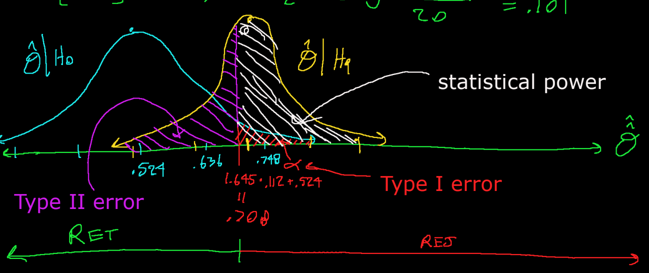
The definition of "statistical significance" is as follows. An estimate is "statistically significant at level alpha" if the null hypothesis is rejected (or equivalently the pvalue is less than alpha).

Type II errors and statistical power
DGP: iid Bern(θ) and n = 20 and we'll say it's large enough to use the CLT (normal approx)

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

$\alpha = 5\% \Rightarrow z = 1.645$

$SE[\hat{\theta} | H_0] = .112, \quad SE[\hat{\theta} | H_a] = \sqrt{\frac{.716(1-.716)}{20}} = .101$



Type II errors are retaining the null when the alternative is true.
P(Type II error) = $P(\hat{\theta}|H_a \in RET)$

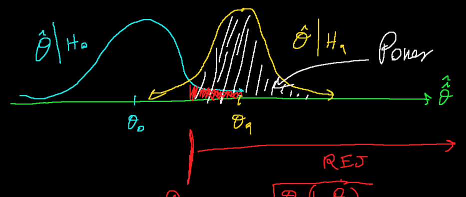
$$= P(\hat{\theta}|H_a \leq .708) = P\left(\frac{\hat{\theta}|H_a - .716}{.101} \leq \frac{.708 - .716}{.101}\right)$$

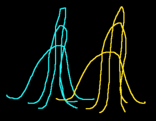
$$= P(Z \leq -0.079) = 46.9\%$$

Power is the probability you reject the null if the alternative is true which is 1 - P(Type II error).

DGP: iid Bern(θ) and n is large enough to use the CLT.

$H_0: \theta = \theta_0$ vs. $H_a: \theta = \theta_a$ s.t. $\theta_a > \theta_0$ at level α.





$\theta_0 + z_{1-\alpha}\sqrt{\frac{\theta_0(1-\theta_0)}{n}}$

P(Z > z) = 1 - Φ(z)
CDF of the std normal

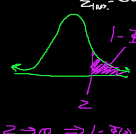
$$Pow = P\left(\hat{\theta}|H_a > \theta_0 + z_{1-\alpha}\sqrt{\frac{\theta_0(1-\theta_0)}{n}}\right)$$

$$= P\left(\underbrace{\frac{\hat{\theta}|H_a - \theta_a}{\sqrt{\frac{\theta_a(1-\theta_a)}{n}}}}_{Z} > \frac{\theta_0 + z_{1-\alpha}\sqrt{\frac{\theta_0(1-\theta_0)}{n}} - \theta_a}{\sqrt{\frac{\theta_a(1-\theta_a)}{n}}}\right)$$

$$= 1 - \Phi\left(\frac{\theta_0 + z_{1-\alpha}\sqrt{\frac{\theta_0(1-\theta_0)}{n}} - \theta_a}{\sqrt{\frac{\theta_a(1-\theta_a)}{n}}}\right)$$

$$= 1 - \Phi\left(\frac{-\sqrt{n}(\theta_a - \theta_0) + z_{1-\alpha}\sqrt{\theta_0(1-\theta_0)}}{\sqrt{\theta_a(1-\theta_a)}}\right)$$

$z_{5\%} = 1.645$
$z_{2.5\%} = 1.96$
$z_{100\%} = \infty$



$z \to \infty \Rightarrow 1 - \Phi(z) \to 0$
$z \to -\infty \Rightarrow 1 - \Phi(z) \to 1$

as n --> ∞, POW --> 100%
as θ_a --> 1, POW --> 100% (need a rigorous proof)
as α --> 0, POW --> 0%

Scientific standards: alpha = 5% (or maybe 1%) and power is set at 80% in clinical trials. Then, you do a "power calculation" to determine the n you need to get POW = 80%.

Let's do a different type of survey. How tall are you in inches (men only). $\vec{x} = $ <70, 72, 73, 68, 69, 70, 67, 72, 71, 73> $\Rightarrow \bar{x} = 70.0''$, $n = 10$

We assume that these realization are from a normal DGP with known variance. DGP: $X_1, \ldots, X_{10} \overset{iid}{\sim} N(\theta, \sigma^2 = 4^2)$
The std dev of 4 comes from a very large sample of Americans. This is also how we know it's normal.

What point estimate should we use? Thetahat = sample average is unbiased (it's always unbiased regardless of the DGP).

$$\hat{\theta} \sim N\left(\theta, \frac{\sigma^2}{n}\right)$$ fact from math 241

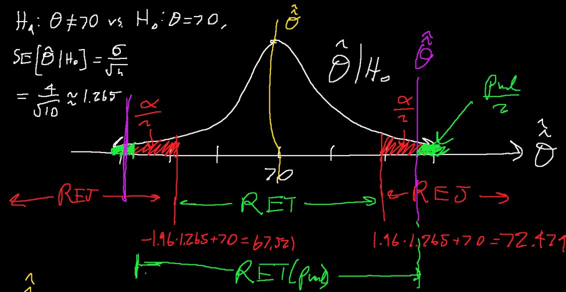
$H_a: \theta \neq \theta_0$ vs. $H_0: \theta = \theta_0$ two sided one-sample z-test
$H_a: \theta < \theta_0$ vs $H_0: \theta \geq \theta_0$ left-sided one-sample z-test
$H_a: \theta > \theta_0$ vs $H_0: \theta \leq \theta_0$ right-sided one-sample z-test

These are all *exact* tests because we know the distribution of the estimator exactly.

Let's test if our class is different from the American population which has mean of exactly 70" at a 5% level. $\Rightarrow z = 1.96$

$H_a: \theta \neq 70$ vs. $H_0: \theta = 70$,

$SE[\hat{\theta}|H_0] = \frac{\sigma}{\sqrt{n}} = \frac{4}{\sqrt{10}} \approx 1.265$



$\hat{\theta} \in RET \Rightarrow$ retain H_0. There is not sufficient evidence to claim that the mean height of male students in this class is different from the American population.

How does one calculate a 2-sided p-value? From the definition:

$$p_{val} := \max\{\alpha : \hat{\theta} \in RET(\alpha)\} = P\left(|\hat{\theta}|H_0| > |\hat{\theta} - \theta|\right)$$

$$= 2P\left(Z > \left|\frac{\hat{\theta} - \theta_0}{SE[\hat{\theta}|H_0]}\right|\right)$$