Sampling Distribution:

Population
μ, σ^2

Sample 1 of n — $\bar{x}_1$ | Sample 2 of n — $\bar{x}_2$ | Sample i, of n — $\bar{x}_i$

- population is normal $\Rightarrow \bar{X} \sim N(\mu, \frac{\sigma^2}{n})$
- population is not normal or sample size is > 30 (CLT)

  $\Rightarrow \bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

- $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$ = sampling distribution

  $\frac{\sigma}{\sqrt{n}}$: standard error.

- CLT: Central Limit Theorem states that a sampling distribution is normal.

- Statistical Inferences
  1. Estimation : use sample to predict popution's parametors
  2. Hypothesis tests : use sample to decide whether a statement is true about the population's parameters.
- Estimation :
  1. point estimates
  2. interval estimates

- standard error $SE = \frac{\sigma}{\sqrt{n}}$; σ : populative std
- Estimated standard error $\hat{SE} = \frac{s}{\sqrt{n}}$; s : sample std
- Confidence interval : Sample statistics $\pm$ margin of error
- Margin of error = $M \times \hat{SE}$

multiplier derived from signrficance level. (α)

- Confidence interval for population mean

① Known σ : $\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$

② unknown σ = $\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$

(t distribution has n-1 dof)

- Sample Size Computation :

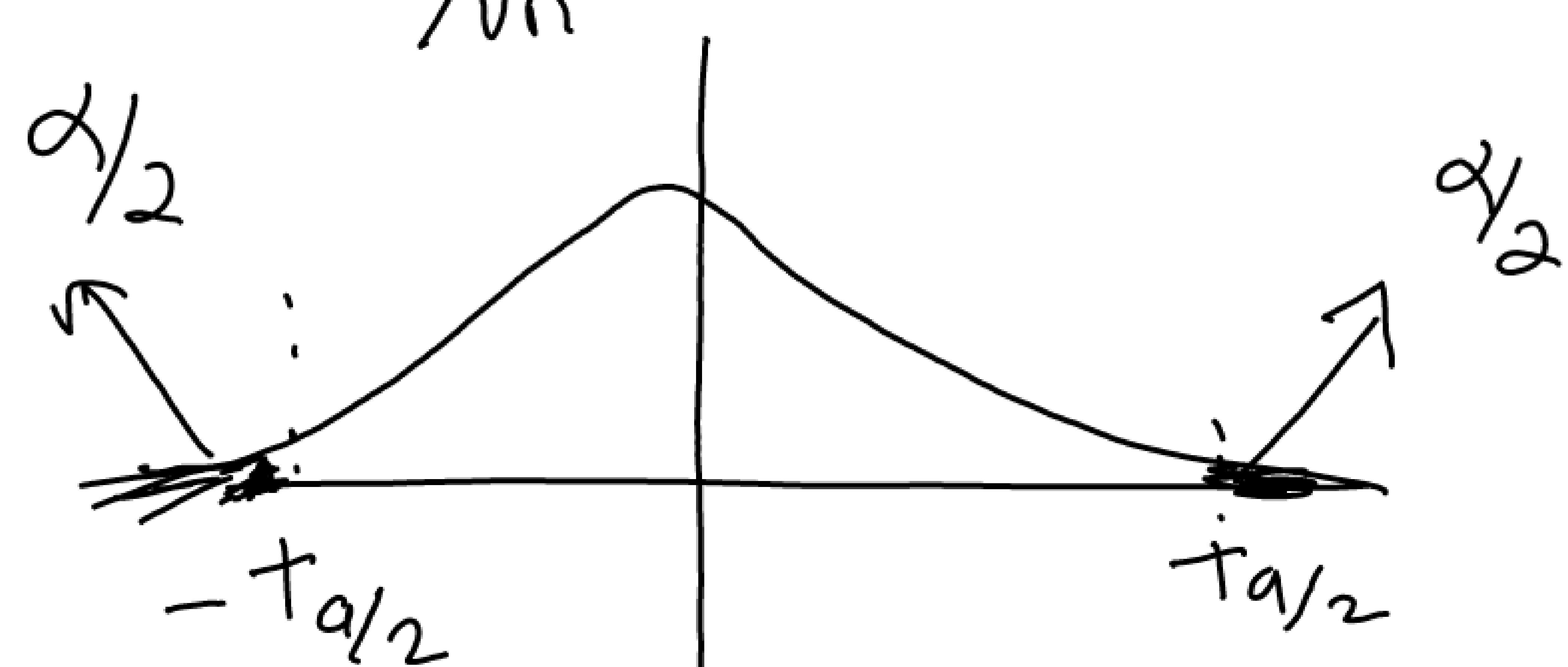
For a given margin of error (E):

$$E = t_{\alpha/2} \cdot \frac{s}{\sqrt{n}} \Rightarrow n = \left(\frac{t_{\alpha/2} s}{E} \right)^2$$

- One-sample mean test

① $H_0 : \mu = \mu_0$; $H_a : \mu \neq \mu_0$

② $t^* = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$; assuming H_0 is true



③ Reject H_0 if $t^* < -t_{\alpha/2}$ or $t^* > t_{\alpha/2}$

- Difference b/w 2 normal variables

$$X \sim N(\mu_1, \sigma_1^2) \; ; \; Y \sim N(\mu_2, \sigma_2^2)$$

$$X - Y \sim N(\mu_1 - \mu_2 ; \sigma_1^2 + \sigma_2^2)$$

- σ_1^2 and σ_2^2 are unknown

① pooled variance: $S_p = \sqrt{\dfrac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}}$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_p\sqrt{1/n_1 + 1/n_2}} \; ; \; DOF = n_1 + n_2 - 2$$

② unpooled variance:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{S_1^2/n_1 + S_2^2/n_2}} \; ; \; DOF = \frac{(n_1-1)(n_2-1)}{(n_1-1)c^2 + (1-c^2)(n_1-1)}$$

$$c = \frac{S_1^2/n_1}{S_1^2/n_1 + S_2^2/n_2}$$

- Comparing 2 population means:

① pooled variance: $CI = \bar{X}_1 - \bar{X}_2 \pm t_{\alpha/2} S_p \sqrt{1/n_1 + 1/n_2}$

② unpooled variance: $CI = \bar{X}_1 - \bar{X}_2 \pm t_{\alpha/2} \sqrt{S_1^2/n + S_2^2/n}$

- For pair samples, $x_1, x_2 \dots x_n$ and $y_1, y_2 \dots y_n$

$\Rightarrow d_1 = x_1 - x_2 \; ; \; d_2 = x_2 - x_2 \; \dots \; d_n = x_n - y_n$

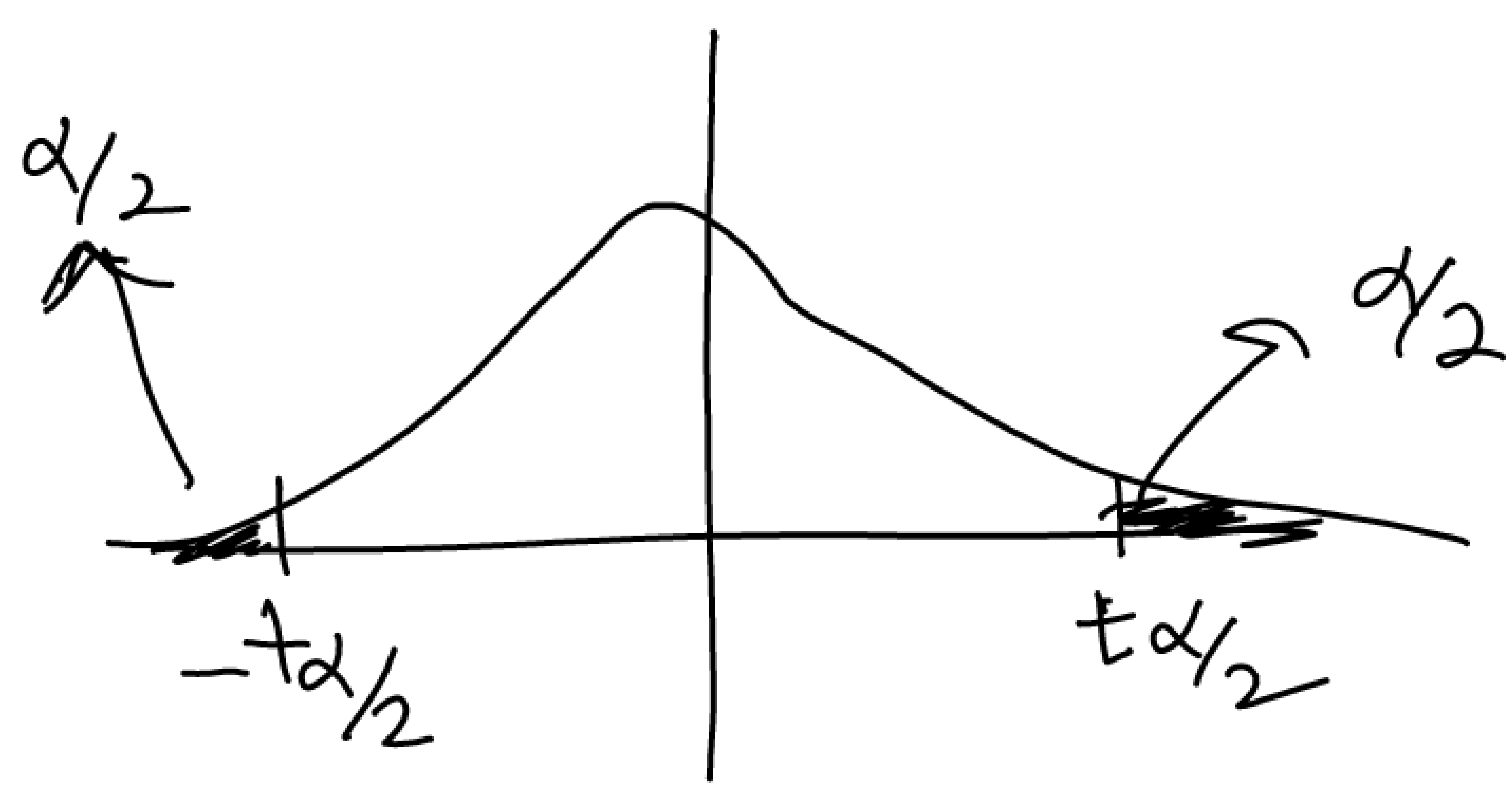
$$d \sim N(\mu_d, \sigma_d^2/n)$$

confidence interval is:

$$\bar{d} \pm t_{\alpha/2} \, s_d/\sqrt{n}$$

test statistic:

$$t^* = \frac{\bar{d}}{s_d/\sqrt{n}}$$



$H_0: \mu_1 = \mu_2$ or $\mu_d = 0$

$H_a: \mu_d \neq 0$