Ans: $var(\hat{\lambda}_1) = \frac{var(X_1)}{n} = \frac{\lambda}{n}$ $\quad var(\hat{\lambda}_2) = \frac{var(X_1)}{2} = \frac{\lambda}{2}$

Since $n > 2$, we have $var(\hat{\lambda}_1) < var(\hat{\lambda}_2)$. Also, $E[\hat{\lambda}_1] = E[\hat{\lambda}_2] = \lambda$.

Hence, $\hat{\lambda}_1$ is more efficient.

Order statistic.

Let $X_1, \cdots, X_n$ be i.i.d. r.v.s with CDF F and pdf f.

Sort $X_1, \cdots, X_n$ by an increasing order:

$$X_{(1)} \leq \cdots \leq X_{(n)}$$

where $\underbrace{X_{(1)}}_{\text{first order statistic}} = \min\{X_1, \cdots, X_n\}$, $\underbrace{X_{(n)}}_{n\text{-th order statistic}} = \max\{X_1, \cdots, X_n\}$

Find the CDF and pdf of $X_{(1)}$ and $X_{(n)}$.

Ans: ① CDF of $X_{(1)}$:

$$F_{X_{(1)}}(x) \overset{\text{def}}{=} \mathbb{P}(X_{(1)} \leq x) = 1 - \mathbb{P}(X_{(1)} > x) = 1 - \mathbb{P}(X_1 > x, \cdots, X_n > x)$$

$$\overset{\text{ind}}{=} 1 - \prod_{i=1}^{n} \mathbb{P}(X_i > x) \overset{X_i \sim F \Leftrightarrow \mathbb{P}(X_i > x) = 1 - F(x)}{=} 1 - \prod_{i=1}^{n}(1 - F(x)) = 1 - (1 - F(x))^n$$

pdf of $X_{(1)}$:

$$f_{X_{(1)}}(x) = \frac{d}{dx} F_{X_{(1)}}(x) = n\left(\frac{d}{dx}F(x)\right)\cdot(1 - F(x))^{n-1} = nf(x)\cdot(1 - F(x))^{n-1}$$

② CDF of $X_{(n)}$:

$$F_{X_{(n)}}(x) \overset{\text{def}}{=} \mathbb{P}(X_{(n)} \leq x) = \mathbb{P}(X_1 \leq x, \cdots, X_n \leq x)$$

$$\overset{\text{ind}}{=} \prod_{i=1}^{n} \mathbb{P}(X_i \leq x) = \prod_{i=1}^{n} F(x) = (F(x))^n$$

pdf of $X_{(n)}$:

$$f_{X_{(n)}}(x) = \frac{d}{dx} F_{X_{(n)}}(x) = nf(x)\cdot(F(x))^{n-1} \quad (*)$$

Ex: $X_1, \cdots, X_n \overset{\text{iid}}{\sim} Unif(0, \theta)$. We have $\hat{\theta}_{MLE} = X_{(n)}$. then

$$bias(\hat{\theta}_{MLE}) = E[\hat{\theta}_{MLE}] - \theta$$

$$f_{\hat{\theta}_{MLE}}(x) \overset{(*)}{=} n\cdot\frac{1}{\theta}\cdot\left(\frac{x}{\theta}\right)^{n-1}, \quad 0 < x < \theta$$

then $E[\hat{\theta}_{MLE}] = \int_0^\theta x \cdot f_{\hat{\theta}_{MLE}}(x)\,dx = \int_0^\theta n\cdot\frac{1}{\theta}\cdot\frac{x^n}{\theta^{n-1}}\,dx = \frac{n}{\theta^n}\int_0^\theta x^n\,dx$

$$= \frac{n}{\theta^n}\cdot\frac{1}{n+1}\cdot x^{n+1}\Big|_0^\theta = \frac{n}{n+1}\theta$$

$bias(\hat{\theta}_{MLE}) = \frac{n}{n+1}\theta - \theta = -\frac{1}{n+1}\theta \xrightarrow{n\to\infty} 0$ asymptotically unbiased.

MLE theory:

Recall CLT: asymptotic dist of $\bar{X}_n$ or S_n

we use: $\bar{X}_n \overset{approx}{\sim} N\left(E[X_1], \frac{var[X_1]}{n}\right)$

$S_n \overset{approx}{\sim} N\left(n E[X_1], n\, var[X_1]\right)$

} in practice for large, finite n (eg $n=100$)

in theory: $\sqrt{n}\left(\bar{X}_n - E[X_1]\right) \xrightarrow{d} N\left(0, var[X_1]\right)$

r.v. sequence: $n=1$ | $n=2$ ... $n=n$

$X_1 - E[X_1]$ | $\sqrt{2}\left(\frac{X_1+X_2}{2} - E[X_1]\right)$ ... $\sqrt{n}\left(\bar{X}_n - E[X_1]\right)$

MLE asymptotics: $\sqrt{n}\left(\hat{\theta}_{MLE} - \theta\right) \xrightarrow{d} N\left(0, \frac{1}{I(\theta)}\right)$

where $\hat{\theta}_{MLE}$ is based on $X_1, \cdots, X_n$ iid

In practice, $\hat{\theta}_{MLE} \overset{approx}{\sim} N\left(\theta, \frac{1}{nI(\theta)}\right)$ for large, finite n

asymptotically, $\hat{\theta}_{MLE}$ is unbiased and achieves the C-R lower bound

---

Confidence intervals (CIs):

$X_1, \cdots, X_n$ iid w/ $E[X_1] = \mu$, $\underline{var[X_1] = \sigma^2 \text{ (known)}}$

Use $\bar{X}_n$ as an estimator of μ.

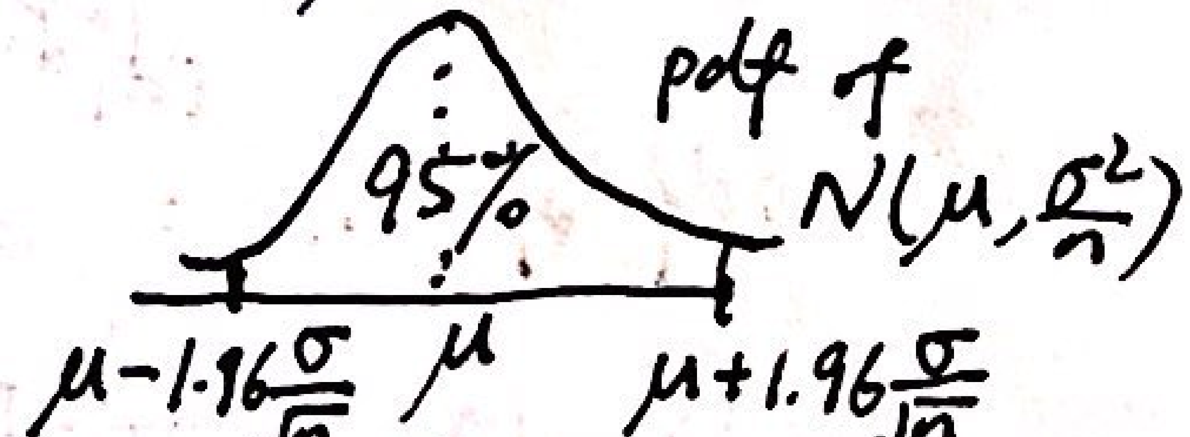
Q: Find a "random interval" that covers μ w/ 95% probability

— 95% CI

Approach: use the distribution of $\bar{X}_n$ around μ

CLT says when n is large, $\bar{X}_n \overset{approx}{\sim} N\left(\mu, \frac{\sigma^2}{n}\right)$

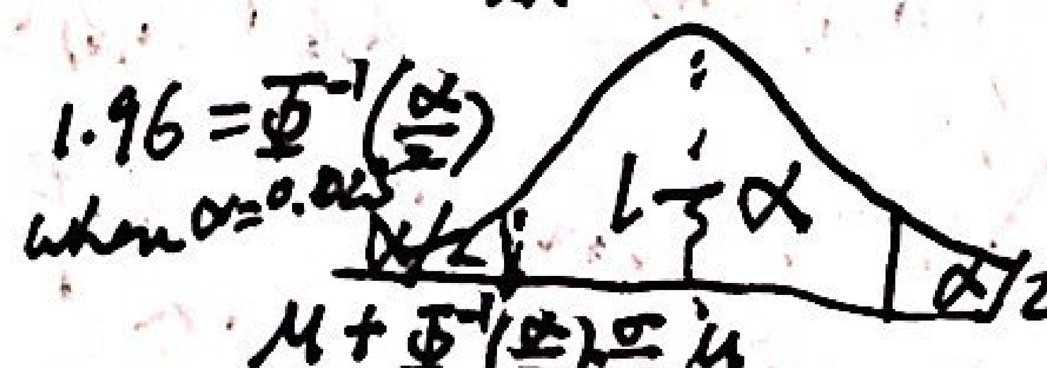
$$\mathbb{P}\left(\underbrace{\mu - 1.96\frac{\sigma}{\sqrt{n}} \le \bar{X}_n \le \mu + 1.96\frac{\sigma}{\sqrt{n}}}\right) = 0.95$$



$\Updownarrow$ put μ in the middle

& boundaries are random, irrelevant to μ

$$\mathbb{P}\left(\bar{X}_n - 1.96\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X}_n + 1.96\frac{\sigma}{\sqrt{n}}\right) = 0.95$$



$(1-\alpha)$ CI

$\alpha = 0.05$ or 0.1

So $\left[\bar{X}_n - 1.96\frac{\sigma}{\sqrt{n}}, \bar{X}_n + 1.96\frac{\sigma}{\sqrt{n}}\right]$ is a random interval that covers μ w/ 95% probability

so it is a 95% CI.

when σ is unknown, use $\hat{\sigma}$ as plug-in:

$$\left[\bar{X}_n - 1.96\frac{\hat{\sigma}}{\sqrt{n}}, \bar{X}_n + 1.96\frac{\hat{\sigma}}{\sqrt{n}}\right]$$

a good $\hat{\sigma}$ should be very close to σ when n is large — "consistency"

Say given a set of realizations : $x_1 = 0.5$, ..., $x_n = 1$

the 95% CI has realization $[0.25, 0.75]$. What can you say about μ?

mistake 1: μ falls into $[0.25, 0.75]$ with 0.95 probability.

Wrong because μ and $[0.25, 0.75]$ are both non-random.

so $\mu \in [0.25, 0.75]$ is NOT a random event in frequentist statistics.

mistake 2: 95% is about $[0.25, 0.75]$

Wrong because 95% is about a "random" interval $\left[\bar{X}_n - 1.96\frac{\sigma}{\sqrt{n}}, \bar{X}_n + 1.96\frac{\sigma}{\sqrt{n}}\right]$ and $[0.25, 0.75]$ is just one realization of it.

Analogy : $X \sim N(0,1)$ with one realization $x = 0.5$

V : X falls in the interval $[-1.96, 1.96]$ w/ 95% probability.

X : 0.5 falls ... ...

More generally, $X_1, \cdots, X_n \overset{iid}{\sim}$ pdf $f(\cdot;\theta)$

$$\text{MLE} : \hat{\theta}_{MLE} = \underset{\theta}{\arg\max} \underbrace{\prod_{i=1}^{n} f(X_i;\theta)}_{\substack{\| \\ \text{likelihood } L(\theta)}} = \underset{\theta}{\arg\max} \underbrace{\sum_{i=1}^{n} \log f(X_i;\theta)}_{\substack{\| \\ \text{log-likelihood } \ell(\theta)}}$$

$$I(\theta) = \underset{X_1}{E}\left[\left(\frac{d}{d\theta}\log f(X_1;\theta)\right)^2\right] = -\underset{X_1}{E}\left[\frac{d^2}{d\theta^2}\log f(X_1;\theta)\right]$$

$\hat{\theta}_{MLE} \overset{approx}{\sim} N\left(\theta, \frac{1}{nI(\theta)}\right)$ when n is large

$$\mathbb{P}\left(\theta - \frac{1.96}{\sqrt{nI(\theta)}} \le \hat{\theta}_{MLE} \le \theta + \frac{1.96}{\sqrt{nI(\theta)}}\right) = 0.95$$

$$\Updownarrow$$

$$\mathbb{P}\left(\hat{\theta}_{MLE} - \frac{1.96}{\sqrt{nI(\theta)}} \le \theta \le \hat{\theta}_{MLE} + \frac{1.96}{\sqrt{nI(\theta)}}\right) = 0.95$$

$\downarrow \hat{\theta}$ plug-in $\hat{\theta}_{MLE} \approx \theta$ $\downarrow \hat{\theta}$

$$\mathbb{P}\left(\hat{\theta}_{MLE} - \frac{1.96}{\sqrt{nI(\hat{\theta}_{MLE})}} \le \theta \le \hat{\theta}_{MLE} + \frac{1.96}{\sqrt{nI(\hat{\theta}_{MLE})}}\right) \approx 0.95$$

so $\left[\hat{\theta}_{MLE} - \frac{1.96}{\sqrt{nI(\hat{\theta}_{MLE})}}, \hat{\theta} + \frac{1.96}{\sqrt{nI(\hat{\theta}_{MLE})}}\right]$ is a 95% CI of θ.