geq F_0^{-1}(x - (\alpha - \varepsilon)) - \alpha + \varepsilon \geq F_0^{-1}(\varepsilon) - 1 > -\infty$$

as $x \in [\alpha, 1-\alpha]$. Therefore, $\varepsilon^* = \alpha$.

For any $\alpha > 0$ this estimator is robust. On the other hand, it's an asymptotically normal estimator in $\mathcal{F} = \mathcal{F}_{symm}$

Theorem 1. *As* $n \to \infty$

$$\sqrt{n}(f(\widehat{F}_n) - f(F)) \xrightarrow{d} Z \sim \mathcal{N}(0, \sigma_\alpha^2),$$

where

$$\sigma_\alpha^2 = \frac{2}{(1-2\alpha)^2}\left(\int_{\theta}^{x_{1-\alpha}} (t-\theta)^2 f(t)dt + \alpha(x_{1-\alpha} - \theta)^2\right).$$

Доказательство. Let

$$F_n(u) = F(u) + \varepsilon D(u).$$

Then

$$F_n^{-1}(x) = F^{-1}\left(-\varepsilon D(F_n^{-1}(x))\right) = F^{-1}(x) - (F^{-1}(x))'\varepsilon D(F_n^{-1}(x)) + o(\varepsilon).$$

Therefore

$$\frac{F_n^{-1}(x) - F^{-1}(x)}{\varepsilon} = -(F^{-1}(x))'(D(F^{-1}(x) + O(\varepsilon))) + o(\varepsilon) =$$
$$-\frac{1}{p(F^{-1}(x))}D(F^{-1}(x)) + o(\varepsilon),\ \varepsilon \to 0.$$

Last equality holds as $F^{-1}(x)$ is a continuity poinf of D, therefore it holds for a.s. $x \in [\alpha, 1-\alpha]$. Notice that $o(\varepsilon)$ is uniformly small on $x \in [\alpha, 1-\alpha]$. Thus,

$$\frac{f(F+\varepsilon D) - f(F)}{\varepsilon} \to L_F(D; f) = -\frac{1}{1-2\alpha}\int_{\alpha}^{1-\alpha} \frac{1}{p(F^{-1}(u))}D(F^{-1}(u))du.$$

Therefore,

$$I_F(x) = \frac{1}{1-2\alpha}\int_\alpha^{1-\alpha} \frac{u}{p(F^{-1}(u))}du - \frac{1}{1-2\alpha}\int_\alpha^{1-\alpha} \frac{1}{p(F^{-1}(u))}\delta_x(F^{-1}(u)))du. \qquad (1)$$

The first term in (1) equals to

$$\int_\alpha^{1-\alpha} \frac{u}{p(F^{-1}(u))}du = \int_{x_\alpha}^{x_{1-\alpha}} \frac{F(v)}{f(v)}dF(v) = \int_{x_\alpha}^{x_{1-\alpha}} F(v)dv = \frac{x_{1-\alpha} - x_\alpha}{2}$$

since $F(v-\theta) - 1/2$ is odd function. The second term of (1) is

$$\int_\alpha^{1-\alpha} \frac{1}{p(F^{-1}(u))}\delta_x(F^{-1}(u))du = \int_{x_\alpha}^{x_{1-\alpha}} \delta_x(v)dv.$$

As $x > x_{1-\alpha}$ function $\delta_x(v)$ equals to 0, as $x < x_\alpha$ $\delta_x(v)$ equals to 1, as $x_\alpha < x < x_{1\alpha}$ integral of δ_v equals to $x_{1-\alpha} - x$. Therefore

$$I_F(f) = \begin{cases} -\frac{x_{1-\alpha}-x_\alpha}{2(1-2\alpha)}, & x < x_\alpha, \\ \frac{x-\theta}{1-2\alpha}, & x \in (x_\alpha, x_{1-\alpha}), \\ \frac{x_{1-\alpha}-x_\alpha}{2(1-2\alpha)}, & x \geq x_{1-\alpha}, \end{cases} \qquad \sigma^2(F) = \frac{2}{(1-2\alpha)^2}\left(\int_\theta^{x_{1-\alpha}} (t-\theta)^2 f(t)dt + \alpha(x_{1-\alpha}-\theta)^2\right).$$

□

Problem 1. Prove that

$$\sigma^2(F) \leq \frac{\sigma^2}{(1-2\alpha)^2},$$

where $\sigma^2 = \mathbf{D}X$ and the inequality becomes an equality for some F.

So, the α-trimmered mean is a good robust alternative to the sample mean. However, as the asymptotic breakdown point of X_α equals $1/3$, $\sigma^2(F)$ can be 9 times larger than σ^2.