Roadmap

- Hypothesis Testing for Normal problems for μ and σ
- Binomial problem for n large.
- Relationship between C.I and H.T.

$$H_0 : \theta = \theta_0$$

θ = unknown parameter

θ_0 = suggested value of θ.

$\{y_1, \dots y_n\}$

Based on the sample,

we want to find evidence to support/against H_0.

Construct the Discrepancy Measure D

(i) $D \geq 0$

(ii) $D = 0$ best evidence for H_0

(iii) Larger ~~extre~~ values of D correspond to stronger evidence against H_0.

typical, but not necessary

(iv) $P(D \geq d)$ can be calculated,

i.e ~~X~~. the distribution of D is known.

· Calculate the value of D from your sample

d

· p-value : $P(D \geqslant d \,;\, H_0 \text{ is true})$

p-value "small" · might be because

(a) H_0 is true, and we observed a very rare event.

OR

(b) H_0 is not true.

Example: A random sample of 25 observations are drawn from a Gaussian population with mean μ and σ unknown

$$H_0 : \mu = 6 \leftarrow \mu_0$$

$\{y_1, \ldots y_{25}\}$ $\bar{y} = 8$ $s = 5$

$$\boxed{\frac{\bar{Y} - \mu}{\frac{S}{\sqrt{n}}} \sim T_{n-1}}$$

$$D = \left| \frac{\bar{Y} - \overset{6}{\mu_0}}{\frac{S}{\sqrt{n}}} \right|.$$

$$D = \left| \frac{\bar{Y} - \overset{\mu_0}{6}}{\frac{S}{\sqrt{n}}} \right|. \quad \ldots \ldots (1)$$

All the properties of D are satisfied by (1)

Calculate $d = \dfrac{\bar{y} - 6}{\frac{s}{\sqrt{n}}}$

$= \dfrac{8 - 6}{\frac{5}{\sqrt{25}}} = \dfrac{2}{1} = \underline{\underline{2}}$

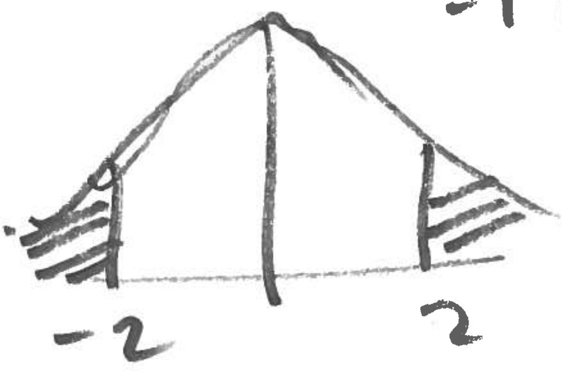
---

p-value = $P(D \geq d)$

$= P(D \geq 2)$

df = 24

$= P(|T_{24}| \geq 2)$

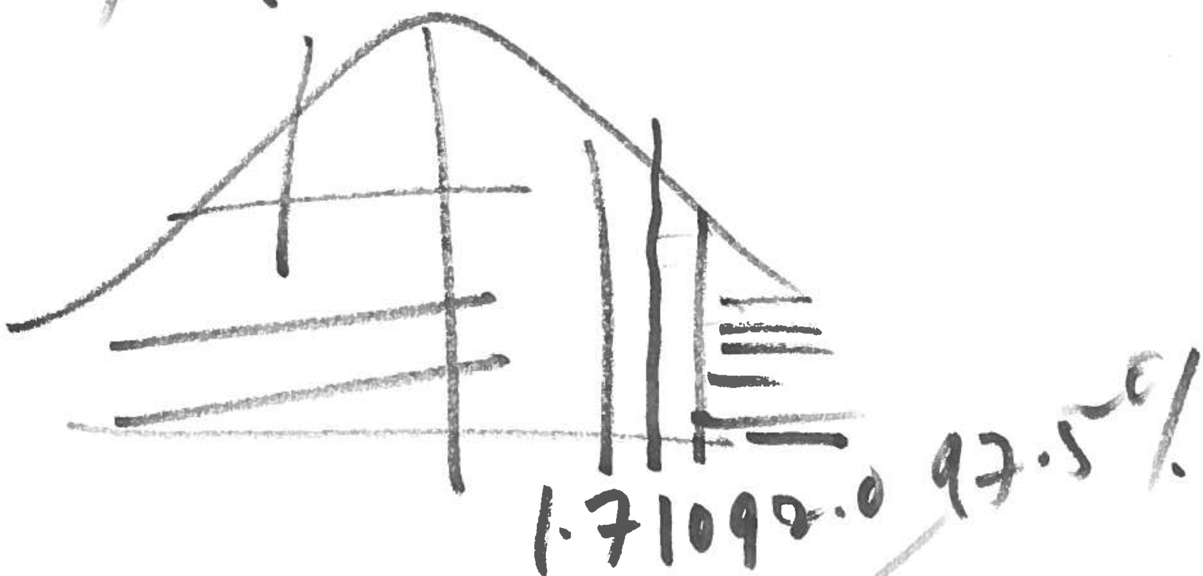


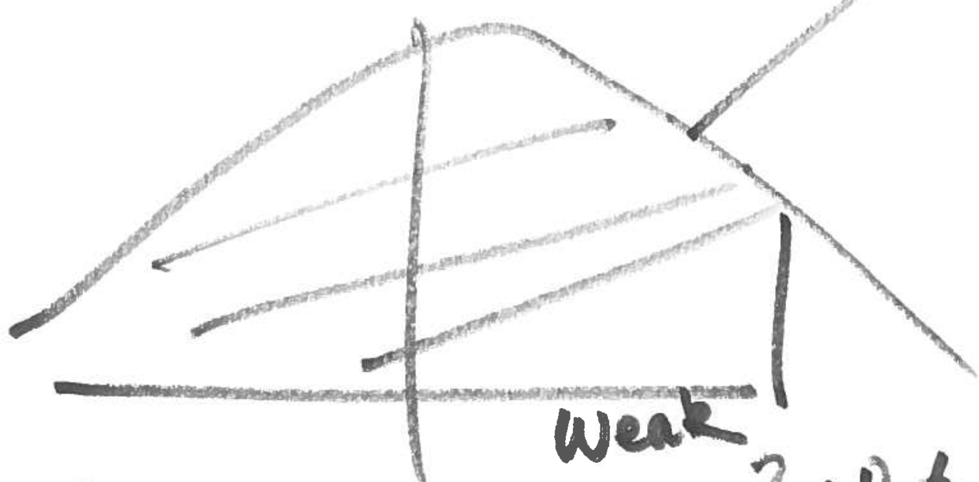
(i) Use R

(ii) Use the T-table → to find a range for the p-value

df = 24

$P(|T_{24}| \le 2)$





p-value > 0.05 ⇒ ~~D~~ not ~~enough~~ enough evidence against

Example: If σ was known, then

$$D = \left| \frac{\bar{Y} - \mu}{\frac{\sigma}{\sqrt{n}}} \right| .$$

$$\cancel{\frac{\bar{Y} - \mu}{\frac{\sigma}{\sqrt{n}}} = Z} .$$

Example: A sample of 1000 UW students are taken and 450 of them ~~are~~ are brown haired.

θ = proportion of Brown-Haired people in UW.

Test: $H_0 : \theta = 0.5$.

$$\underline{\text{CLT}}: \quad \frac{\tilde{\theta} - \theta}{\sqrt{\frac{\theta(1-\theta)}{n}}} = Z \sim G(0,1)$$

$H_0 : \theta = 0.5$

$$D = \left| \frac{\tilde{\theta} - 0.5}{\sqrt{\frac{(0.5)(1-0.5)}{n}}} \right|$$

$$d = \left| \frac{\hat{\theta} - 0.5}{\sqrt{\frac{0.5 \times 0.5}{1000}}} \right| = \underbrace{\left| \frac{0.45 - 0.5}{\sqrt{0.25/1000}} \right|}_{\underline{\underline{3}}\ \times}$$

p-value $= P(D \geq 3)$

$= P(|Z| \geq 3)$

Equivalence between C.I and H.T.

$Y_1, \ldots Y_n \sim G(\mu, \sigma) \quad \mu, \sigma$ unknown

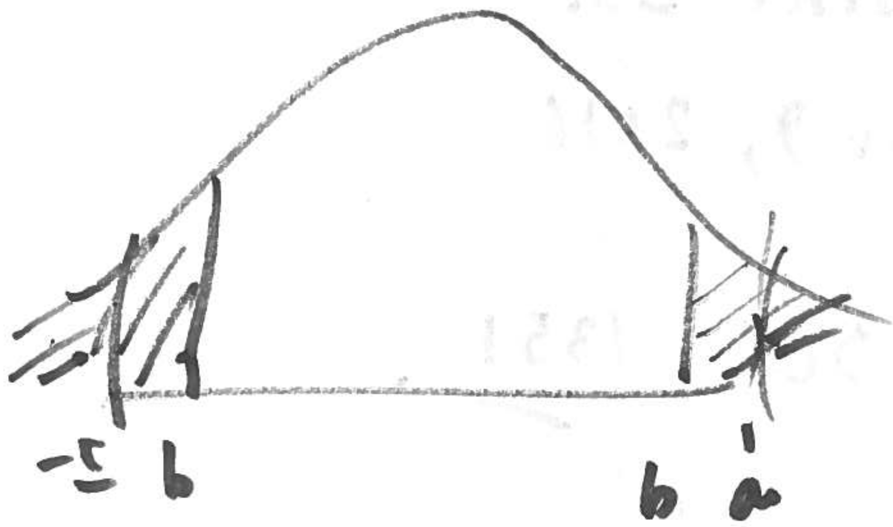
$$\left.\begin{array}{l} H_0: \mu = \mu_0 \\ H_1: \mu \neq \mu_0 \end{array}\right\}$$

p-value ≥ 0.05. Then it must be true that μ_0 is in the 95% C.I.

$$P(D \geq d) \geq 0.05$$

$$P\left(\left|\frac{\bar{Y} - \mu_0}{\frac{S}{\sqrt{n}}}\right| \geq \left|\frac{\bar{y} - \mu_0}{s/\sqrt{n}}\right|\right) \geq 0.05$$

$$= P(|T_{n-1}| \geq b) \geq 0.05$$



Shaded areas ≥ 0.05

$[-a, \quad a] = 95\%$ C.I.

$a > b \Rightarrow$ b must be in the $95^{\text{th}}\%$ interval for T_{n-1}