Econ 6190 Problem Set 5

Fall 2024

1. [Hansen 7.12] Take a random variable Z such that $\mathbb{E}[Z] = 0$ and $\text{var}[Z] = 1$. Use Chebyshev's inequality to find a δ such that $P[|Z| > \delta] \leq 0.05$. Contrast this with the exact δ which solves $P[|Z| > \delta] = 0.05$ when $Z \sim \text{N}(0, 1)$. Comment on the difference.

2. [Second exam, 2022] Let X be a random variable following a normal distribution with mean μ and variance $\sigma^2 > 0$. We draw a random sample $\{X_1, X_2, \ldots X_n\}$ from X and construct a sample mean statistic $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

   (a) Fix $\delta > 0$. Find an upper bound of $P\{|\bar{X} - \mu| > \delta\}$ by using Markov inequality with $r = 2$.

   (b) Repeat the exercise (a) but using Markov inequality with $r = 4$.

   (c) Compare the two bounds in (a) and (b) above when $\delta = \sigma$ and when n is at least 2. Which one of them gives you a tighter bound of $P\{|\bar{X} - \mu| > \delta\}$?

   (d) Since we know X is normal, find the exact value of $P\{|\bar{X} - \mu| > \delta\}$.

   (e) From (d), we see that the tail probability of a normal sample mean is much thinner than what Markov inequality predicts. In fact, we can show that if $Z \sim N(\mu, \sigma^2)$, then

   $$P\{|Z - \mu| > \delta\} \leq 2\exp\left(-\frac{\delta^2}{2\sigma^2}\right). \tag{1}$$

   Given (1), find a constant c such that

   $$P\{|\bar{X} - \mu| \leq c\} > 0.95.$$

   That is, we can predict that with a probability of at least 0.95, sample average is within $c-$distance of its true mean. What is the prediction of c if you only use Chebyshev's inequality?

   (f) Given your answer to (e), how much more data do we have to collect if we want the prediction of c based on Chebyshev's inequality to be the same as that based on (1)

3. Consider a sample of data $\{X_1, \ldots X_n\}$, where

$$X_i = \mu + \sigma_i e_i, i = 1 \ldots n,$$

where $\{e_i\}_{i=1}^n$ are iid and $\mathbb{E}[e_i] = 0$, $\text{var}(e_i) = 1$, $\{\sigma_i\}_{i=1}^n$ are n finite and positive constants, and $\mu \in \mathbb{R}$ is the parameter of interest.

(a) Let

$$\hat{\mu}_1 = \frac{1}{n} \sum_{i=1}^{n} X_i$$

be the sample mean estimator. Under what condition is $\hat{\mu}_1$ a consistent estimator of μ? Under what condition is $\hat{\mu}_1 - \mu = O_p(\frac{1}{\sqrt{n}})$?

(b) Let

$$\hat{\mu}_2 = \frac{\frac{1}{n} \sum_{i=1}^n \frac{X_i}{\sigma_i^2}}{\frac{1}{n} \sum_{i=1}^n \frac{1}{\sigma_i^2}}$$

be an alternative estimator of μ. Under what condition is $\hat{\mu}_2$ a consistent estimator of μ? Under what condition is $\hat{\mu}_2 - \mu = O_p(\frac{1}{\sqrt{n}})$?

(c) Compare the MSE of $\hat{\mu}_1$ and $\hat{\mu}_2$. Which one is more efficient and why?

4. Suppose that $X_n Y_n \xrightarrow{d} Y$ and $Y_n \xrightarrow{p} 0$. Suppose a function f is continuously differentiable at 0, show that $X_n(f(Y_n) - f(0)) \xrightarrow{d} f'(0)Y$, where $f'(0)$ is the first derivative of f at 0.

5. Let $\{X_1 \ldots X_n\}$ be a sequence of i.i.d random variables with mean μ and and variance σ^2. Let $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$.

(a) If $\mu \neq 0$, how would you approximate the distribution of $(\bar{X})^2$ in large samples as $n \to \infty$?

(b) If $\mu = 0$, how would you approximate the distribution of $(\bar{X})^2$ in large samples as $n \to \infty$?