(I don't think I'll give you exam question on this.

9/9/20

"Level of a test" alpha is defined as P(Type I error) >= alpha

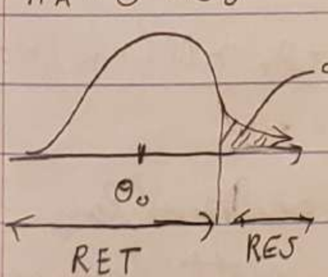
"Size of a test" is exactly P(Type I error).

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

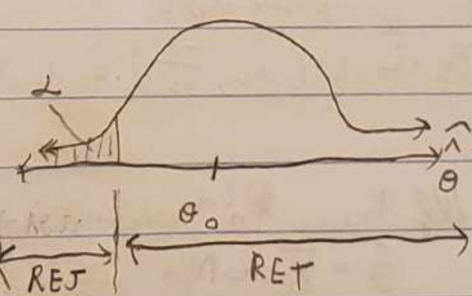
If thetahat | H_0 is continuous, then level = size = alpha.
If it's discrete, some sizes won't be attainable

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

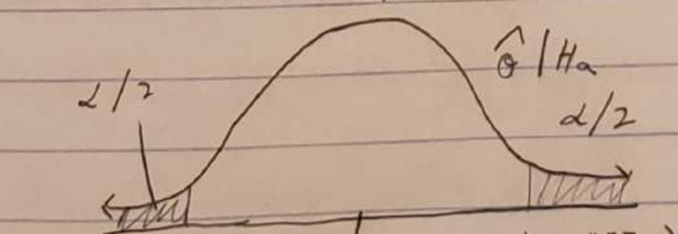
$H_A: \theta > \theta_0$



$H_a: \theta < \theta_0$



$H_a: \theta \neq \theta_0$



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 $x_1, x_2, \ldots x_n \sim$ iid some distribution with mean mu and variance sigma^2 (sigsq) The central limit theorem (CLT) shows that:

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

"convergence in distribution". It means as n gets large, the CDF of the left hand side (lhs) looks more and more like the CDF of the right hand side (rhs)

use this big time ← *

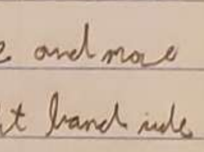
approx distr.

$$\Rightarrow \bar{x} \,\dot\sim\, N\left(\mu, \frac{\sigma^2}{n}\right) \text{ and } T = x_1 + \ldots + x_n \,\dot\sim\, N(n\mu, n\sigma^2)$$

If $x_1, \ldots x_n \sim$ iid Bern(theta) 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 approximation if theta is not too close to 0 or 1



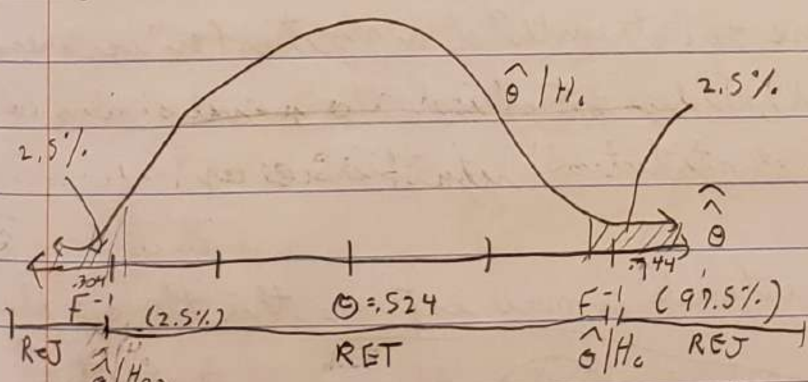
How to peform an "approximate test"? There are many, many options even for the same DGP. The protocol goes as follows:

1. you think of a "test statistic" that could measure the departure away from H_0.
2. Derive the statistical estimator's *approx* distribution under H_0 thetahat | H_0
3. Gauge the departure of thetahathat from the bulk of the distribution thetahat | H_0 at level alpha

$H_0: \theta = .524$, $H_a: \theta \neq .524$, $n = 20$, $\hat{\hat{\theta}} = 0.6$ (same as last class

$$\hat{\theta} \mid H_0 \sim N\left(.524, \frac{.524(1-.524)}{20}\right) = N(.524, \underbrace{.112}_{\sigma}{}^{2})$$

Set $\alpha = 5\% \Rightarrow \frac{\alpha}{2} = 2.5\%$ (241 exercise)



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

$$\Rightarrow P\left(\underbrace{\frac{\hat{\theta} \mid H_0 - .524}{.112}}_{N(0,1)} \leq \frac{\hat{\hat{\theta}} - .524}{.112}\right) = 2.5\%$$

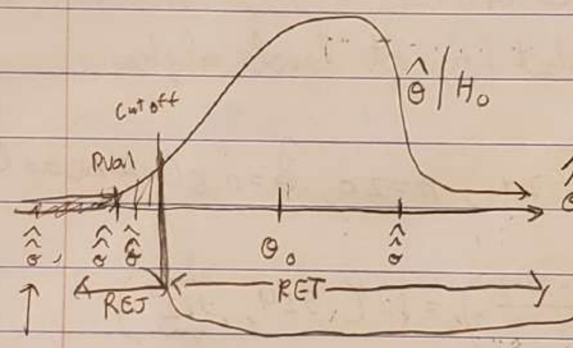
Look at table or use computer

$$\Rightarrow P\left(Z \le \frac{\hat{\theta} - .524}{.112}\right) = 2.5\,\% \Rightarrow \frac{\hat{\theta} - .524}{.112} \simeq 1.96 \simeq 2$$

$$\Rightarrow \hat{\theta} = .304$$

One proportion Z test (approximate test).

$H_1: \theta < \theta_0$

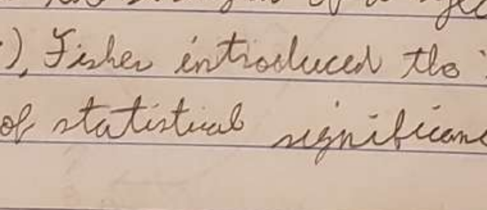


this estimate should imply a "stronger" rejection than this estimate

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:

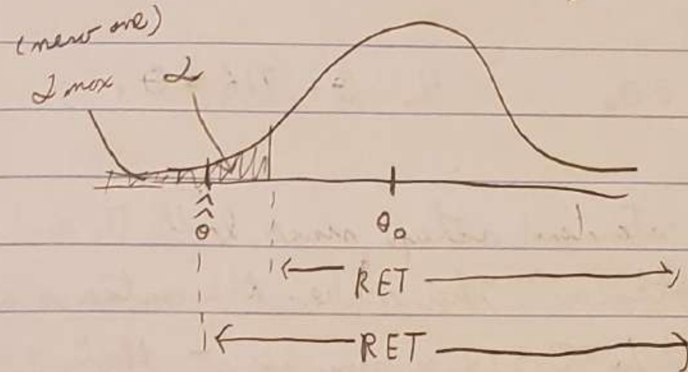
Pval := P(estimate ($\hat{\theta}$) is more <u>extreme</u> than the one observed | H_0)

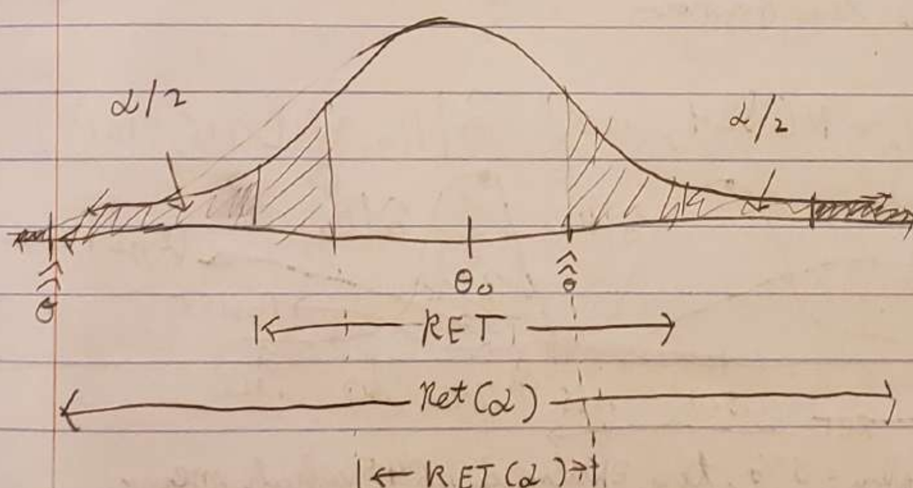
nebulous



Real definition:

$P_{val} = \max\{\alpha : \hat{\theta} \in RET(\alpha)\}$





If H_0 is retained, that means
pval >= alpha
and if H_0 is rejected, that means
pval < alpha

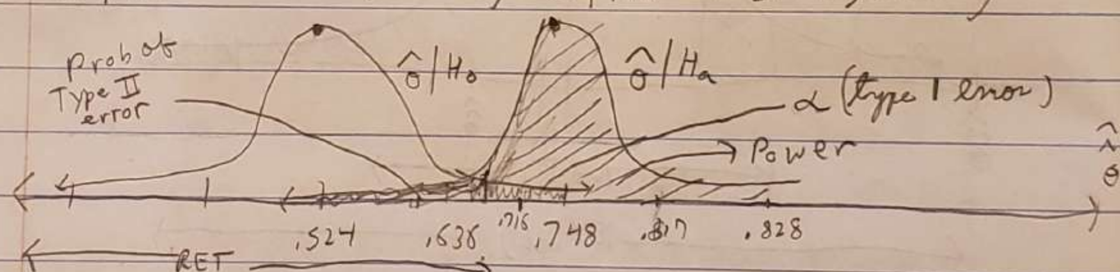
Type II error and POWER

DGP: $X_1, \dots, X_n \sim$ iid Bern(theta)

$H_0: \theta = .524 = \theta_0$ $\quad$ $H_a: \theta = .716 = \theta_1$

This is a non-standard 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

$\hat{\theta} \mid H_0 \sim N(.524, .112^2)$, $\hat{\theta} \mid H_a \sim N(.716, .101^2)$



at alpha = 5%, the z value is 1.645 which means the rejection region ends at, thetahathat = .524 + 1.645 * .112 = .708

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

Errors

Truth	Decision: RET	Decision: REJ
H_0		Type I
H_a	Type II	

Power is the probability of providing your theory is true !! you want power to be large i.e near 100%

$$P(\text{Type II error}) = P(\hat{\theta} \mid H_a \in RET) = P(\hat{\theta} \mid H_a \leq .708)$$

$$= P\left(\frac{\hat{\theta} \mid H_a - .716}{.101} \leq \frac{.708 - .716}{.101}\right) = P(Z \leq -.079) \simeq 47\%$$

$$\Rightarrow \text{Power} \simeq 53\%$$

pretty bad, like a coin flip