Homework 5

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

Problem 1 *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

$$T = E(X - EX)^3 = EX^3 - 3EXEX^2 + 2(EX)^3.$$

So

$$E(X_1 + X_2 + X_3)^3 = 3EX^3 + 18EXEX^2 + 6(EX)^3.$$

Then

$$\begin{aligned} T &= -\frac{1}{6}E\left(X_1 + X_2 + X_3\right)^3 + \frac{3}{2}EX^3 + 3(EX)^3 \\ &= E\left(-\frac{1}{6}\left(X_1 + X_2 + X_3\right)^3 + \frac{1}{2}\left(X_1^3 + X_2^3 + X_3^3\right) + 3\left(X_1 X_2 X_3\right)\right) \end{aligned}$$

Thus

$$h\left(x_1, x_2, x_3\right) = -\frac{1}{6}\left(x_1 + x_2 + x_3\right)^3 + \frac{1}{2}\left(x_1^3 + x_2^3 + x_3^3\right) + 3\left(x_1 x_2 x_3\right)$$

■

Problem 2 *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

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

So, we have $\zeta_1=1/9$. ■

Problem 3 *Let $X_1,\dots,X_n i.i.d\sim$ 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

G_n is a U-statistic of order $r=2$ with kernel $h(x_1,x_2)=|x_1-x_2|$ can also an unbiased estimate of $E|X_1-X_2|$ (Dasgupta, 2008).

From the CLT for U-statistics, it follows that

$$\sqrt{n}(G_n-E|X_1-X_2|)\rightsquigarrow N(0,r^2\zeta_1)$$

where ζ_1 can be calculated as following.

We have $F(x)=x/\tau$, $x\in(0,\tau)$ because $X\sim$ Uniform$(0,\tau)$. Thus

$$
\begin{aligned}
Eh(x,X_2)&=E|x-X_2|\\
&=\int_0^\tau|x-y|\frac{1}{\tau}dy\\
&=\frac{1}{\tau}\left(\int_0^x x-ydy+\int_x^\tau y-xdy\right)\\
&=\frac{x^2}{\tau}-x+\frac{\tau}{2},
\end{aligned}
$$

and

$$
\begin{aligned}
Eh(X_1, X_2) &= E[E[h(X_1, X_2)|X_1]] \\
&= E\left[\frac{X_1^2}{\tau} - X_1 + \frac{\tau}{2}\right] \\
&= \int_0^\tau \frac{1}{\tau}\left(\frac{x^2}{\tau} - x + \frac{\tau}{2}\right) dx \\
&= \frac{\tau}{3},
\end{aligned}
$$

and

$$
h_1(x) = Eh(x, Y) - T = \frac{x^2}{\tau} - x + \frac{\tau}{6}
$$

and

$$
\begin{aligned}
\zeta_1 &= E\left[h_1^2(x)\right] \\
&= E\left(\frac{x^2}{\tau} - x + \frac{\tau}{6}\right)^2 \\
&= \frac{\tau^2}{180}.
\end{aligned}
$$

Thus,

$$
\sqrt{n}\left(G_n - \frac{\tau}{3}\right) \to N\left(0, \frac{\tau^2}{45}\right)
$$

■

References

Anirban Dasgupta. *Asymptotic Theory of Statistics and Probability*. 2008.