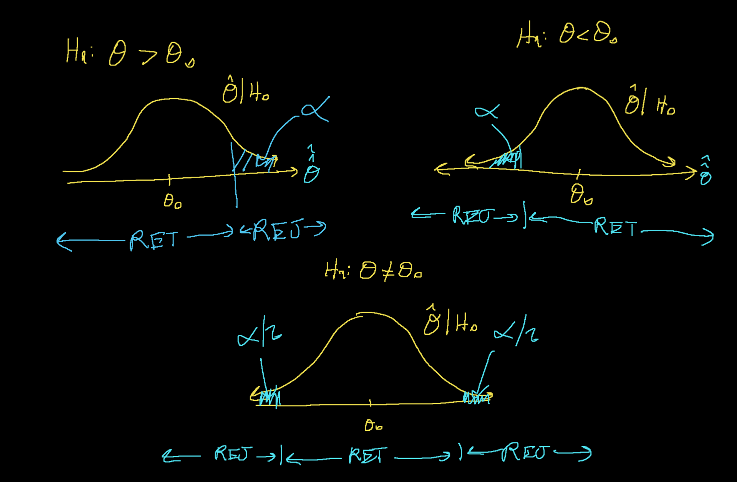


What we did in the previous lecture was called a "binomial exact test of one proportion". Downsides: (1) you need a binomial PMF calculator and it's a lot of work to get the retainment region (2) not all sizes are attainable. This is the recommended test.

Let X_1, X_2, ..., X_n ~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).

approx distr.

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

If X_1, ..., X_n ~iid Bern(theta) and n is "large" then:

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

this is a pretty good approximation if theta is not too close to 0 or 1.

good

BAD

BAD

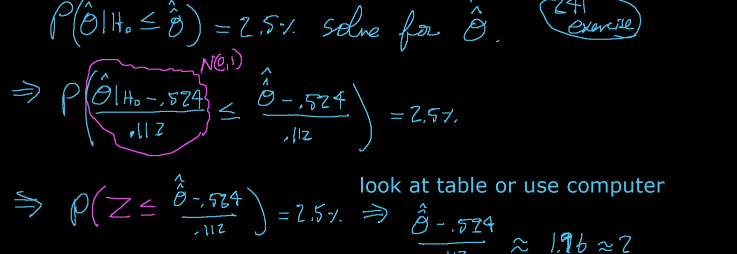
How to perform an "approximate test"? There are many, many options even for the same DGP. The protocol goes as follows:
(1) you think of a "test statistic" that could measure the departure away from H_0.
(2) Derive the statistical estimator's *approx* distribution under H_0, thetahat | H_0.
(3) Gauge 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)

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

σ

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



$P(\hat{\theta} | H_0 \leq \hat{\hat{\theta}}) = 2.5\%$ solve for $\hat{\hat{\theta}}$. 241 exercise

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

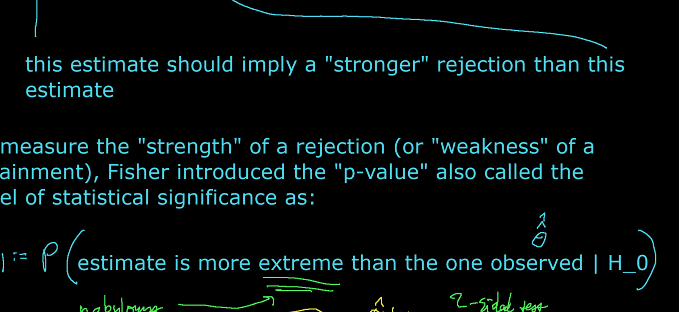
$$\Rightarrow P\left(Z \leq \frac{\hat{\hat{\theta}} - .524}{.112}\right) = 2.5\% \Rightarrow$$

look at table or use computer

$$\frac{\hat{\hat{\theta}} - .524}{.112} \approx 1.96 \approx 2$$

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

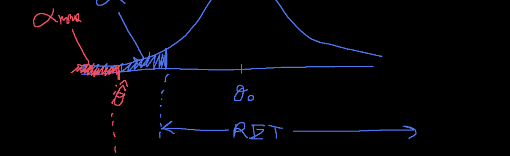
One proportion z-test (approximate test).



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

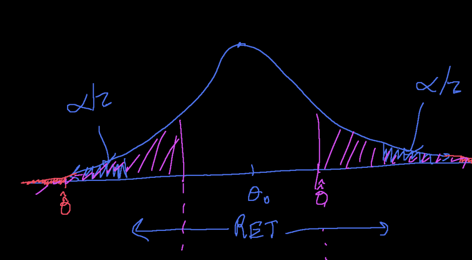
To measure the "strength" of a rejection (or "weakness" of a retainment), Fisher introduced the "p-value" also called the level of statistical significance as:

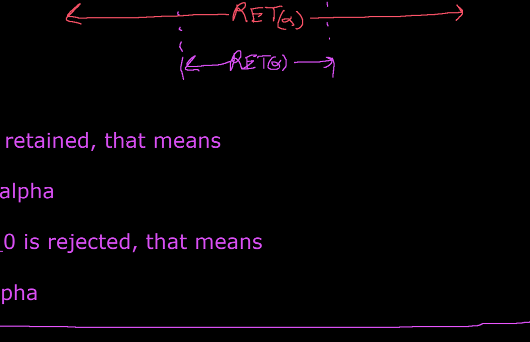
Pval := P(estimate is more extreme than the one observed $\hat{\hat{\theta}}$ | H_0)



Real definition:

$$\text{Pval} = \max\{\alpha : \hat{\hat{\theta}} \in RET(\alpha)\}$$





If H_0 is retained, that means

pval >= alpha

and if H_0 is rejected, that means

pval < alpha

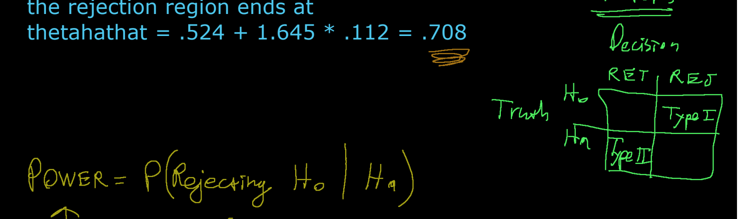
Type II errors and POWER

DGP: X_1, ..., X_n ~iid Bern(theta)

$H_0: \theta = .524 = \theta_0$ but $H_a: \theta = .716 = \theta_a$

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.

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



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

Errors

		Decision RET	Decision REJ
Truth	H_0		Type I
	H_a	Type II	

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

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

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

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

$$\Rightarrow \text{POWER} \approx 53\%$$