Econ 6190 Problem Set 6

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 $\{X_1, \dots X_n\}$, where

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

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 $\{X_1 \dots X_n\}$ 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 $n \to \infty$?

1. Note since $\mathbb{E}Z = 0$, $\mathbb{E}Z^2 = \text{var}(Z)$. Hence by Chebyshev's inequality,

$$P[|Z| > \delta] \leq \frac{\mathbb{E}[Z^2]}{\delta^2} = \frac{\text{var}(Z)}{\delta^2} = \frac{1}{\delta^2}.$$

Let $\frac{1}{\delta^2} = 0.05$, we find $\delta = \sqrt{20} \approx 4.47$

On the other hand, if we know Z is standard normal, let $\Phi(\cdot)$ be the cdf of a standard normal. It follows

$$\begin{aligned} P[|Z| > \delta] =& P\{Z > \delta\} + P\{Z < -\delta\} \\ =& 1 - \Phi(\delta) + \Phi(-\delta) \\ =& 2(1 - \Phi(\delta)) \end{aligned}$$

Setting $2(1 - \Phi(\delta)) = 0.05$, we get $\Phi(\delta) = 1 - 0.025 = 0.975$. That is, δ is the 97.5 percent quantile of a standard normal. Looking from the statistical tables, $\delta \approx 1.96$.

If we do not know the distribution of Z, we get

$$P[|Z| > 4.47] \leq 0.05, \tag{1}$$

which holds for all distributions with mean 0 and variance 1. On the other hand, if we know the distribution of Z (say standard normal), we can get a much sharper bound:

$$P[|Z| > 1.96] = 0.05 \tag{2}$$

which only holds for this specific distribution. Note even when Z is standard normal, (1) is still a correct statement. It is just less sharp than (2).

2.

(a) **[5 pts]** Fix $\delta > 0$. Find an upper bound of $P\{|\bar{X} - \mu| > \delta\}$ by using Markov inequality when $r = 2$.

Answer: $P\{|\bar{X} - \mu| > \delta\} \leq \frac{\mathbb{E}[\bar{X}-\mu]^2}{\delta^2} = \frac{bias(\bar{X})+var(\bar{X})}{\delta^2}$. *Since* $\bar{X}$ *is unbiased,* $bias(\bar{X}) = 0$. *Also,* $var(\bar{X}) = \frac{\sigma^2}{n}$. *Thus,* $P\{|\bar{X} - \mu| > \delta\} \leq \frac{\mathbb{E}[\bar{X}-\mu]^2}{\delta^2} = \frac{\sigma^2}{\delta^2 n}$.

(b) **[5 pts]** Repeat the exercise (b) but using Markov inequality when $r = 4$.

Answer:$P\{|\bar{X} - \mu| > \delta\} \leq \frac{\mathbb{E}[\bar{X}-\mu]^4}{\delta^4}$. *Notice since* X *is normal,* $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$. *Therefore,* $\mathbb{E}[\bar{X} - \mu]^4 = \mathbb{E}[\bar{X} - \mathbb{E}[\bar{X}]]^4$, *which is the fourth-th centralized moment of* $\bar{X}$, *equalling* $3\frac{\sigma^4}{n^2}$. *It follows* $P\{|\bar{X} - \mu| > \delta\} \leq \frac{3\sigma^4}{\delta^4 n^2}$

(c) **[5 pts]** 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 $P\{|\bar{X} - \mu| > \sigma\}$.

Answer: When $\delta = \sigma$, *using* $r = 2$ *yields* $P\{|\bar{X} - \mu| > \delta\} \leq \frac{1}{n}$, *while using* $r = 4$ *yields* $P\{|\bar{X} - \mu| > \delta\} \leq \frac{3}{n^2}$.

Therefore, when $n > 3$, $\frac{3}{n^2} < \frac{1}{n}$, *applying* $r = 4$ *gives a tighter bound; if* $n = 3$, *they give the same bound. If* $n = 2$, *then applying* $r = 2$ *gives a tighter bound.*

(d) **[5 pts]** Since we know X is normal, find the exact value of $P\{|\bar{X} - \mu| > \delta\}$.

Answer: $P\{|\bar{X} - \mu| > \delta\} = P\left\{\left|\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right| > \frac{\delta}{\frac{\sigma}{\sqrt{n}}}\right\} = 2\left(1 - \Phi\left(\frac{\delta\sqrt{n}}{\sigma}\right)\right)$

(e) **[10 pts]** 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?

Answer: It suffices to find c *such that* $P\{|\bar{X} - \mu| > c\} \leq 0.05$. *Note again* $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$. *Therefore, to use (1), set* $c = 2\exp\left(-\frac{c^2 n}{2\sigma^2}\right) = 0.05$. *It follows* $c = \frac{\sigma}{\sqrt{n}}\sqrt{2\log 40} \approx 2.72\frac{\sigma}{\sqrt{n}}$. *If* Chebyshev's inequality were used, then we need to set $\frac{\sigma^2}{c^2 n} = 0.05$, i.e., $c = \sqrt{20}\frac{\sigma}{\sqrt{n}} \approx 4.47\frac{\sigma}{\sqrt{n}}$.

(f) **[5 pts]** 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)?

Answer: let n_c *be the sample size based on Chebyshev's prediction, and let* n_1 *be the sample size based on (1). Setting* $4.47\frac{\sigma}{\sqrt{n_c}} = 2.72\frac{\sigma}{\sqrt{n_1}}$ *implies* $\frac{\sqrt{n_c}}{\sqrt{n_1}} = \frac{4.47}{2.72}$. *That is,* $n_c \approx \left(\frac{4.47}{2.72}\right)^2 n_1$, *i.e., we have to collect around 1.7 times more data if we only uses* Chebyshev's inequality.

3.

(a) **[5 pts]** 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 conditions is $\hat{\mu}_1 - \mu = O_p(\frac{1}{\sqrt{n}})$?

Answer: Clearly $\mathbb{E}[X_i] = \mu$, *i.e.,* $\hat{\mu}_1$ *is unbiased. Also,* $\hat{\mu}_1 = \frac{1}{n}\sum_{i=1}^{n}(\mu + \sigma_i e_i) = \mu + \frac{1}{n}\sum_{i=1}^{n}\sigma_i e_i$. *Thus,* $var(\hat{\mu}_1) = var(\frac{1}{n}\sum_{i=1}^{n}\sigma_i e_i) = \frac{1}{n^2}\sum_{i=1}^{n}\sigma_i^2$ *(by iid assumption of* $\{e_i\}_{i=1}^{n}$). Thus, by Chebyshev's inequality, $\hat{\mu}_1$ *is consistent if* $\frac{1}{n^2}\sum_{i=1}^{n}\sigma_i^2 = o(1)$, *and* $\hat{\mu}_1 - \mu = O_p(\frac{1}{\sqrt{n}})$ if $\frac{1}{n}\sum_{i=1}^{n}\sigma_i^2 = O(1)$ *(or equivalently,* $\frac{1}{n}\sum_{i=1}^{n}\sigma_i^2$ *is asymptotically bounded). [an answer of i.i.d leads to consistency gets 0 points.]*

(b) **[10 pts]** 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 conditions is $\hat{\mu}_2 - \mu = O_p(\frac{1}{\sqrt{n}})$?

Answer: First, note $\hat{\mu}_2$ *is also unbiased. Also,* $\hat{\mu}_2 = \frac{\frac{1}{n}\sum_{i=1}^{n}\frac{X_i}{\sigma_i^2}}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{\sigma_i^2}} = \mu + \frac{\frac{1}{n}\sum_{i=1}^{n}\frac{e_i}{\sigma_i}}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{\sigma_i^2}}$. *Thus,*

$var(\hat{\mu}_2) = var\left(\frac{\frac{1}{n}\sum_{i=1}^{n}\frac{e_i}{\sigma_i}}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{\sigma_i^2}}\right) = \frac{1}{\left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)^2}\frac{1}{n^2}\sum_{i=1}^{n}\frac{1}{\sigma_i^2} = \frac{1}{n}\frac{1}{\left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)} = \frac{1}{\sum_{i=1}^{n}\frac{1}{\sigma_i^2}}$. *Thus,* $\hat{\mu}_2$ *is consistent if* $\sum_{i=1}^{n}\frac{1}{\sigma_i^2} \to \infty$ *as* $n \to \infty$ *(or equivalently,* $\frac{1}{\sum_{i=1}^{n}\frac{1}{\sigma_i^2}} = o(1)$*). And* $\hat{\mu}_2 - \mu = O_p(\frac{1}{\sqrt{n}})$ *if* $\frac{n}{\sum_{i=1}^{n}\frac{1}{\sigma_i^2}}$ *is asymptotically bounded.*

(c) **[10 pts]** Compare the MSE of $\hat{\mu}_1$ and $\hat{\mu}_2$. Which one is more efficient?

Answer: Both of them are unbiased. The one with a smaller variance is more efficient.

$$var(\hat{\mu}_1) = \frac{1}{n}\underbrace{\frac{1}{n}\sum_{i=1}^{n}\sigma_i^2}_{arithmetic\ mean}$$

$$var(\hat{\mu}_2) = \frac{1}{n}\underbrace{\frac{1}{\left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)}}_{harmonic\ mean}$$

Since harmonic mean is always no bigger than arithmetic mean for positive numbers, it follows $var(\hat{\mu}_2) \le var(\hat{\mu}_1)$, *i.e.,* $\hat{\mu}_2$ *is at least as efficient as* $\hat{\mu}_1$. *In fact, as long as there is some* $\sigma_i \neq \sigma_j$ *for* $i \neq j$, *then* $var(\hat{\mu}_2) < var(\hat{\mu}_1)$.

4.

By mean value theorem or Taylor expansion:

$f(Y_n) - f(0) = f'(\bar{Y})(Y_n - 0) = f'(\bar{Y})Y_n$, where $\bar{Y}$ lies on the line between Y_n and 0. Therefore we have:

$$X_n\left[f(Y_n) - f(0)\right] = f'(\bar{Y})X_nY_n$$

Note:

- $X_nY_n \xrightarrow{d} Y$ as given.
- $f'(\bar{Y}) \xrightarrow{p} f'(0)$ ($Y_n \xrightarrow{p} 0$. Therefore, as $\bar{Y}$ lies on the line between Y_n and 0, it implies $\bar{Y} \xrightarrow{p} 0$ too. The claim follows by continuous mapping theorem.)

Conclusion follows by continuous mapping theorem.

5.

(a)

Let $f(x) = x^2$. So we are required to derive the asymptotic distribution of $f(\bar{x})$ using delta method.

Step 1 Do the expansion(of $f(\bar{x})$ around $f(u)$)

- $f(\bar{x}) - f(u) = f'(\tilde{x})(\bar{x} - u)$, where $\tilde{x}$ lies on the line between $\bar{x}$ and u.
- Therefore we have:

$$\sqrt{n}\left[f(\bar{x}) - f(u)\right] = f'(\tilde{x})\sqrt{n}(\bar{x} - u)$$

- $\sqrt{n}(\bar{x} - u) \xrightarrow{d} N(0, \sigma^2)$ by central limit theorem for i.i.d. data.
- $f'(\tilde{x}) \xrightarrow{p} f'(u)$ ($\bar{x} \xrightarrow{p} u$ by Khintchine Law of large numbers. Therefore, as $\tilde{x}$ lies on the line between $\bar{x}$ and u, it implies $\tilde{x} \xrightarrow{p} u$ too. The claim follows by continuous mapping theorem.)

Step 2 Therefore we have form