Homework 1

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

Problem 1 *Let $X_1, \dots, X_n \sim F$ and let $F_n(x)$ be the empirical distribution function, for a fixed x, find limiting distribution of $\sqrt{F_n(x)}$.*

Solution

We have $F_n = \frac{1}{n}\sum_{i=1}^n I(X_i \leq x)$, and $F_n \sim \frac{1}{n} Binomial(n, F(x))$. By CLT, we get $\frac{nF_n(x)-nF(x)}{\sqrt{n}\sqrt{nF(x)(1-F(x))}} = \frac{F_n(x)-F(x)}{\sqrt{(F(x)(1-F(x)))}} \rightsquigarrow N(0,1)$. Thus, $F_n \rightsquigarrow N(F(x), F(x)(1-F(x)))$. Finally, we get $\sqrt{F_n} \rightsquigarrow ?$

So, $\hat{\tau} = \sqrt{\frac{1}{n}\sum_{i=1}^n F_n(X_i)}$.

Thus, $T(F_n) \rightsquigarrow N(\sqrt{F}, \frac{1}{n^2}\sum_{i=1}^n F_n(X_i))$

■

Problem 2 *Let x, y be two distinct real numbers, find $Cov(F_n(x), F_n(y))$, where F_n be the empirical distribution function.*

Solution

Assume that $x < y$. We have $F_n(x) = \frac{1}{n}\sum_{i=1}^n I(X_i \leq x)$, and $F_n(x) \sim \frac{1}{n} Binomial(n, F(x))$. This gives

$$\begin{aligned} Cov(F_n(x), F_n(y)) &= E[F_n(x) - E(F_n(x))][F_n(y) - E(F_n(y))] \\ &= E[F_n(x)F_n(y)] - E[F_n(x)]E[F_n(y)] \end{aligned}$$

where

$$\begin{aligned} n^2 E[F_n(x)F_n(y)] &= E\left[\sum_{i=1}^n I(X_i \leq x) \sum_{j=1}^n I(X_j \leq y)\right] \\ &= E\left[\sum_{i=1}^n \sum_{j=1}^n I(X_i \leq x) I(X_j \leq y)\right] \\ &= \sum_{i=1}^n \sum_{j=1}^n E\left[I(X_i \leq x) I(X_j \leq y)\right] \\ &= \sum_{i=1}^n E\left[I(X_i \leq x) I(X_i \leq y)\right] + \sum_{i,j=1; i\neq j}^n E\left[I(X_i \leq x)\right] E\left[I(X_j \leq y)\right] \\ &= nF(x) + n(n-1)F(x)F(y) \end{aligned}$$

and $E[F_n(x)]E[F_n(y)] = F(x)F(y)$.

Thus, $Cov(F_n(x), F_n(y)) = \frac{F(x)+(n-1)F(x)F(y)}{n} - F(x)F(y) = \frac{F(x)-F(x)F(y)}{n}$ ■

Problem 3 *Let* $X_{(1)} \leq \cdots \leq X_{(n)}$ *be order statistics from continuous popolation* F*, prove that for any* $0 < \beta < 1$

$$P(F(X_{(n)}) - F(X_{(1)}) > \beta) = 1 - n\beta^{n-1} + (n-1)\beta^n.$$

Solution

We have $\mathbf{P}\left(F(X_{(n)}) \leq x, F(X_{(1)}) \leq y\right) = x^n\left(1-(1-y)^n\right)$ which leads to the joint probability density of $F(X_{(n)})$ and $F(X_{(1)})$, $f_{F(X_{(n)}),F(X_{(1)})}(x,y) = \frac{\partial}{\partial x \partial y} x^n\left(1-(1-y)^n\right) = n^2 x^{n-1}(1-y)^{n-1}$.

Thus, $\mathbf{P}(F(X_{(n)}) - F(X_{(1)}) > \beta) = \int_{x-y>\beta} f_{F(X_{(n)}),F(X_{(1)})}(x,y)dxdy = \int_0^{1-\beta}\int_{y+\beta}^1 n^2x^{n-1}(1-y)^{n-1}dydx = \int_0^{1-\beta}(1-(y-\beta)^n)(1-y)^{n-1}dy = 1-\beta^n - \int_0^{1-\beta}(y+\beta)^n(1-y)^{n-1}dy =$

$F(X_{(n)}$ ∎

Problem 4 *Let* $X_1, \ldots, X_n$ *be simple samples from* $U(0,1)$*, prove that sample median* $\hat{\xi}_{n,1/2}$ *has asymptotic distribution* $N(\frac{1}{2}, \frac{1}{4n})$.