M339G: August 27th, 2025.

Some Math·Stats Highlights.

Def'n. A random sample from distribution D is a random vector $(X_1, X_2, \ldots, X_n)$ such that

- $X_1, \ldots, X_n$ are independent, and
- $X_i \sim D$ for all $i=1,\ldots,n$.

Def'n. A statistic is a function of the random sample.

Def'n. An estimator for a parameter θ is a statistic $\hat{\theta}$ that doesn't depend on θ and is used to estimate θ.

Def'n. We say that the estimator $\hat{\theta}$ is unbiased if

$$\boxed{\mathbb{E}[\hat{\theta}] = \theta}$$

In general, the bias of the estimator $\hat{\theta}$ is

$$\boxed{bias(\hat{\theta}) := \mathbb{E}[\hat{\theta} - \theta]}$$

Def'n. The mean squared error of the estimator $\hat{\theta}$ is

$$\boxed{MSE(\hat{\theta}) := \mathbb{E}\left[(\hat{\theta} - \theta)^2\right]}$$

Confidence Intervals.

Let $X_1, \dots, X_n$ be a normal random sample, i.e.,

$\{X_i : i = 1..n\}$ are independent

and

$X_i \sim \text{Normal}(\text{mean} = \underset{?}{\mu}, \text{sd} = \underset{\text{know}}{\sigma})$

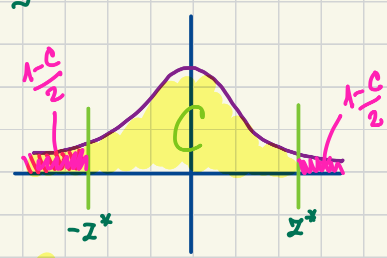
We know the sampling dist'n of the sample mean

$$\bar{X} = \frac{1}{n}(X_1 + \dots + X_n) \sim \text{Normal}\left(\text{mean} = \mu, \text{sd} = \frac{\sigma}{\sqrt{n}}\right)$$

We know that $\bar{X}$ is a "good" estimator for the population mean μ.

C... confidence level

$$\mathbb{P}\left[-z^* \le \underbrace{\overset{\sim N(0,1)}{\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}}}_{\text{Pivotal Quantity}} \le z^*\right] = C$$



$$\frac{1-C}{2} + C = \frac{1+C}{2}$$

$$\Phi(z^*) = \frac{1+C}{2}$$

$$z^* = \Phi^{-1}\left(\frac{1+C}{2}\right) = \text{qnorm}((1+C)/2)$$

$$\mathbb{P}\left[\bar{X} - z^* \cdot \frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X} + z^* \cdot \frac{\sigma}{\sqrt{n}}\right] = C$$

Confidence Interval

Q: If σ not known $\longrightarrow$ t. distribution.

Q: What if the sample is not normal ... ?