Homework 4

Liu Huihang
SA18017026
QQ: 184050473
MAIL: huihang@mail.ustc.edu.cn

Problem 1 *Read the first four chapters of Owen (2001).*

Problem 2 *Coding*

Problem 3 *Assume that $(X_i, Y_i), i = 1, \cdots, n$ are i.i.d. samples from population $(\mathbf{X}, \mathbf{Y})$. Parameter of interest is $\tau = \sigma_X^2/\sigma_Y^2$, where σ_X^2 and σ_Y^2 are variance of X and Y respectively. Denote θ by $(\tau, \eta')'$ where η is nuisance parameter. Try to give a set of simultaneous equations to estimate θ and represent the empirical likelihood confidence intercal of τ.*

Solution

To define τ by X and Y, we formulate the following five estimationg equations:

$$\begin{aligned}
0 &= E[X - \mu_x], \\
0 &= E[Y - \mu_y], \\
0 &= E[(X - \mu_x)^2 - \sigma_X^2], \\
0 &= E[(Y - \mu_y)^2 - \sigma_Y^2], \\
0 &= E[\sigma_X^2/\sigma_Y^2 - \tau].
\end{aligned}$$

for the parameter $\theta = (\tau, \mu_x, \mu_y, \sigma_X^2, \sigma_Y^2)'$.

To handle nuisance parameter, write the estimating function as $m(X, Y, \tau, \eta') = 0$, $\tau \in \mathbb{R}, \eta \in \mathbb{R}^4$. The parameters (τ, η') satisfy the equations $E[m(X, Y, \tau, \eta')] = 0$.

Now we define

$$\mathcal{R}(\tau, \eta') = \max \left\{ \prod_{i=1}^{n} n\omega_i | \sum_{i=1}^{n} \omega_i m(X, Y, \tau, \eta') = 0, \omega_i \geq 0, \sum_{i=1}^{n} \omega_i = 1 \right\}$$

and

$$\mathcal{R}(\tau) = \max_{\eta} \mathcal{R}(\tau, \eta').$$

Under mild conditions as described in Chapter 3.10 (Owen, 2001), $-2log\mathcal{R}(\tau) \to \chi^2_{(1)}$.

Thus, a $100(1-\alpha)\%$ empirical likelihood confidence intercal of τ can be represented as

$$\left\{ \tau | -2log\mathcal{R}(\tau) \leq \chi^{2,1-\alpha}_{(1)} \right\}.$$

∎

Problem 4 *Let $X, X_1, \ldots, X_n i.i.d \sim F$, find kernel $h(x_1, x_2, x_3)$ such that $E_F h(X_1, X_2, X_3) = E(X - E_F X)^3$*

Solution

Let's consider the following equations

$$
\begin{aligned}
s_n^3 &= \frac{n}{(n-1)(n-2)} \sum_{i=1}^{n} (X_i - \bar{X}_n)^3 \\
&= -\frac{1}{6n(n-1)(n-2)} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \left(\left(X_i - \bar{X}_n\right)^3 + \left(X_j - \bar{X}_n\right)^3 - 8\left(X_k - \bar{X}_n\right)^3 \right) \\
&= -\frac{1}{6n(n-1)(n-2)} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \left(\left(X_i - \bar{X}_n\right) + \left(X_j - \bar{X}_n\right) - 2\left(X_k - \bar{X}_n\right) \right)^3 \\
&= -\frac{1}{6n(n-1)(n-2)} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \left(X_i + X_j - 2X_k\right)^3 \\
&= \frac{1}{\binom{n}{3}} \sum_{i,j,k} \frac{1}{36} \left(2X_i - X_j - X_k\right)^3
\end{aligned} \tag{1}
$$

Thus, $h(x_1, x_2, x_3) = \frac{1}{36}\left(2x_1 - x_2 - x_3\right)^3$

■

Problem 5 *Prove the* $\zeta_1 = 1/9$ *in slide (page 25).*

Solution

Let $X \sim F(x), Y \sim G(y)$. Because X and Y are independent, $\tau = 0$. Thus we have

$$
\begin{aligned}
\zeta_1 &= Cov(h(P_1, P_2), h(P_1, P_3)) \\
&= E[h(P_1, P_2)h(P_1, P_3)] \\
&= 1 \times P(h(P_1, P_2)h(P_1, P_3) = 1) + (-1) \times P(h(P_1, P_2)h(P_1, P_3) = -1) \\
&= 2P(h(P_1, P_2)h(P_1, P_3) = 1) - 1
\end{aligned}
$$

Denote $P_i = (X_i, Y_i), i = 1, 2, 3$, we have

$$
\begin{aligned}
&P(h(P_1, P_2)h(P_1, P_3) = 1 | P_1, h(P_1, P_2) = 1) \\
=&P(h(P_1, P_3) = 1 | P_1) \\
=&P(X_3 > X_1, Y_3 > Y_1 \text{ or } X_3 < X_1, Y_3 < Y_1) \\
=&(1 - F(X_1))(1 - G(Y_1)) + F(X_1)G(Y_1)
\end{aligned}
$$

and

$$
\begin{aligned}
&P(h(P_1, P_2)h(P_1, P_3) = 1 | P_1, h(P_1, P_2) = -1) \\
=&P(h(P_1, P_3) = -1 | P_1) \\
=&P(X_3 < X_1, Y_3 > Y_1 \text{ or } X_3 > X_1, Y_3 < Y_1) \\
=&F(X_1)(1 - G(Y_1)) + (1 - F(X_1))G(Y_1)
\end{aligned}
$$

and

$$
\begin{aligned}
&P(h(P_1,P_2)h(P_1,P_3)=1|P_1)\\
&=P(h(P_1,P_2)=1,h(P_1,P_3)=1|P_1)+P(h(P_1,P_2)=-1,h(P_1,P_3)=-1|P_1).
\end{aligned}
$$

Thus

$$
\begin{aligned}
&P(h(P_1,P_2)h(P_1,P_3)=1)\\
=&E[P(h(P_1,P_2)h(P_1,P_3)=1|P_1)]\\
=&\int\int[(1-F(X_1))(1-G(Y_1))+F(X_1)G(Y_1)]^2+\\
&\qquad\qquad[F(X_1)(1-G(Y_1))+(1-F(X_1))G(Y_1)]^2dFdG\\
&\frac{5}{9}
\end{aligned}
$$

So, we have $\zeta_1 = 1/9$. ■

Problem 6 *Let $X_1,\ldots,X_n i.i.d \sim Uniform(0,\tau)$ kernel $h(x,y)=|x-y|$ and U-statistics $G_n=\frac{1}{\binom{n}{2}}\sum_{i<j}|X_i-X_j|$, find the limit distribution of G_n.*

Solution

The Hajek projection of $U-\theta$ os ■

References

Art B. Owen. *Empirical Likelihood*. 2001.