3 ESTIMATORS

Let a population: A random variable-X with distribution-$f(X)$ and population size $N \sim \infty$.

In general, the distribution function is either unknown or with known μ or σ^2.

We need to learn important parameters of the population: $\mu \equiv E[X]$, $\sigma^2 \equiv var[X]$, others...

To do that we take a sample $n \ll N$ and use it to infer (estimate) the populations parameters.

Take a random sample $\{X_1, \dots, X_n\}$ s.t. $\left.\begin{array}{l} \to X_i: \text{independent to each other} \\ \to X_i: \text{identically distributed-}f \end{array}\right\}$ idd

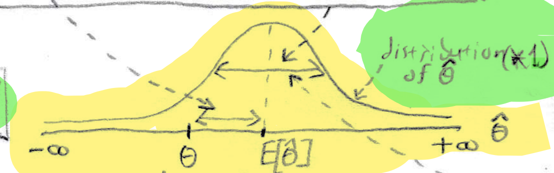
To estimate a parameter-θ of $f(X)$ $\{$such as $\mu \equiv E[X]$, $\sigma^2 \equiv var[X] = E[(X-\mu)^2]$, ...$\}$

define an estimator-$\hat{\theta} = \hat{\theta}(X_1, \dots, X_n)$ on sample with function $\hat{\theta}(X_1, \dots, X_n)$ appropriately defined (*2)

- $\hat{\theta} \neq \theta$ (true value) $\left\{\begin{array}{l} \hat{\theta} = \hat{\theta}(X_1, \dots, X_n) \text{ depends} \\ \text{on sample and takes} \\ \text{different value for} \\ \text{each new } n\text{-sample} \end{array}\right\} \to$ take virtually ∞-times n-samples and calculate
  - $\to E[\hat{\theta}] \equiv E[\hat{\theta}(X_1, \dots, X_n)] \overset{idd}{=}$ function of $E[X_i], E[X_i^2], \dots$ (*3)
  - $\to var[\hat{\theta}] \equiv E[(\hat{\theta} - E[\hat{\theta}])^2] \overset{idd}{=}$ function of $E[X_i], E[X_i^2], \dots$
- $\hat{\theta}$: good estimator if ideally $\to E[\hat{\theta}] \equiv \theta$ AND $\to V_{\hat{\theta}} = var[\hat{\theta}] \equiv 0$. So we define: $\boxed{\text{bias}(\hat{\theta}) \equiv E[\hat{\theta}] - \theta, \quad \text{sampling variance}: var[\hat{\theta}]}$

and the standard measure of accuracy of the estimator

$$\boxed{\text{mean squared error}: \ mse[\hat{\theta}] \equiv E[(\hat{\theta} - \theta)^2] = \dots = [\text{bias}(\hat{\theta})]^2 + var[\hat{\theta}]}$$



(*1) The estimator-$\hat{\theta}$ follows its own (unknown) distribution with $E[\hat{\theta}]$ and $var[\hat{\theta}]$

(*2) We seek to define $\hat{\theta}$ such that the $mse[\hat{\theta}]$ is minimum ($E[\hat{\theta}] \to \theta$ and $var[\hat{\theta}] \to 0$)

(*3) In general both $E[\hat{\theta}]$ and $V_{\hat{\theta}} = var[\hat{\theta}]$ are calculable functions of n, μ, σ^2, other parameters of $f(X)$.
If we know the true μ, σ^2 we simply plug them in the known function to get a numerical result.
If we don't know μ, σ^2 we use their estimates $\hat{\mu}, \hat{\sigma}^2$ to define the estimator of the function.
In particular,
Say $V_{\hat{\theta}} \equiv var[\hat{\theta}] = g(\mu, \sigma^2, n, \text{other parameters})$ $\to$ define appropriate $\hat{\mu}, \hat{\sigma}^2$ $\to$ plug $\hat{\mu}, \hat{\sigma}^2$ in $g(\quad)$ to define the : estimator of $V_{\hat{\theta}} \equiv \hat{V}_{\hat{\theta}} = g(\hat{\mu}, \hat{\sigma}^2, n, \dots) \Rightarrow$ $\boxed{\begin{array}{c}\text{Standard error of } \hat{\theta} \\ S(\hat{\theta}) \equiv \hat{V}_{\hat{\theta}}^{1/2}\end{array}}$

! A standard recipe for defining a good estimator is to define the analog estimator-$\hat{\theta}$ on sample. For example
- $\to$ for $\mu \left(= \frac{1}{N} \sum_1^N X_i\right)$ define the $\hat{\mu} \equiv \bar{X}_n \equiv \frac{1}{n} \sum_1^n X_i$
- $\to$ for $\sigma^2 \left(= \frac{1}{N} \sum (X_i - \mu)^2\right)$ define the $\hat{\sigma}^2 = \frac{1}{n} \sum_1^n (X_i - \hat{\mu})^2 \equiv \frac{1}{n} \sum X_i^2 - \hat{\mu}^2$