Date: Tuesday, March 26, 2019

Hypothesis test continued...

So far:

1) State a hypothesis in terms of population metric.

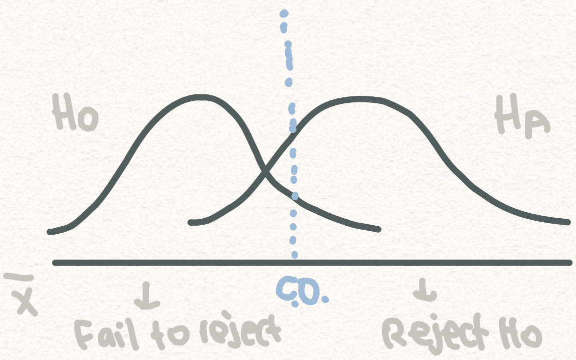
   Ex: $H_0: \mu = 0$ vs $H_A: \mu > 0$

2) Calculate sample statistic that summarizes the metric of interest (Ex. $\bar{x}$)

3) Define rejection region based on significance.

   Chosen by you.
   Represented as α.

   Example: p-value $\leq \alpha$



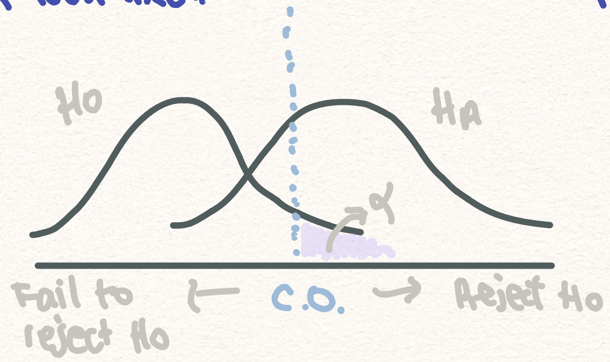
↳ We assume that the sample data has a probability distribution under both Ho and HA.

↳ Our rejection region depends on a cut-off (C.O.)

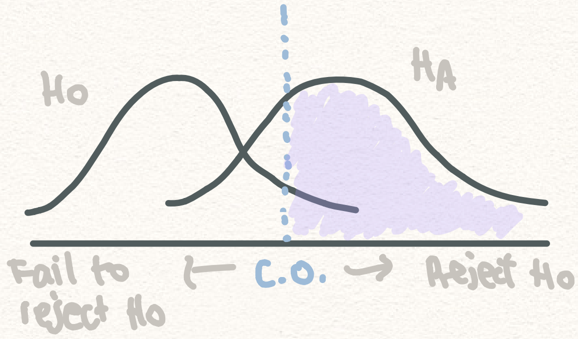
↳ We reject Ho if $\bar{x} >$ C.O. (Example data)

Not the same thing as α
α - p-value
C.O - Statistic

What does α look like?



What does power look like?



$\alpha = \mathbb{P}(\text{Type I error})$
$\quad = \mathbb{P}(\text{reject } H_0 \mid H_0 \text{ true})$

$\text{Power} = 1-\beta$
$\quad = 1 - \mathbb{P}(\text{Type II error})$
$\quad = \mathbb{P}(\text{reject } H_0 \mid H_A \text{ true})$

* If we decrease α, we decrease power

* Overall, there is a tradeoff α and β for a fixed n and C.O.

$\alpha = \mathbb{P}(\text{Type I})$
$\beta = \mathbb{P}(\text{Type II})$

For fixed n and C.O.

1) $\alpha \uparrow \Rightarrow \beta \downarrow$ and power $\uparrow$

2) $\alpha \downarrow \Rightarrow \beta \uparrow$ and power $\downarrow$

Also, the movement of α will move the C.O. but which direction depends on the direction in HA.

- What does n do?
- Work with the example of a population mean, μ.

↳ Our sample statistic is $\bar{X} = \frac{1}{n}\sum_{j=1}^{n} X_j$

$\sigma^2 = \text{var}(X_j)$

By the Central Limit Theorem, we know that $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

Example:

$H_0: \mu = \mu_0$ vs $H_A: \mu = \mu_1$

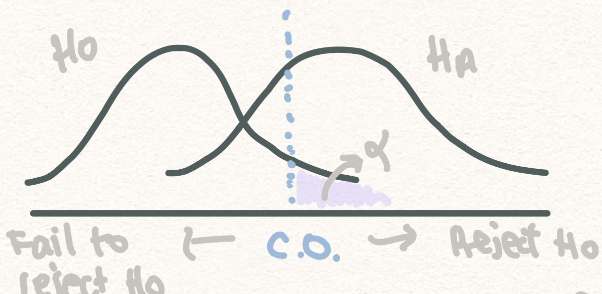
Under H_0, we have that:

$$\bar{x} \sim N\left(\mu_0, \frac{\sigma^2}{n}\right)$$

Under H_A, we have that:

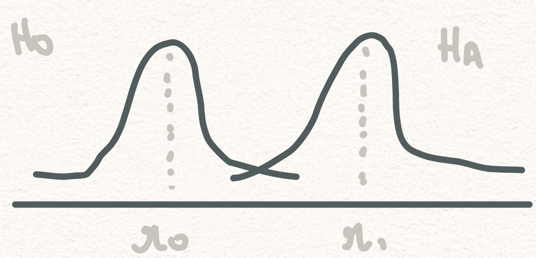
$$\bar{x} \sim N\left(\mu_1, \frac{\sigma^2}{n}\right)$$

In general,



→ power increases because you have more separability

n increases ⇒ spread decreases under both Ho and HA.



* The distribution under Ho and HA depend on the sample statistic and are specifically the sampling distribution of the statistic given the value of the parameter.

* We use Central Limit Theorem to approximate these distributions.

* In general, when n increases, we get more separability of H_0 and H_A due to decreased variance. Thus, power increases!

* Note that for a fixed α, the CO will also change.

* The test and calculations that you make depend now on a lot of moving parts:

  1) Metric being compared (mean, proportion, count, etc?)
  2) Do groups have equal variance?
  3) How long do you run the experiment/sample size*
  4) The equality statements under H_0 and H_A.

* The procedure is the same, but we need to think about our choice of α and what power we'd like to have.

* *Based on α and desired power, you can then do sample size calculations to determine your needed n.

  ↳ This becomes an important part of planning the experiment.

Example: Comparing Proportions

Question: Does your ability to pass a test depend on which section you take?

Data: $y_{ij} = \begin{cases} 1 & \text{, student } i \text{ passes from section } j \\ 0 & \text{, student } i \text{ fails from section } j \end{cases}$

$j = \{1, 2\}$

(section 1 is the cool one!!! )

Model: Bernoulli

$y_{ij} \sim \text{Bern}(\pi_j)$ Probability of pairing in session j

$i = 1, \ldots, n_j$
$j = 1, 2$

Sample statistic: $p_j = \frac{1}{n_j} \sum_{i=1}^{n_j} y$ $\Big\}$ unbiased estimator for π_j

By CLT, $p_j \sim N(\pi_j, \text{var}(p_j))$

$$\text{var}(p_j) = \frac{1}{n_j^2} \overset{\text{linear if indep.}}{\sum_{i=1}^{n_j}} \text{var}(y_{ij})$$

$$= \frac{1}{n_j^2} \sum_{i=1}^{n_j} \pi_j (1-\pi_j)$$

$$= \frac{1}{n_j^2} n_j (\pi_j (1-\pi_j))$$

$$= \frac{1}{n_j} \pi_j (1-\pi_j)$$

Hypothesis test: $H_0: \pi_1 = \pi_2$ $\overset{\text{equivalent}}{\equiv}$ $H_0: \pi_1 - \pi_2 = 0$

$H_A: \pi_1 \neq \pi_2$ $\quad$ $H_A: \pi_1 - \pi_2 = \delta$ for some $\delta \neq 0$.

What we actually care about is the difference in population parameters in each section, ie effect size.

$\boxed{\delta = \text{effect size!}}$

What we care about is $\delta = \pi_1 - \pi_2$
$\Rightarrow$ our statistic is $p_1 - p_2$.

From our previous CLT, we can calculate the distribution for $p_1 - p_2$:

$$p_1 - p_2 \sim N\left(\pi_1 - \pi_2,\ \frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2}\right)$$

Quick fact:
$X_1 \sim N(\mu_1, \sigma_1^2)$; $X_2 \sim N(\mu_2, \sigma_2^2)$ and X_1, X_2 independent

$\Rightarrow X_1 - X_2 \sim N(\mu_1 - \mu_2,\ \sigma_1^2 + \sigma_2^2)$

$$\Rightarrow \underbrace{\frac{(p_1 - p_2) - (\pi_1 - \pi_2)}{\sqrt{\frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2}}}}_{\text{sample statistic } z^*} \sim N(0,1)$$

Note: we have to estimate π_j w/ p_j to calculate z^*

↳ We want a cutoff so that

$\mathbb{P}(\text{rejecting } H_0 \mid H_0 \text{ true}) = \alpha$

$\mathbb{P}(|z^*| > c) = \alpha$

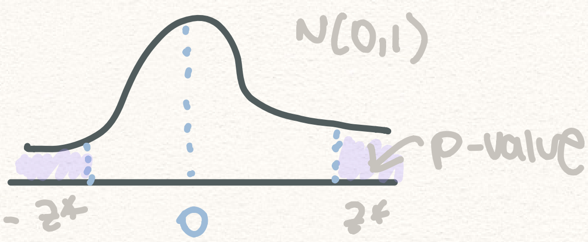
$\mathbb{P}(z^* > c \text{ or } z^* < -c) = \alpha$

↓ we can calculate this using quantiles of a $N(0,1)$

↳ we can also use a p-value:

$\mathbb{P}$ (observing a sample stat more extreme than what we just saw under H_0)

$= \mathbb{P}(|z| > z^*)$ where $z \sim N(0,1)$



If p-value $< \alpha$, we reject H_0.

Next time: Calculate power/sample size & multiple comparisions.