ECON6190 Section 11

Nov. 8. 2024

Yiwei Sun

- *If T follows a binomial distribution with parameter n (the total number of iid Bernoulli trials) and $\mu \in (0,1)$ (probability of "successes" in each Bernoulli trial), then:*

$$Var(T) = n\mu(1-\mu),$$
$$E(T) = n\mu,$$
$$E\left[\left(\frac{T-E(T)}{\sqrt{Var(T)}}\right)^4\right] = \frac{1-6\mu(1-\mu)}{n\mu(1-\mu)}.$$

1. **[60 pts]** Let X be a Bernoulli random variable (that is, $X \in \{0,1\}$, with $Pr\{X=1\} = \mu \in (0,1)$). 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) **[20 pts]** State and prove Markov inequality.

For each $r>0$, assume $E[|X|^r] < \infty$, $P(|X|>\delta) \leq \frac{E[|X|^r]}{\delta^r}$, $\forall \delta > 0$.

Proof. $P(|X|>\delta) = E[\mathbb{1}\{|X|>\delta\}]$

$\leq E\left[\mathbb{1}\{|X|>\delta\}\frac{|X|^r}{\delta^r}\right]$

For $|X|>\delta$, $\left(\frac{|X|}{\delta}\right)^r > 1$

For $|X| \leq \delta$, multiplying by $\frac{|X|^r}{\delta^r}$ doesn't matter b/c $\mathbb{1}\{|X|>\delta\} = 0$.

$= \frac{1}{\delta^r} E[\mathbb{1}\{|X|>\delta\}|X|^r]$

$\leq \frac{E[|X|^r]}{\delta^r}$

take expetation of $|X|^r$, $\forall |X|>\delta$. For $|X|<\delta$, let it be 0.

$|X|^r \geq 0$, for $|X|<\delta$.

   (b) **[10 pts]** Fix $\delta > 0$. Find an upper bound of $Pr\{|\bar{X}-\mu| > \delta\}$ by using Markov inequality with $r = 2$.

$P(|\bar{X}-\mu|>\delta) \leq \frac{E[(\bar{X}-\mu)^2]}{\delta^2} = \frac{\mu(1-\mu)}{n\delta^2}$, for $\delta>0$.

$$E[(\bar{X}-\mu)^2] = (\text{bias}(\bar{X}))^2 + var(\bar{X})$$
$$= 0 + var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)$$
$$= \underbrace{\frac{1}{n^2} var\Big(\underbrace{\sum_{i=1}^{n} X_i}_{\text{binominal}(n,\mu)}\Big)}_{=n\mu(1-\mu)}$$
$$= \frac{\mu(1-\mu)}{n}$$

(c) **[10 pts]** Repeat the exercise (b) but with $r = 4$.

$$P(|\bar{X}-\mu|>\delta) \leq \frac{E[(\bar{X}-\mu)^4]}{\delta^4} = \frac{\mu(1-\mu)(1-6\mu(1-\mu))}{n^3\delta^4}, \quad \delta>0.$$

Notice: $T = n\bar{X}$, $E[T]=n\mu$, $var(T)=n\mu(1-\mu)$

$$\begin{aligned}
E[(\bar{X}-\mu)^4] &= E\left[\left(\frac{T}{n}-\frac{E[T]}{n}\right)^4\right] \\
&= E\left[\left(\frac{T-E[T]}{n}\right)^4\right] \\
&= E\left[\left(\frac{T-E[T]}{\sqrt{var(T)}}\cdot\frac{\sqrt{var(T)}}{n}\right)^4\right] \\
&= \frac{(var(T))^2}{n^4}\cdot\overbrace{\frac{1-6\mu(1-\mu)}{n\mu(1-\mu)}}^{\text{by the hint}} \\
&= \frac{\cancel{n^2}\mu^{\cancel{2}}(1-\mu)^{\cancel{2}}}{n^{\cancel{4}3}}\cdot\frac{1-6\mu(1-\mu)}{\cancel{n}\cancel{\mu}(1-\cancel{\mu})} \\
&= \frac{\mu(1-\mu)(1-6\mu(1-\mu))}{n^3}
\end{aligned}$$

(d) **[20 pts]** Markov inequality applies to any distribution with finite moments. Since X in this case is bounded, we can also apply the so-called Hoeffding's inequality to get a different bound:

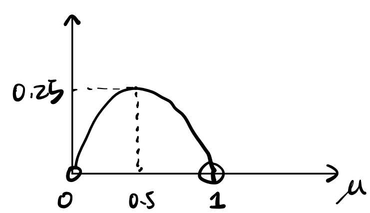
$$Pr\left\{\left|\bar{X}-\mu\right|>\delta\right\} \leq 2\exp\left(-2\delta^2 n\right). \tag{1}$$

Given (1), What is the prediction of the tail probability $Pr\left\{\left|\bar{X}-\mu\right|>\delta\right\}$ when $\delta = 0.1$ and sample size $n = 100$? What is the prediction of the same tail probability if you use Markov inequality with $r = 2$? Which one gives you a better (i.e., tighter) bound? (You may take that $\exp(-2) = 0.14$)

Hoeffding inequality: $P(|\bar{X}-\mu|>\delta) \leq 2\exp(-2\cdot\overbrace{(0.1)^2\cdot 100}^{1})$

$$= 2\underbrace{\exp(-2)}_{0.14}$$

$$= 0.28$$



Markov $r=2$: $P(|\bar{X}-\mu|>\delta) \leq \dfrac{\mu(1-\mu)}{\underbrace{100(0.1)^2}_{1}}$

Since $\mu\in(0,1)$, $\mu(1-\mu)\in(0, 0.25]$.

So Markov always give a tighter bound.

2. **[40 pts]** Let $\{X_1 \dots X_n\}$ be a sequence of i.i.d random variables with mean μ and and variance σ^2. Let $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$.

(a) **[10 pts]** If $\mu \neq 0$, how you would approximate the distribution of $(\bar{X}_n)^3$ (after suitable normalization) in large samples as $n \to \infty$?

(b) **[10 pts]** If $\mu \neq 0$, derive the sharpest possible stochastic order of magnitude for $(\bar{X}_n)^3$.

(a) see section 9 for detailed notes

By CLT, $\sqrt{n}(\bar{x} - \mu) \xrightarrow{d} N(0, \sigma^2)$

By delta method $\sqrt{n}((\bar{X})^3 - (\mu)^3) \xrightarrow{d} N(0, (3\mu^2)^2\sigma^2)$

$\Rightarrow \sqrt{n}((\bar{x})^3 - \mu^3) \xrightarrow{d} N(0, 9\mu^4\sigma^2)$

(b) $\sqrt{n}((\bar{x})^3 - \mu^3) = O_p(1)$

$(\bar{x})^3 - \underbrace{\mu^3}_{O_p(1)} = O_p(\frac{1}{\sqrt{n}})$

$(\bar{x})^3 = O_p(\frac{1}{\sqrt{n}}) + O_p(1) = O_p(\max\{\frac{1}{\sqrt{n}}, 1\}) = O_p(1)$.

Or alternatively, $\bar{X} - \underbrace{\mu}_{O_p(1)} = O_p(\frac{1}{\sqrt{n}})$ by CLT / $\hat{\theta} - \theta = O_p(\sqrt{MSE(\hat{\theta})})$

$\bar{X} = O_p(\frac{1}{\sqrt{n}}) + O_p(1) = O_p(1)$

$(\bar{x})^3 = O_p(1)\,O_p(1)\,O_p(1) = O_p(1)$.

Is $(\bar{x})^3$ unbiased? $\underbrace{E[(\bar{x})^3]}_{} \overset{?}{=} \mu^3$, No! b/c $f(x) = x^3$ is not linear

$= E[(\frac{1}{n}\sum_{i=1}^{n} x_i)^3] \neq E[\frac{1}{n}\sum_{i=1}^{n}(x_i^3)]$

$\rightarrow$ need to apply Jensen's inequality.

(c) **[10 pts]** If $\mu = 0$, how would you approximate the distribution of $(\bar{X}_n)^3$ (after suitable normalization) in large samples as $n \to \infty$?

(d) **[10 pts]** If $\mu = 0$, derive the sharpest possible stochastic order of magnitude for $(\bar{X}_n)^3$.

(c) By CLT, $\sqrt{n}(\bar{x} - 0) \xrightarrow{d} N(0, \sigma^2) \Rightarrow \frac{\sqrt{n}}{\sigma}\bar{x} \xrightarrow{d} N(0,1)$

$(\frac{\sqrt{n}}{\sigma}\bar{x})^3 \xrightarrow{d} (N(0,1))^3$

or $(\sqrt{n}(\bar{x}))^3 \xrightarrow{d} (N(0, \sigma^2))^3$.

(d) $n^{\frac{3}{2}}(\bar{X})^3 = O_P(1)$

$$\Rightarrow (\bar{X})^3 = O_P(n^{-\frac{3}{2}})$$

or alternatively, $\bar{X} - \underbrace{\mu}_{=0} = O_P(\frac{1}{\sqrt{n}})$

$$\bar{X} = O_P(\frac{1}{\sqrt{n}})$$

$$(\bar{X})^3 = O_P(\frac{1}{\sqrt{n}}) \cdot O_P(\frac{1}{\sqrt{n}}) \cdot O_P(\frac{1}{\sqrt{n}}) = O_P\left(\frac{1}{\sqrt{n}^3}\right)$$