Estimation

Saturday, 2 December 2023 1:50 PM

Population parameters	Sample statistics
μ	$\bar{x}$
σ^2	s^2
p	$\hat{p}$...

- We would like to estimate μ & σ^2 (or any other population parameters) in terms of sample statistics.

- We can have

Point estimates

For example $\bar{x} = \sum_{i=1}^{n} \frac{x_i}{n}$ can be a point estimation of the population mean.

Interval estimates. (Here we would like to find an interval where our population parameter will belong).

- Let us first talk about point estimators.

Unbiasedness :- A sample statistics $\hat{\Theta}$ is said to be an unbiased estimator of a population parameter θ if

$$E(\hat{\Theta}) = \theta$$

For example s^2 is an unbiased estimator of σ^2.

How?

$$s^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{(n-1)}$$

$$(n-1)s^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2$$

$$= \sum_{i=1}^{n} (x_i - \bar{x} + \mu - \mu)^2 \quad \text{(adding \& subtracting } \mu)$$

$$= \sum_{i=1}^{n} \left((x_i - \mu) - (\bar{x} - \mu)\right)^2$$

$$= \sum_{i=1}^{n} (x_i - \mu)^2 + (\bar{x} - \mu)^2 - 2(x_i - \mu)(\bar{x} + \mu)$$

$$= \sum_{i=1}^{n} (x_i - \mu)^2 + \sum_{i=1}^{n} \underbrace{(\bar{x} - \mu)^2}_{\text{independent of } i} - 2\sum_{i=1}^{n} (x_i - \mu) \underbrace{(\bar{x} + \mu)}_{\text{independent of } i}$$

$$(n-1)s^2 = \sum_{i=1}^{n} (x_i - \mu)^2 + (\bar{x} - \mu)^2 \underbrace{\sum_{i=1}^{n} 1}_{=n} - 2(\bar{x} + \mu)\left(\underbrace{\sum_{i=1}^{n} x_i}_{=n\mu} - \mu \underbrace{\sum_{i=1}^{n} 1}_{=n}\right)$$

$$= \sum_{i=1}^{n} (x_i - \mu)^2 + n(\bar{x} - \mu)^2 - 2n(\bar{x} - \mu)^2$$

$$(n-1)s^2 = \sum_{i=1}^{n} (x_i - \mu)^2 - n(\bar{x} - \mu)^2$$

$$E\left((n-1)s^2\right) = E\left(\sum_{i=1}^{n} (x_i - \mu)^2 - n(\bar{x} - \mu)^2\right)$$

$$(n-1)E(s^2) = \sum_{i=1}^{n} E\left((x_i - \mu)^2\right) - n \sum_{i=1}^{n} E\left((\bar{x} - \mu)^2\right)$$

Variance of sampling distribution of mean

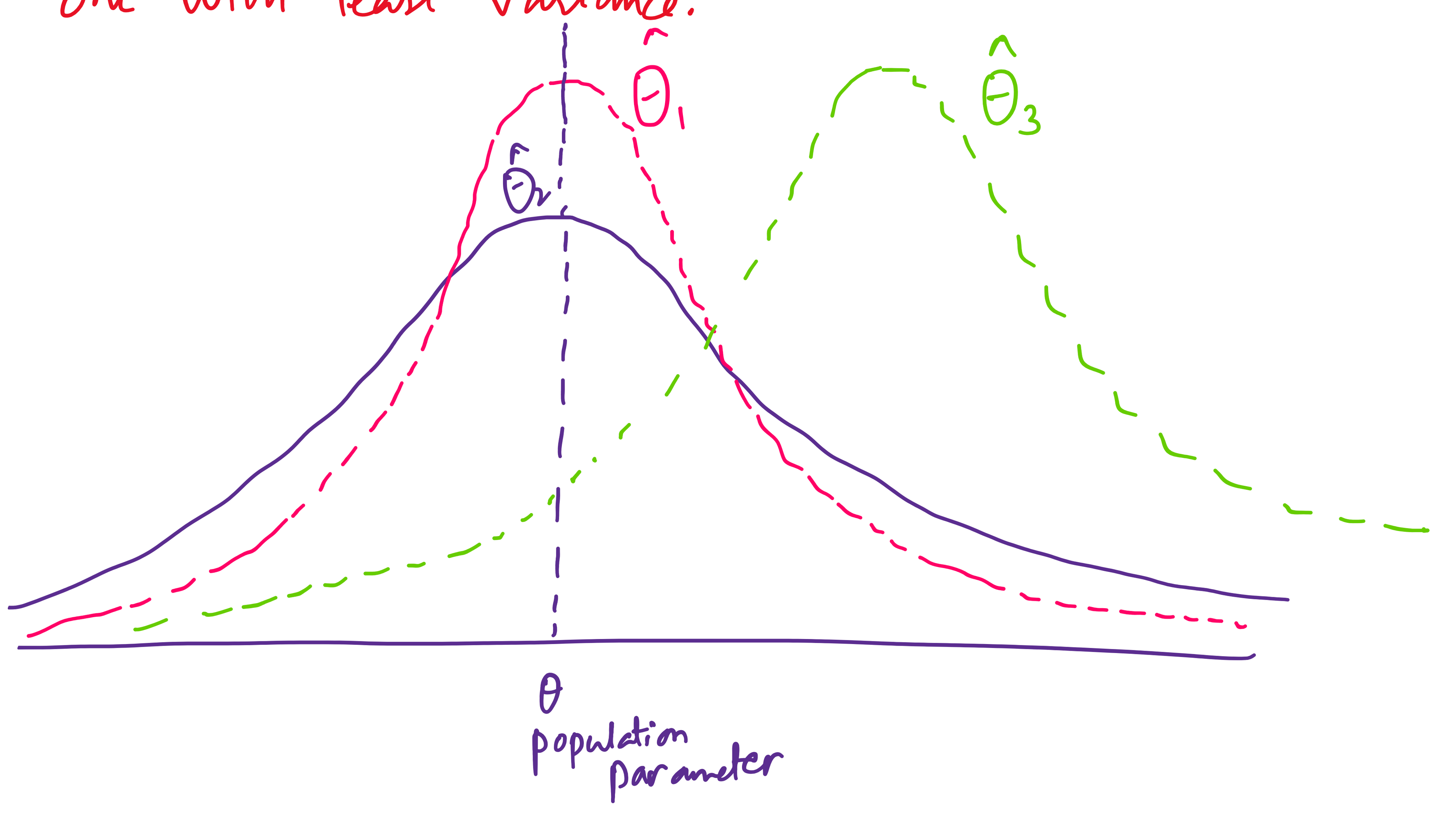
$$(n-1)E(s^2) = \sum_{i=1}^{n} \sigma^2 - n\left(\frac{\sigma^2}{n}\right)$$

$$(n-1)E(s^2) = n\sigma^2 - \sigma^2 = (n-1)\sigma^2$$

$$\Rightarrow \boxed{E(s^2) = \sigma^2}$$

Note

- Out of a biased and an unbiased estimator, we always choose the unbiased estimator.
- Out of many unbiased estimators available we choose the one with least variance.



- Out of $\hat{\Theta}_1$ $\hat{\Theta}_2$ $\hat{\Theta}_3$

$\hat{\Theta}_1$ & $\hat{\Theta}_2$ are unbiased & $\hat{\Theta}_3$ is biased, so we choose $\hat{\Theta}_1$ or $\hat{\Theta}_2$.

- Out of $\hat{\Theta}_1$ and $\hat{\Theta}_2$, $\hat{\Theta}_1$ has less variance so we choose $\hat{\Theta}_1$.