M358K: October 11th, 2021.

Confidence Intervals.

While 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:

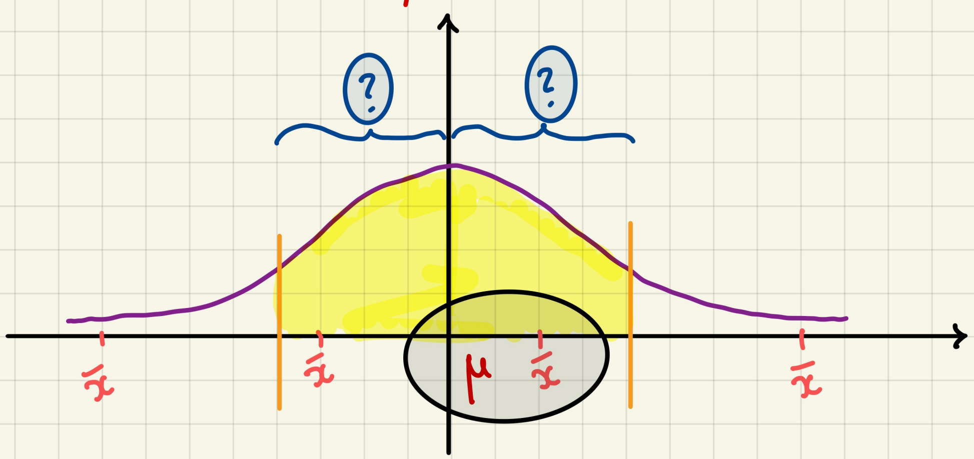
$$\bar{X}_n \sim \text{Normal}\left(\text{mean} = \mu, \boxed{\text{sd} = \frac{\sigma}{\sqrt{n}}}\right)$$?

For now: assume that the 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.

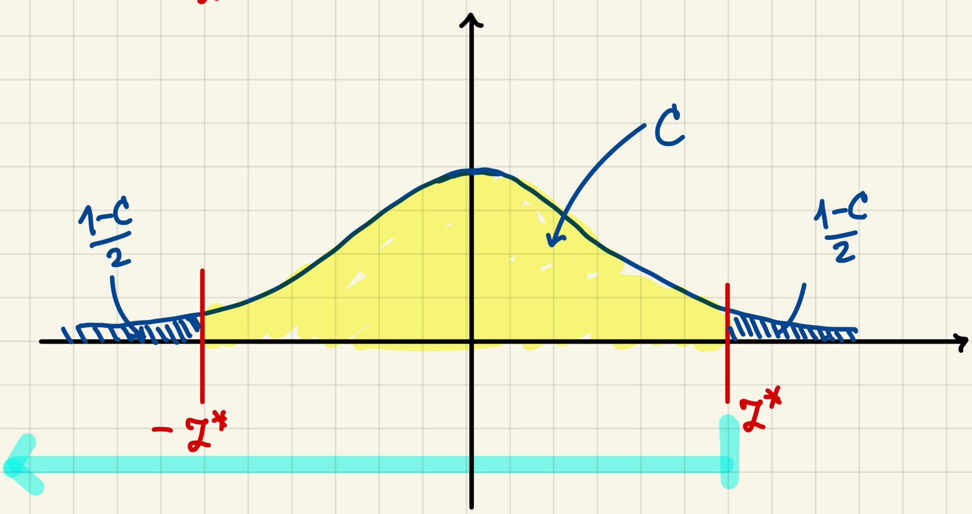
Say $C = 0.95, 0.90, 0.99, 0.80$.

Look @

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

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

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



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

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

z^* is the CRITICAL VALUE of $N(0,1)$ such that

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

$\checkmark \longrightarrow \quad z^* = \dfrac{?}{\frac{\sigma}{\sqrt{n}}} \implies ? = z^* \cdot \dfrac{\sigma}{\sqrt{n}}$

$\checkmark\checkmark \longrightarrow \quad \mathbb{P}\left[-z^* \cdot \frac{\sigma}{\sqrt{n}} < \bar{X}_n - \mu < z^* \cdot \frac{\sigma}{\sqrt{n}} \right] = C$

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

my RANDOM Interval
which we call our CONFIDENCE INTERVAL.

- Every time that you collect a sample and construct a confidence interval, you obtain a <u>DIFFERENT INTERVAL</u>.
- With the probability C, the interval will contain the mean parameter μ and w/ probability $1-C$ it will not.

For a particular data set $x_1, x_2, \dots, x_n$,
we calculate $\bar{x} = \frac{1}{n}(x_1 + x_2 + \dots + x_n)$.
Then, we provide a C·confidence interval:

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

We usually write:

$$\bar{x} \pm z^* \cdot \frac{\sigma}{\sqrt{n}}$$

w/ words : " we are $\boxed{C\cdot\text{confident}}$ that the mean μ is within the interval."

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

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

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

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

spread in the population

sample size

critical value which depends on the confidence level C