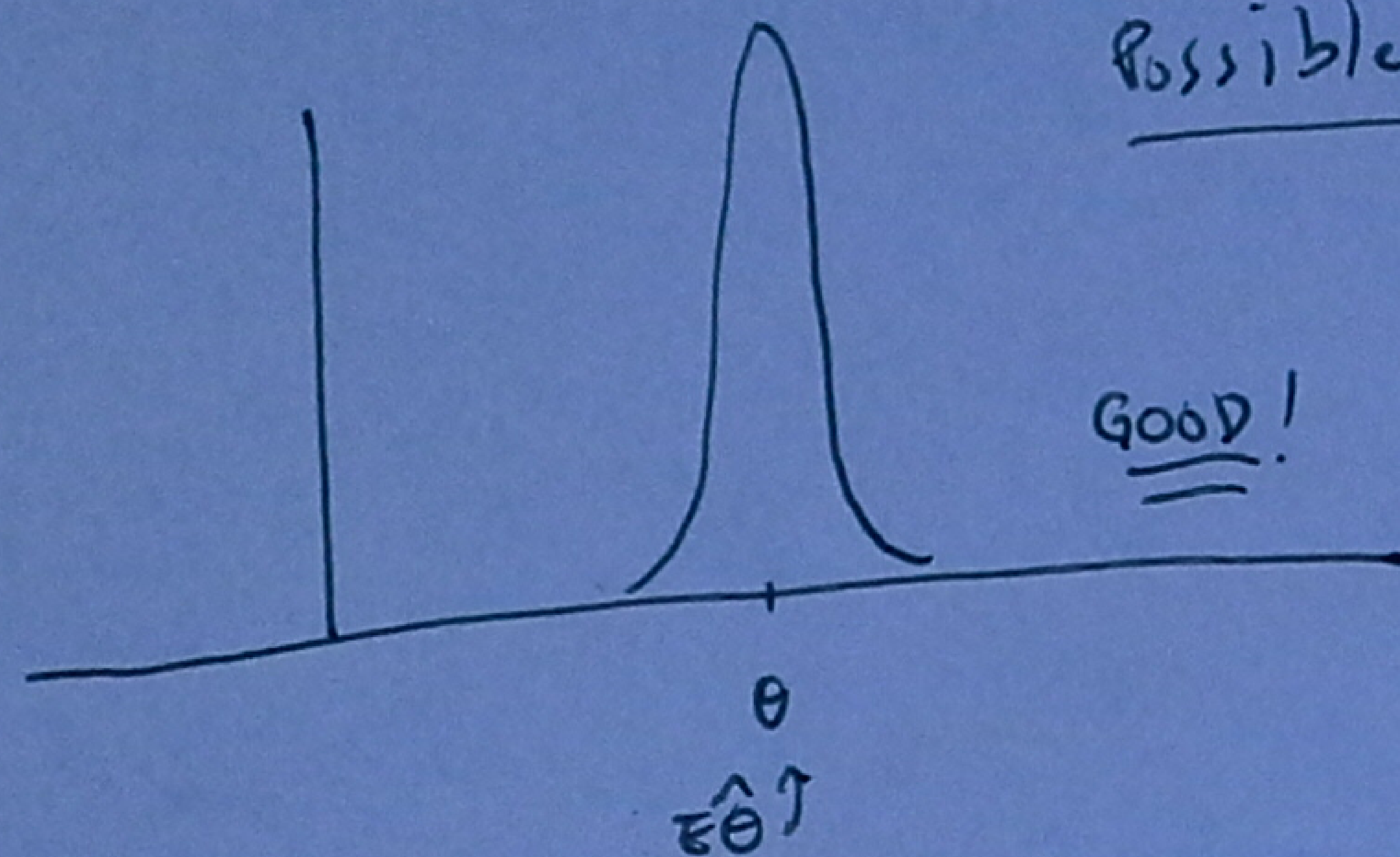
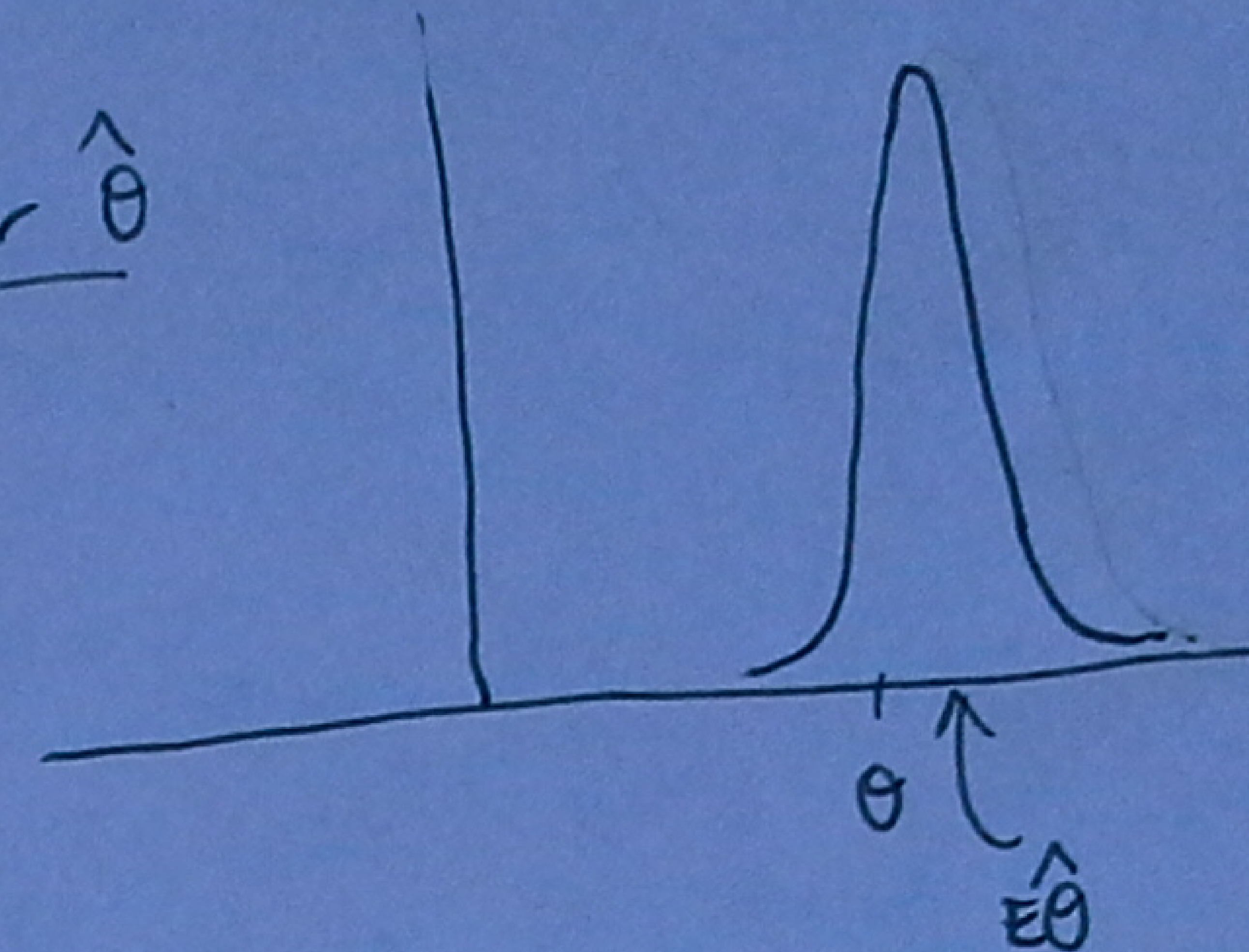
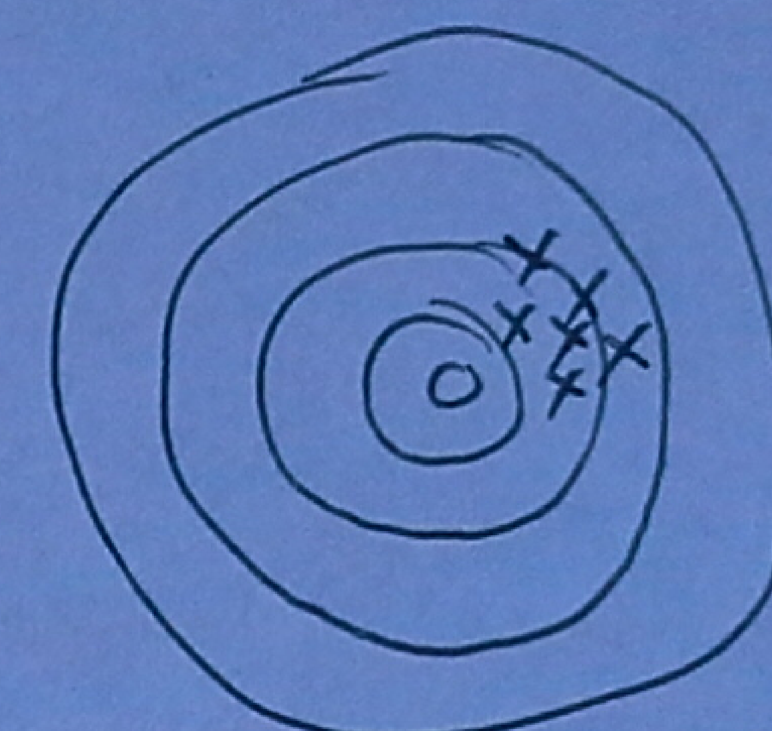
Statistic $\hat{\theta}$ used to estimate θ

Possible Distributions for $\hat{\theta}$





$$\boxed{\text{Bias}(\hat{\theta}) = E\hat{\theta} - \theta}$$



systematic error = BIAS

Ideally $\underbrace{\text{Bias}(\hat{\theta}) = 0}$

$\hat{\theta}$ is unbiased for θ

example $X_1, \dots, X_n \sim \text{iid } N(\theta, \sigma^2)$

$$E\left[\sum_{i=1}^{n}(X_i - \bar{X})^2\right] = \dots = (n-1)\sigma^2$$

details in book

$$\Rightarrow \hat{\sigma}^2 = \overbrace{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}^{\text{sample variance}} \text{ is } \underline{\text{unbiased}} \text{ for } \sigma^2$$

$\hat{\sigma}^2_{MLE} = \frac{1}{n}\sum(X_i - \bar{X})^2$ has a small bias

$$E\hat{\sigma}^2 = \sigma^2$$

$$E\hat{\sigma}^2_{MLE} = \left(1 - \frac{1}{n}\right)\sigma^2 = \sigma^2 - \underbrace{\boxed{\frac{\sigma^2}{n}}}_{\text{BIAS}}$$

$$\hat{\sigma}^2_{MLE} = \frac{1}{n}\frac{n-1}{n-1}\sum(X_i - \bar{X})^2 = \frac{n-1}{n}\hat{\sigma}^2$$

$$\left[\ \hat{\sigma}^2_{MLE} = \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2 = \frac{1}{n}\sum_i X_i^2 - \left(\frac{1}{n}\sum_i X_i\right)^2\right.$$

$$\boxed{X_i \sim iid(\mu, \sigma^2)}$$

$$E\left(\frac{1}{n}\sum_{i=1}^{n}X_i^2\right) = \frac{1}{n}\sum_{i=1}^{n} E X_i^2 = \sigma^2+\mu^2$$

$$\sigma^2 = E(X_i-\mu)^2 = E X_i^2 - \mu^2$$

$$\Rightarrow E X_i^2 = \sigma^2 + \mu^2$$

$$E\ \frac{1}{n^2}\left(\sum X_i\right)^2 = \frac{1}{n^2} E \sum_{i=1}^{n} X_i \sum_{j=1}^{n} X_j$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} E X_i X_j$$

$$E X_i X_j = \begin{cases} \sigma^2+\mu^2 & \text{if } i=j \\ \mu^2 & \text{if } i\neq j \end{cases}$$

$$= \frac{1}{n^2}\left[n^2\mu^2 + n\sigma^2\right] = \mu^2 + \frac{\sigma^2}{n}$$

$$E\hat{\sigma}^2_{MLE} = \sigma^2+\mu^2 - \left[\mu^2 + \frac{\sigma^2}{n}\right] = \sigma^2 - \frac{\sigma^2}{n}$$

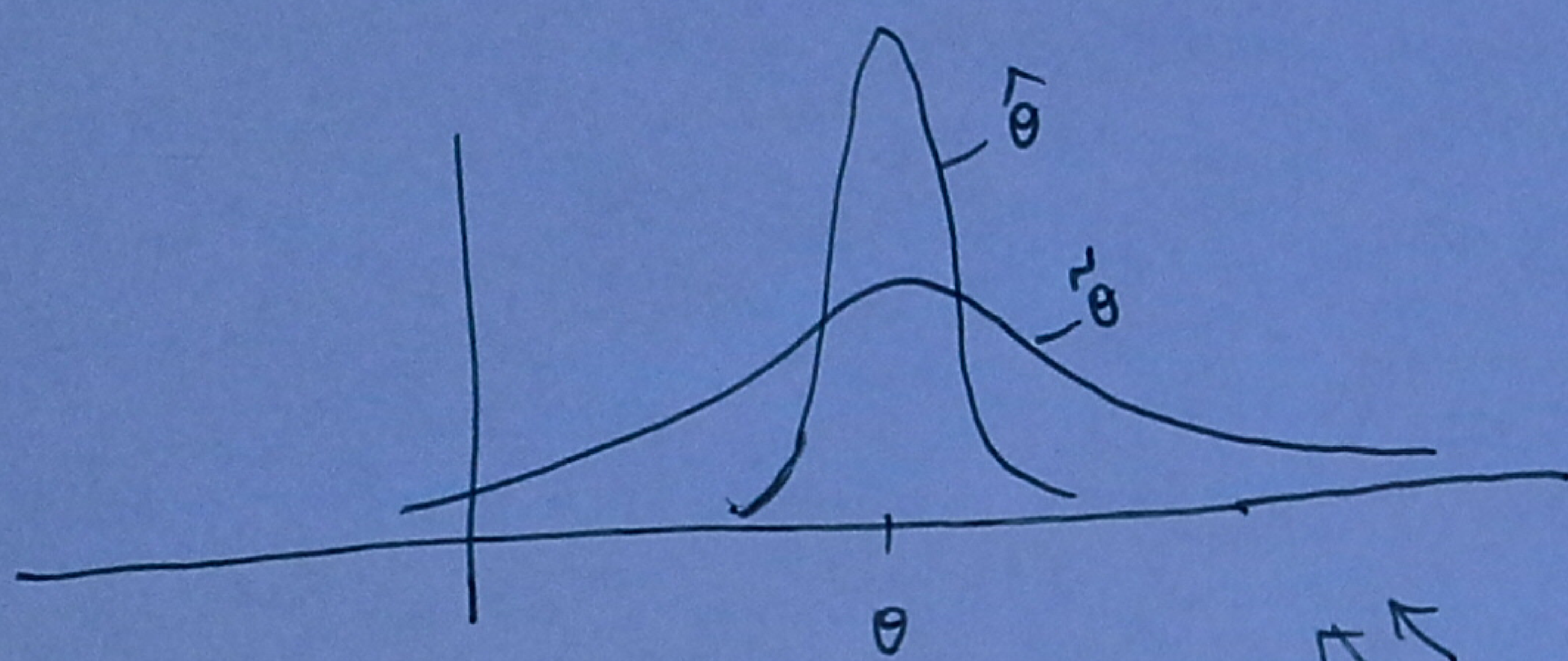
$\hat{\sigma}^2_{MLE}$ is <u>biased</u> but the $\underbrace{\text{bias} = -\frac{\sigma^2}{n} \longrightarrow 0 \text{ as } n\to\infty}$

($\hat{\sigma}^2_{MLE}$ asymptotically unbiased)

Def $\hat{\theta}$ is unbiased if $Bias(\hat{\theta}) = E\hat{\theta} - \theta = 0$

$\hat{\theta}$ is asymptotically unbiased if $Bias(\hat{\theta}) \to 0$ as $n \to \infty$

ex $\hat{\sigma}^2_{MLE}$ is asymptotically unbiased.



$\underbrace{\hat{\theta} - \theta = \text{error in estimation}}_{\text{Ideally small}}$

prefer the estimator with smaller variance (if the two biases are the same)

- Estimators $\hat{\theta}$ & $\tilde{\theta}$
- $Bias(\hat{\theta}) \approx Bias(\tilde{\theta})$
- If $var(\hat{\theta}) < var(\tilde{\theta})$, then $\hat{\theta}$ is called more efficient than $\tilde{\theta}$

Theorem [CRAMER-RAO lower bound] If $\hat{\theta}$ is unbiased, then $var\,\hat{\theta} \geq$ CR lower bd.

Def The Minimum Variance unbiased estimator (AKA "most efficient" unbiased) is an estimator that is unbiased with variance equal to the CR lower bd. (MVUE)

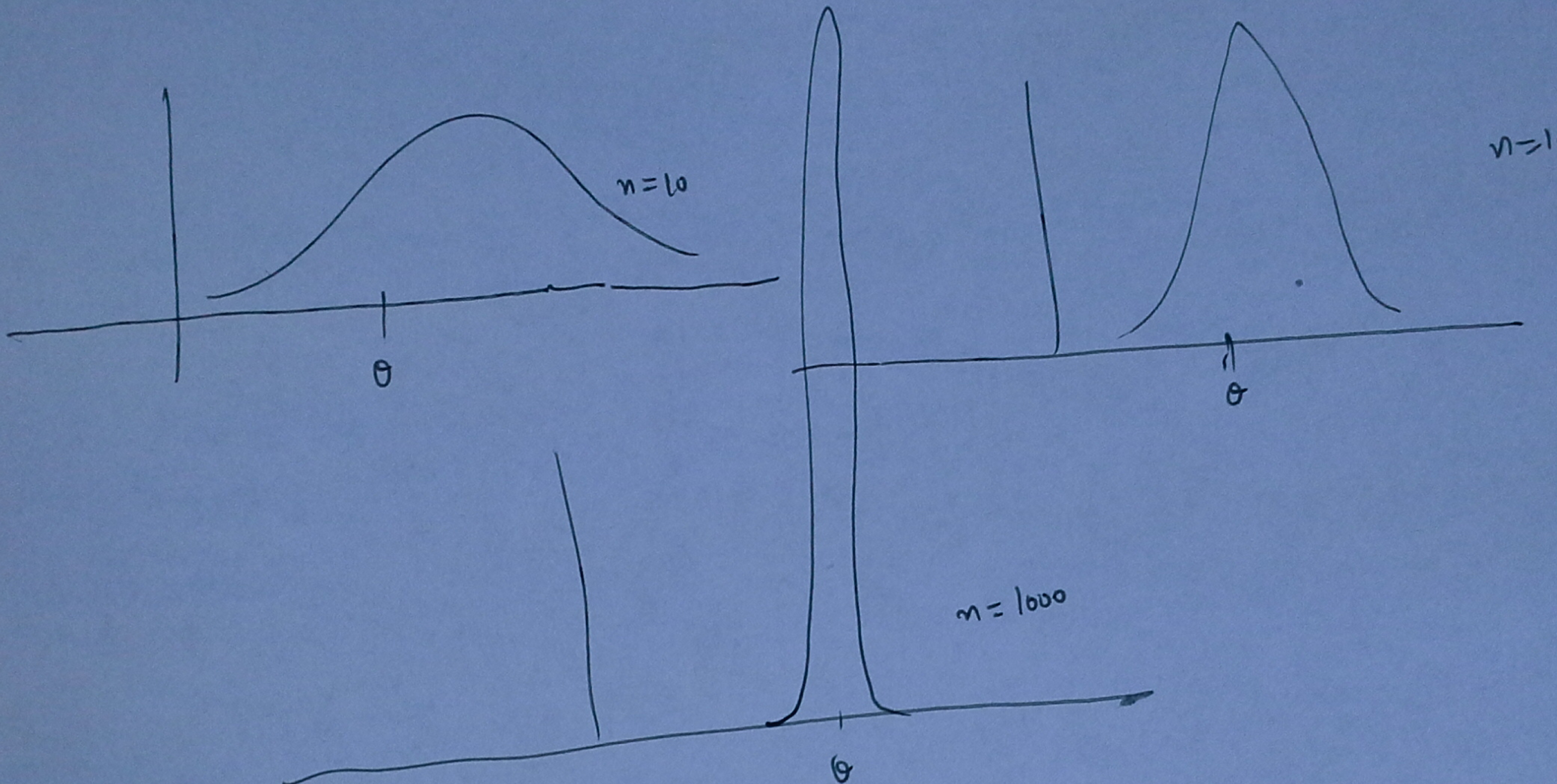
$\hat{\theta} - \theta$ = error

$$MSE = E(\hat{\theta}-\theta)^2 = E\left(\underbrace{\hat{\theta} - E\hat{\theta}} + \underbrace{E\hat{\theta} - \theta}\right)^2 = E\left[(\hat{\theta}-E\hat{\theta})^2 - 2(\hat{\theta}-E\hat{\theta})(E\hat{\theta}-\theta) + (\underbrace{E\hat{\theta}-\theta})^2\right]$$

Mean Squared Error

$$= \underbrace{E(\hat{\theta}-E\hat{\theta})^2}_{\operatorname{var}\hat{\theta}} \underbrace{- 2E(\hat{\theta}-E\hat{\theta})(\underbrace{E\hat{\theta}-\theta}_{ct.})}_{= -2(E\hat{\theta}-\theta)\,\underbrace{E[\hat{\theta}-E\hat{\theta}]}_{=0}} + \cancel{E}\,\underbrace{(E\hat{\theta}-\theta)^2}_{=[\mathrm{Bias}\,\hat{\theta}]^2}$$

$$MSE = E(\hat{\theta}-\theta)^2 = \operatorname{var}\hat{\theta} + [\mathrm{Bias}\,\hat{\theta}]^2$$



Def $\hat{\theta}$ is consistent for θ if $\underbrace{\hat{\theta} \xrightarrow{P} \theta \text{ as } n \to \infty}$.

For any $\delta > 0$; $P(|\hat{\theta} - \theta| < \delta) \longrightarrow 1$ as $n \to \infty$

my tolerance level

If MSE $\to 0$ as $n \to \infty$, then Chebyshev's inequality shows consistency!

Def An estimator $T = t(X_1, \dots, X_n)$ is called <u>sufficient</u> for θ if the likelihood function $L(\theta)$ factors into a product

$$L(\theta) = \underbrace{g(t(X_1,\dots,X_n);\theta)}_{\substack{\text{function of } \theta \\ \text{and } t(X_1,\dots,X_n)}} \cdot \underbrace{b(X_1,\dots,X_n)}_{\text{function of } X_1,\dots,X_n \text{ but not } \theta}$$

→|| For MLE all we need to know is the value of $t(X_1,\dots,X_n)$ ||←

ex. $X_1, \dots, X_n \sim$ iid Bernoulli (p) with $\theta = p$

data 001110 ...

$$L(\theta) = (1-p)(1-p)ppp(1-p)\cdots$$

$$= \theta^X (1-\theta)^{n-X}$$

$$X = \sum_1^n X_i$$

$$\boxed{X \text{ is sufficient for } \theta}$$

X is $Bin(n,p)$ so $P(X=x) = \binom{n}{x}\theta^x(1-\theta)^{n-x}$

density of sufficient st.

$$L(\theta) = \boxed{\binom{n}{X}\theta^X(1-\theta)^{n-X}} \cdot \boxed{\frac{1}{\binom{n}{X}}}$$

not depending on θ

$X_1, \dots, X_n \sim \text{iid } N(\theta, 1)$

$$L(\theta) = \prod_1^n \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{(x_i-\theta)^2}{2}\right\} = \frac{1}{(2\pi)^{n/2}} \exp\left\{-\frac{1}{2}\sum_1^n (x_i-\theta)^2\right\}$$

$$= \frac{1}{(2\pi)^{n/2}} \exp\left\{-\frac{1}{2}\left[\sum_1^n x_i^2 - 2\theta\sum_1^n x_i + n\theta^2\right]\right\}$$

$$= \frac{1}{(2\pi)^{n/2}} \exp\left\{-\frac{1}{2}\left[-2\theta\sum_1^n x_i + n\theta^2\right]\right\} \cdot \exp\left\{-\frac{1}{2}\sum x_i^2\right\}$$

$Y = \sum_1^n X_i$ is <u>sufficient</u> for θ, $\bar{X} = \frac{1}{n}\sum_1^n X_i$ is also <u>sufficient</u> for θ

$EY = E\sum_1^n X_i = \sum_1^n EX_i = n\theta$

$X_1, \dots, X_n \sim \text{iid } N(0, \sigma^2)$ $\qquad \theta = \sigma^2$

$$L(\theta) = \prod_1^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{x_i^2}{2\sigma^2}\right\} = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{-\frac{\sum_1^n x_i^2}{2\sigma^2}\right\}$$

$Z = \sum_1^n X_i^2$ is <u>sufficient</u> (for σ^2)