Lecture 4 9/9/2020

I don't think I'll give you exam questions on this:

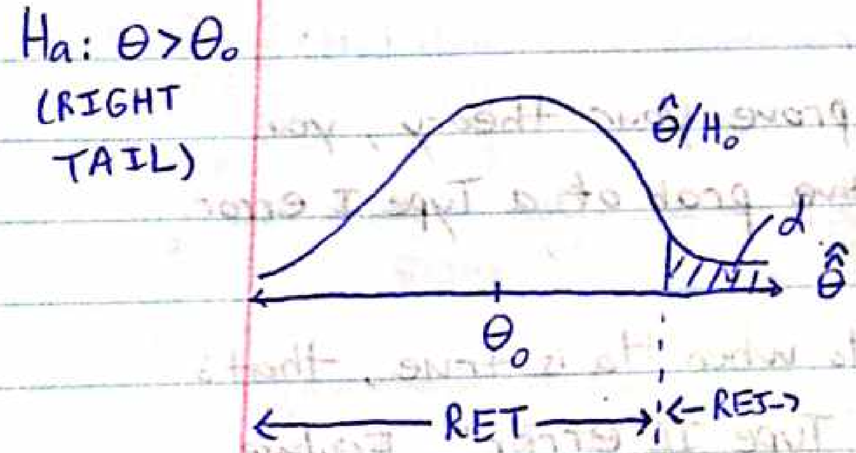
1) "level of a test" α is defined as $P(\text{Type I error}) \geq \alpha$

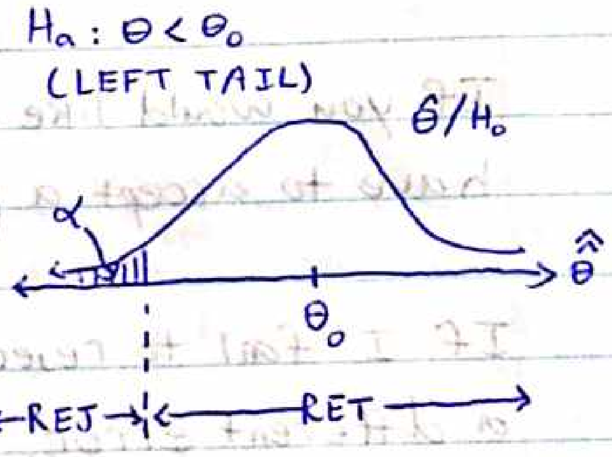
2) "Size of a test" is exactly $P(\text{Type I error})$.

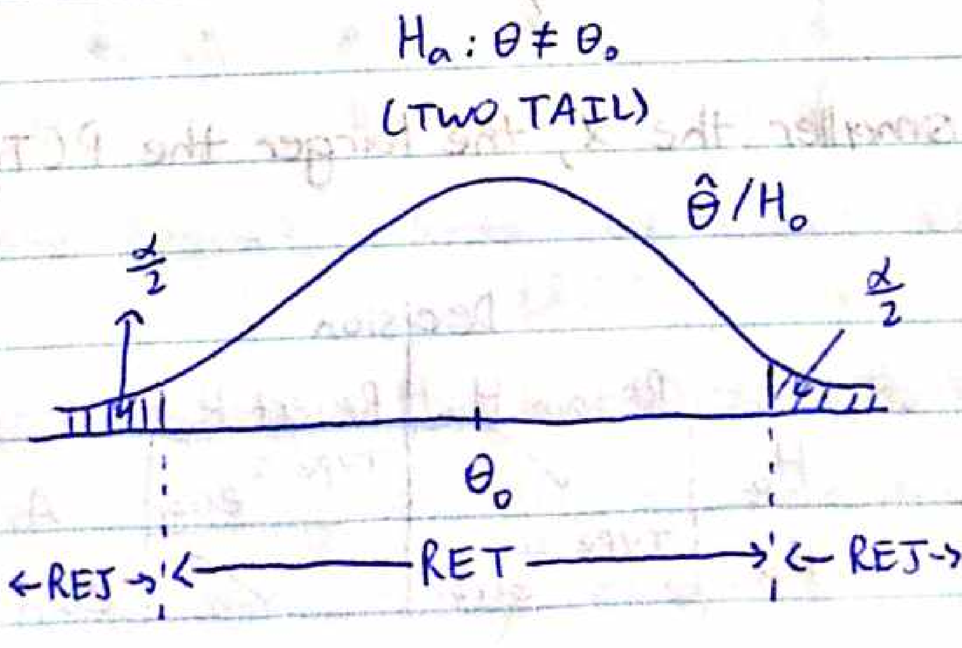
In our example the level was 5% but the size was 7.06% since $\alpha = 5\%$ was "unattainable"

If $\hat{\theta}/H_0$ is continuous, then level = size = α. If its discrete, some sizes won't be attainable.

If I want a level of $\alpha = 5\%$ and the size is lower, then I'm "cheating" (we'll see why next class).







What we did in the previous lecture was called a "Binomial exact test of one proportion". Downsides:

1) you need a binomial PMF calculator and it's a lot of work to get the retainment region

2) NOT all sizes are attainable: This is the recommended test.

Let $X_1, X_2, \ldots, X_n \sim$ i.i.d. some distr. w/ mean μ and var σ^2. The central limit theorem (CLT) shows that:

$$\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \xrightarrow{d} N(0,1)$$

"convergence in dist". It means as n gets large, the CDF of the left hand side (L.H.S.) looks more and more like the CDF of the right hand side (R.H.S.)

prob gonna use this more

approx distr.

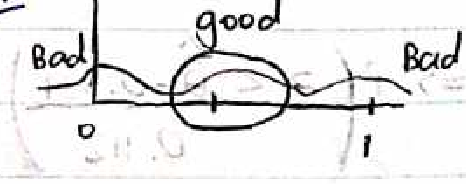
$$\Rightarrow {}^{*}\bar{X} \dot{\sim} N\left(\mu, \frac{\sigma^2}{n}\right) \text{ and } T = X_1 + \ldots + X_n \dot{\sim} N(n\mu, n\sigma^2)$$

If $X_1, \ldots, X_n \sim$ i.i.d. Bern(θ) and n is "large" then:

$$\hat{\theta} = \bar{X} \dot{\sim} N\left(\theta, \frac{\theta(1-\theta)}{n}\right)$$

this is a pretty good approx if θ is not too close to 0 or 1.

Ex:



How to perform an "approximate test"? There are many, many options even for the same DGP. The protocol goes as follows:

1) you think of a "test statistic" that could measure the departure away from H_0.

Scanned with CamScanner

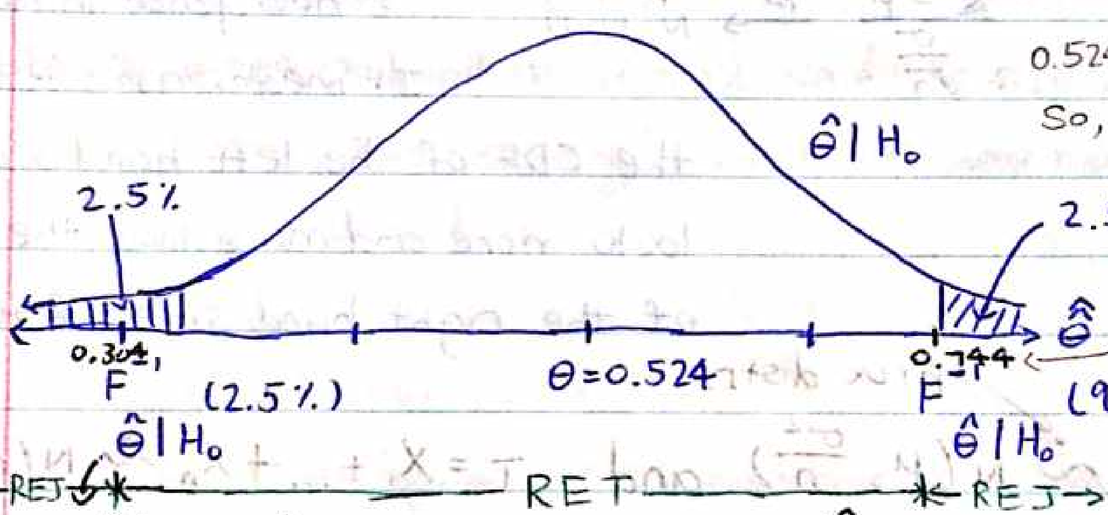
2) Derive the statistical estimator's *approx* distr. under H_0, $\hat{\theta} \mid H_0$.

3) Gauge the departure of $\hat{\hat{\theta}}$ from the bulk of the distr. $\hat{\theta} \mid H_0$ at level α.

$H_0: \theta = 0.524$, $H_a: \theta \neq 0.524$, $n = 20$, $\hat{\hat{\theta}} = 0.6$ (same as last class)

$$\hat{\theta} \mid H_0 \sim N\left(0.524, \frac{0.524(1-0.524)}{20}\right) = N(0.524, \underbrace{\overbrace{0.112^2}^{\sigma^2}}_{\sigma})$$

set $\alpha = 5\%$ $\Rightarrow$ $\frac{\alpha}{2} = 2.5\%$



$0.524 - 0.304 = 0.22$

So, $0.524 + 0.22 = 0.744$

$P(\hat{\theta} \mid H_0 \leq \hat{\hat{\theta}}) = 2.5\%$ Solve for $\hat{\hat{\theta}}$.

$$\Rightarrow P\left(\underbrace{\frac{\hat{\theta} \mid H_0 - 0.524}{0.112}}_{N(0,1)} \leq \frac{\hat{\hat{\theta}} - 0.524}{0.112}\right) = 2.5\%$$

look at table / use PC

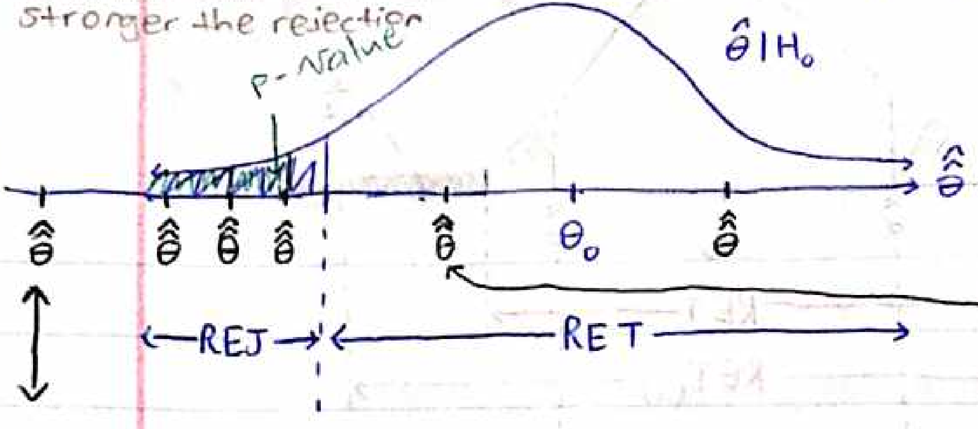
$$\Rightarrow P\left(Z \leq \frac{\hat{\hat{\theta}} - 0.524}{0.112}\right) = 2.5\% \Rightarrow \frac{\hat{\hat{\theta}} - 0.524}{0.112} \approx 1.96 \approx 2$$

$$\Rightarrow \hat{\hat{\theta}} = 0.304$$

One proportion z-test (approx test)

* The smaller the P-value, the stronger the rejection

$H_a: \theta < \theta_0$



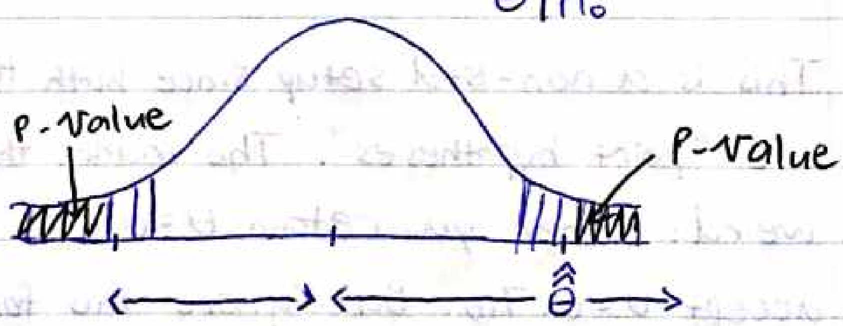
this estimate should imply a "stronger" rejection than this estimate

To measure the "strength" of a rejection (or "weakness" of a retainment), Fisher introduced the "P-value" also called the level of statistical significance as:

P-value = P(estimate $\hat{\hat{\theta}}$ is more extreme than the one observed | H_0)

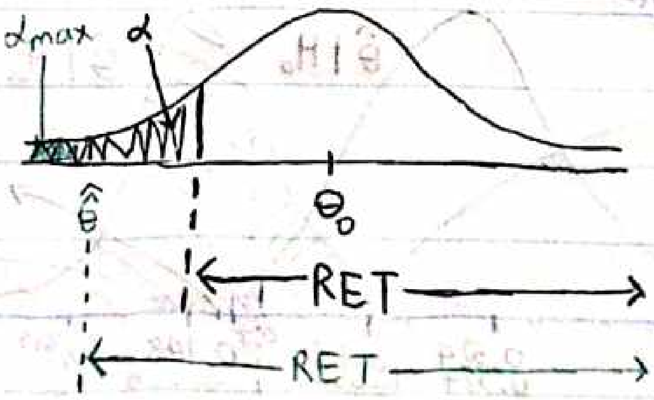
nebulous

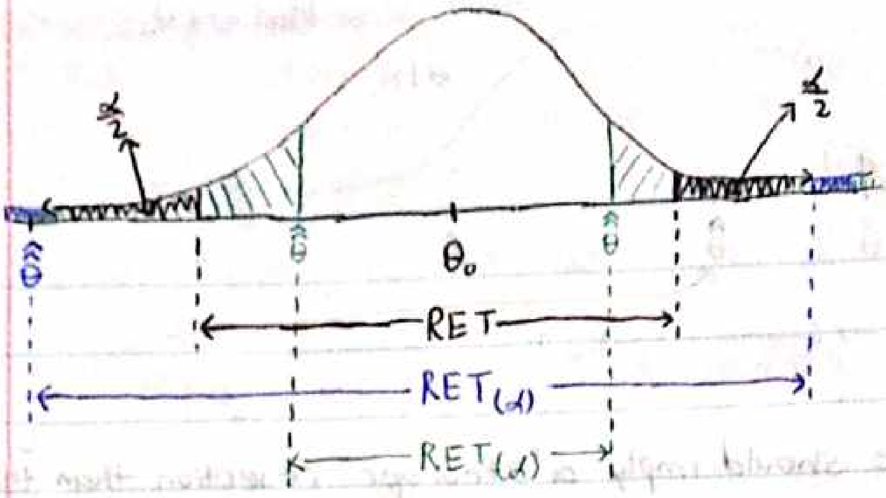
2-sided test



Real definition:

P-value = $\max\{\alpha : \hat{\hat{\theta}} \in RET(\alpha)\}$





If H_0 is retained, that means P-value $\geq \alpha$ and
if H_0 is rejected, that means P-value $< \alpha$

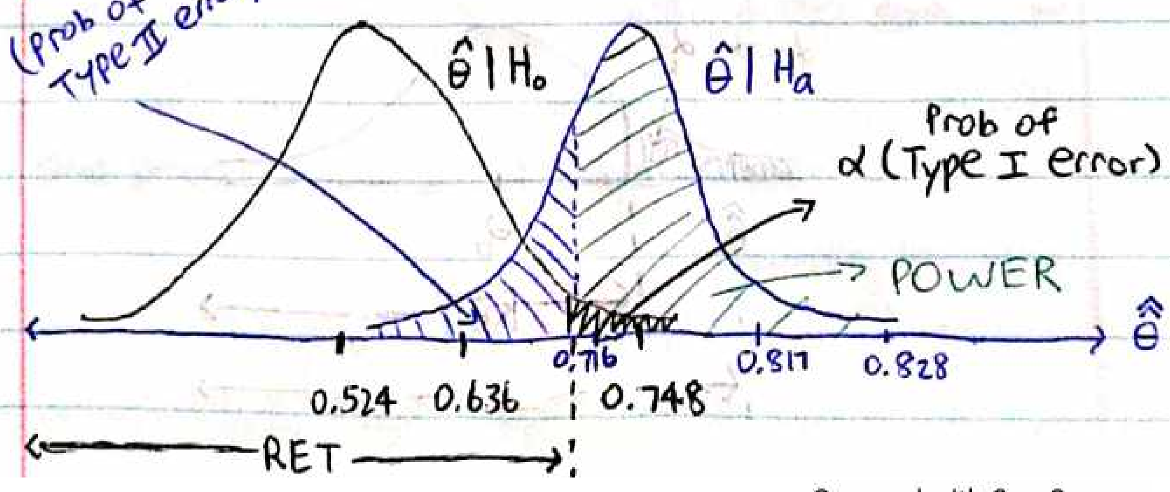
Type II errors and POWER

DGP: $X_1, \dots, X_n \sim$ i.i.d. Bern (θ)

$H_0: \theta = 0.524 = \theta_0$ but $H_a: \theta = 0.716 = \theta_a$

This is a non-std setup since both H_0 and H_a are "point hypotheses". This makes the outcome weird: either you retain $\theta = 0.524$ or you accept $\theta = 0.716$. But ignore this for now.

$\hat{\theta} | H_0 \sim N(0.524, 0.112^2)$, $\hat{\theta} | H_a \sim N(0.716, 0.101^2)$



At $\alpha = 5\%$, the Z-value is 1.645 which means the rejection region ends at

$\hat{\theta} = 0.524 + 1.645 * 0.112 = 0.708$

Errors

		Decision	
		RET	REJ
Truth	H_0		Type 1
	H_a	Type 2	

POWER = P(Rejecting $H_0 | H_a$)

$= 1 - P(\text{Retaining } H_0 | H_a) = 1 - P(\text{Type II error})$

POWER is the probability of proving your theory is true!! You want POWER to be LARGE i.e. near 100%.

$P(\text{Type II error}) = P(\hat{\theta} | H_a \in RET) = P(\hat{\theta} | H_a \leq 0.708)$

$$= P\left(\frac{\hat{\theta} | H_a - 0.716}{0.101} \leq \frac{0.708 - 0.716}{0.101}\right)$$

$= P(Z \leq -0.079) \approx 47\%$

$\Rightarrow$ POWER $\simeq 53\%$