M358K: October 9th, 2023.

Confidence Intervals [The Normal Case].

> We are still in the normal model.
> The same logic will apply to other models as well.

Let $X_1, X_2, \dots, X_n$ be a normal random sample, i.e.,

$\{X_i, i=1..n\}$ are all independent, and

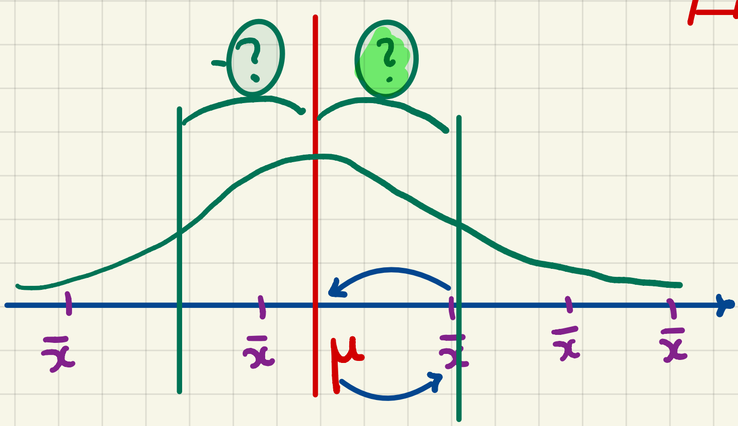
$$X_i \sim \text{Normal}(\text{mean} = \mu, \ \text{sd} = \sigma)$$

We know exactly the distribution of the sample mean:

$$\bar{X}_n \sim \underline{\text{Normal}}\left(\text{mean} = \underline{\mu}, \ \text{sd} = \underline{\frac{\sigma}{\sqrt{n}}}\right)$$

For now: assume that σ is known.

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



Q: How $\boxed{\text{CONFIDENT}}$ are we about the value that we get?

What does "confidence" even mean?

Let C be a large probability, i.e., a confidence level.

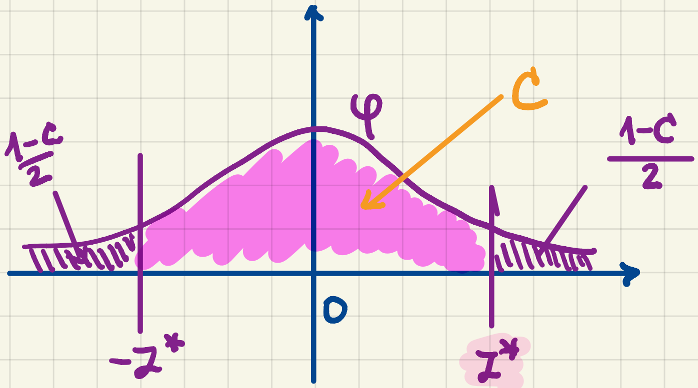
Say $C = 0.95, 0.90, 0.99, 0.80$

Look @

$$\mathbb{P}\left[\,|\bar{X} - \mu| < ?\,\right] = C$$

$$\mathbb{P}\left[-? < \bar{X} - \mu < ?\right] = C \quad ✓✓$$

$$\mathbb{P}\left[\underbrace{\frac{-?}{\frac{\sigma}{\sqrt{n}}}}_{= -z^*} < \underbrace{\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}}_{= Z \,\sim\, N(0,1)} < \underbrace{\frac{?}{\frac{\sigma}{\sqrt{n}}}}_{= z^*}\right] = C \quad ✓$$



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

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

z^* is the critical value of $N(0,1)$ such that

$$\boxed{\mathbb{P}[-z^* < Z < z^*] = C}$$

$$✓ \Rightarrow z^* = \frac{?}{\frac{\sigma}{\sqrt{n}}} \Rightarrow \boxed{? = z^* \frac{\sigma}{\sqrt{n}}}$$

(✓) $\Rightarrow \mathbb{P}\left[\boxed{-z^*\cdot\frac{\sigma}{\sqrt{n}}} < \bar{X} - \mu < \boxed{z^*\cdot\frac{\sigma}{\sqrt{n}}}\right] = C$

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

Random Interval

which we call our CONFIDENCE INTERVAL

- Every time that you collect a sample and construct a confidence interval, you will obtain a different interval.
- The confidence interval will contain the mean parameter μ with probability C, and w/ probability $1-C$, it will NOT!

In Practice:

For a particular data set: $x_1, x_2, \dots, x_n$,

we calculate $\boxed{\bar{x} = \frac{1}{n}(x_1 + \dots + x_n)}$ sample average

Then, we provide $C\cdot$ confidence interval:

$$\left(\bar{x} - z^*\cdot\frac{\sigma}{\sqrt{n}}\ ,\ \bar{x} + z^*\cdot\frac{\sigma}{\sqrt{n}}\right).$$

We usually write: $\boxed{\mu = \bar{x} \pm z^*\cdot\frac{\sigma}{\sqrt{n}}}$

In words: "We are $C\cdot$confident that the mean μ is within the interval...."

$\boxed{z^*\cdot\frac{\sigma}{\sqrt{n}}}$ is called the margin of error in this particular case

More generally, the margin of error is the radius of the confidence interval.

$$z^* \cdot \boxed{\frac{\sigma}{\sqrt{n}}}$$

the standard error (the standard deviation of the point estimator)

spread in the population

critical value which depends on C

sqrt(sample size)