Home Assignment 1

Matthew Zhao

1/14/2023

Lecture 1 HW

Problem 1

$$\hat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X_n})^2$$

$$= \frac{1}{n}\sum_{i=1}^{n} X_i^2 - 2X_i\bar{X_n} + \bar{X_n}^2$$

$$= \frac{1}{n}\sum_{i=1}^{n}(X_i^2) - 2(\frac{1}{n}\sum_{i=1}^{n} X_i\bar{X_n}) + \frac{1}{n}(\sum_{i=1}^{n}\bar{X_n})^2$$

$$= \frac{1}{n}\sum_{i=1}^{n}(X_i^2) - 2(\frac{1}{n}\sum_{i=1}^{n} X_i)^2 + \frac{1}{n}(\sum_{i=1}^{n} X_i)^2$$

$$= \frac{1}{n}\sum_{i=1}^{n}(X_i^2) - (\frac{1}{n}\sum_{i=1}^{n} X_i)^2$$

By LLN, $\frac{1}{n}\sum_{i=1}^{n}(X_i^2) \xrightarrow{p} E_P(X^2)$ and $\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p} E_P(X)$

By CMT, $\hat{\sigma}_n^2 \xrightarrow{p} E_P(X^2) - E_P(X)^2 = Var_P(X)$

Problem 2

Intuitively, as $n \to \infty$, the probability that $X_1,\ldots,X_n$ does not contain θ approaches 0 by definition of uniform distribution. Hence, the probability that our estimator for θ ($\max\{X_1, ..., X_n\}$) does not equal θ also goes to zero, thus $P(|\hat{\theta} - \theta| > \epsilon) \to 0$ as $n \to \infty$.

Formally, we can derive the CDF F_θ of our estimator and use it to show that as $n \to \infty$, our estimator converges in probability to θ. Specifically, we can use order statistics to show that $F_\theta(x) = F(x)^n$ where $F(x)$ is the CDF for each individual X_i drawn from $U[0, \theta]$:

$$F_\theta(x) = \mathbb{P}(X_1, ..., X_n \leq x)$$

$$= \mathbb{P}(X_1 \leq x)\mathbb{P}(X_2 \leq x)...\mathbb{P}(X_n \leq x)$$
$$= F(x)^n$$

Plugging this result in we show:

$$\mathbb{P}(|\hat{\theta}_n - \theta| > \epsilon) = \mathbb{P}(\theta - \epsilon < \hat{\theta}_n < \theta + \epsilon)$$

$$
\begin{aligned}
&= F_\theta(\theta+\epsilon) - F_\theta(\theta-\epsilon) \\
&= F(\theta+\epsilon)^n - F(\theta-\epsilon)^n \\
&= \frac{(\theta+\epsilon)^n}{\theta} - \frac{(\theta-\epsilon)^n}{\theta} \to 0
\end{aligned}
$$

as $\epsilon \to 0$, thus $\hat{\theta}_n \xrightarrow{p} \theta$

Problem 3

```
library(tidyverse)
library(haven)
library(stargazer)
library(xtable)
options(scipen=999)

make_ci <- function(xbar,var){
  c(xbar - 1.96*var, xbar + 1.96*var)
}

ci_coverage <- function(n){
  X <- rnorm(n)
  xbar <- mean(X)
  sd <- sd(X)

  known <- make_ci(xbar,1/sqrt(n))
  est <- make_ci(xbar,sd/sqrt(n))

  c(0 >= known[1] & 0 <= known[2],
    0 >= est[1] & 0 <= est[2])
}

known_results <- c()
est_results <- c()
N <- c(30,100,500)
for (n in N){
  known_temp <- c()
  est_temp <- c()
  for (x in 1:1000){
    known_temp <- c(known_temp,ci_coverage(n)[1])
    est_temp <- c(est_temp,ci_coverage(n)[2])
  }
  known_results <- c(known_results,mean(known_temp))
  est_results <- c(est_results,mean(est_temp))
}
data.frame(N,est_results,known_results)
```

```
##     N est_results known_results
## 1  30       0.951         0.962
## 2 100       0.954         0.949
## 3 500       0.946         0.950
```

Lecture 2 HW

Problem 1

(1)

(2)

(3)

(4)

(5)

(6)

Problem 2

(0)

(1)

(2)

(3)