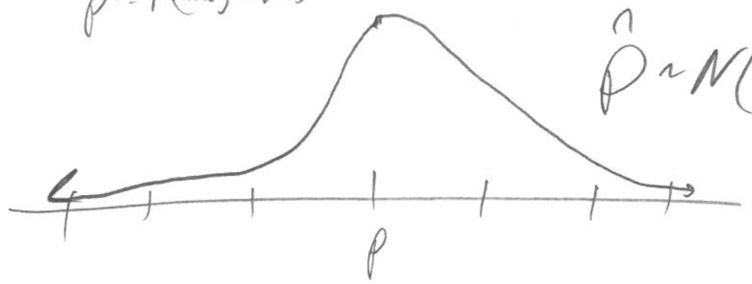
Lec 22 12/3/15 Math 241

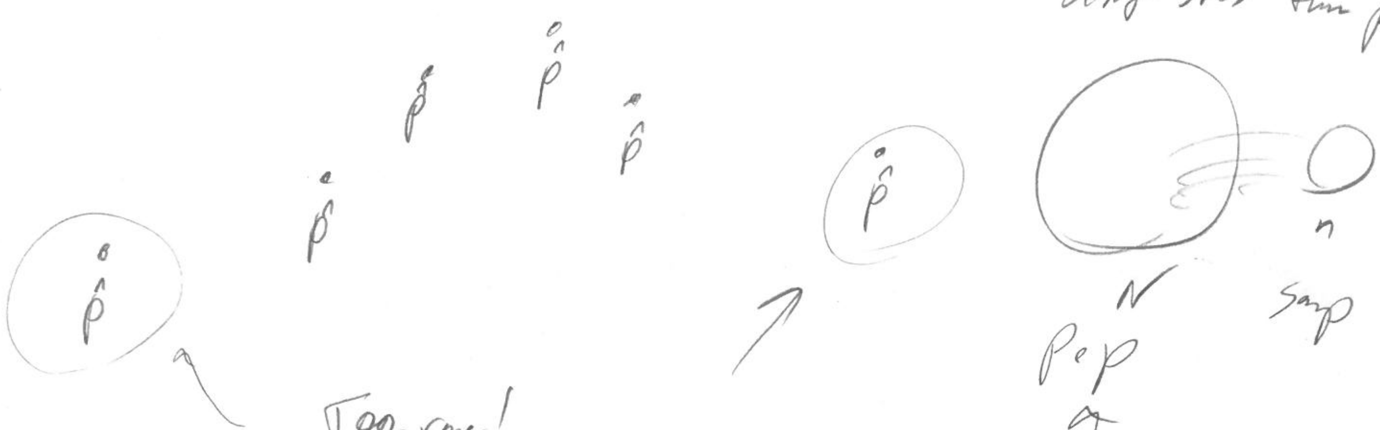
Testing: I have a theory that male/female births are even

$p := P(\text{male}) = 0.5$

$$\hat{p} \sim N\left(p, \left(\sqrt{\frac{p(1-p)}{n}}\right)^2\right)$$

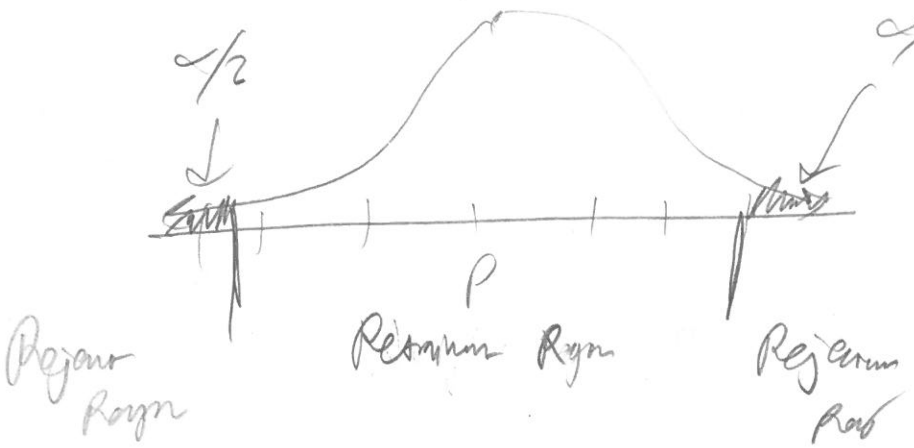


Get data, sample size n using SRS from pop



At some pt you have to say too rare! What pt do you stop?

$$\alpha := P(\hat{p} \text{ being too rare})$$



apportion your alpha on both sides evenly

Rule:

If $\hat{p} \in$ Retainment Region $\Rightarrow$ Retain Theory

If $\hat{p} \in$ Rejection Region $\Rightarrow$ Reject Theory

What is retainment region? $\alpha = 5\% \Rightarrow 1-\alpha = 95\%$

$$1-\alpha = P\left(|Z| \le z_{\frac{\alpha}{2}}\right) = P\left(|Z| \le z_{2.5\%}\right) = P\left(|Z| \le 2\right)$$

$$= P\left(\left|\frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}}\right| \le 2\right) = P\left(-2 \le \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} \le 2\right)$$

$$= P\left(-2\sqrt{\frac{p(1-p)}{n}} \le \hat{p}-p \le 2\sqrt{\frac{p(1-p)}{n}}\right) = P\left(p-2\sqrt{\frac{p(1-p)}{n}} \le \hat{p} \le p+2\sqrt{\frac{p(1-p)}{n}}\right)$$

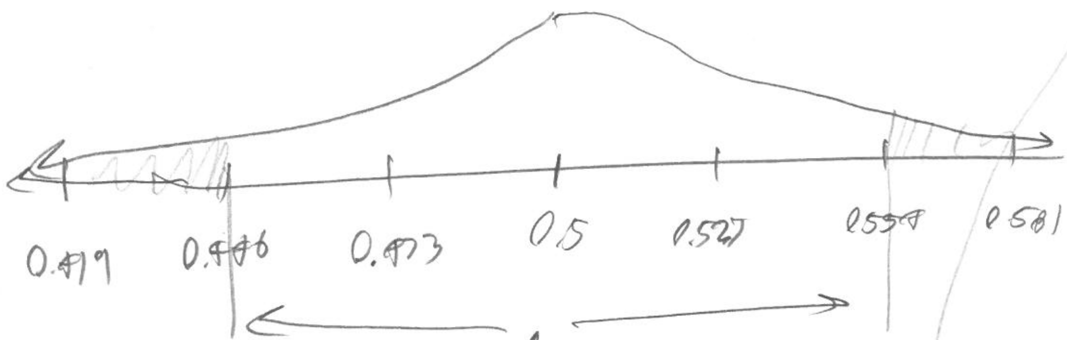
$$= P\Bigg(\hat{p} \in \underbrace{\left[p \pm 2\sqrt{\frac{p(1-p)}{n}}\right]}_{\text{Retainment Region}}\Bigg)$$

Generally: Ret Reg $:= \left[p \pm z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}\right]$

α is your tolerance for error just like for CI's!

Sample $n = 345$ babies ..

$$\hat{p} \sim N\left(0.5, \sqrt{\frac{0.5(1-0.5)}{345}}\right) = N(0.5, .0269^2)$$



At $\alpha = 0.05 \Rightarrow$ Ret Regn $= [0.5 \pm 2 \cdot .0269] = [0.446, 0.554]$

Now, get the $\hat{p}$. 164 babies male $\Rightarrow \hat{p} = \frac{164}{345} = 0.48$

$\hat{p} \in$ Ret Regn $\Longrightarrow$ Retain Theory

Why do we need this? Can't we just tell that $0.48 \approx 0.5$ and call it a day? Flip coin 100 times to see if fair.

Scenario #1: 51 Heads $\Rightarrow$ Fair

Scenario #2: 98 Heads $\Rightarrow$ Unfair

Scenario #3: 61 Heads $\Rightarrow$???

Not so clear

So let's run the test. "null
Make your theory, called a hypothesis," H_0. Then the alternative
hypothesis is just param space $\setminus H_0$

$H_0: p = 0.5$ (coin fair)
$H_a: p \neq 0.5$ (coin unfair)
} Formally, a 2-sided, 1-proportion hypothesis test

Flip 100x: $n = 100 \Rightarrow$ $\hat{p} \sim N\left(0.5, \sqrt{\frac{0.5 \cdot 0.5}{100}}\right) = N(0.5, 0.05^2)$
$\alpha = 5\%$

Retain region $= [0.5 \pm 2 \cdot .05] = [0.4, 0.6]$

$0.61 \notin$ Retain Region $\Rightarrow$ Reject theory!

$\longrightarrow$ But it's so close!!! If α was a little different... it would've passed (more on this later)

Let's do another.

M&M company said prop. of blue M&M's = 20%

$H_0: p = 0.2$

$H_a: p \neq 0.2$

$\alpha = 1\% \Rightarrow z_{0.5\%} = 2.56$ (~~test~~ for all: I will give all possible z values, don't worry)

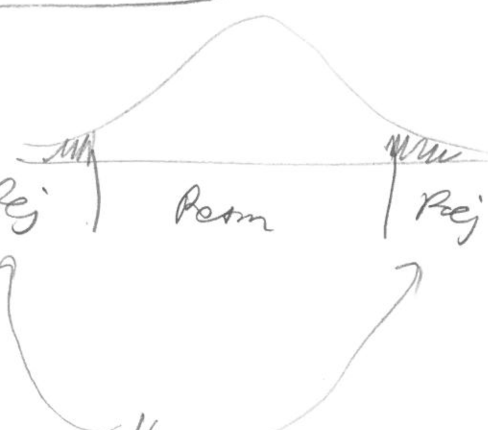
Retainment Region $= \left[0.2 \pm 2.56 \sqrt{\frac{0.2(1-0.2)}{n}}\right]$ $1.024/\sqrt{n}$

Do exp!

Why should I choose a given α?
Where does $\alpha = 5\%$ come from?

		Retain Ho	Reject Ho
The truth	Ho should be retained	✓	Type I err
	Ho is false	Type II err	✓

What is P(Type I err)?
$= \alpha$



Ho is still a good thing here!

You control Type I err

P(Type II err) ... Prob 242

but if Type I err ↑ ⇒ Type II err ↓
(vice versa)

Ho: Drug does not work
Ha: Drug works

which matters more?

Type I err: releasing a useless drug to market!
Type II err: saying the drug doesn't work when it really does!

Alarm system

H_0: no fire Retain: no alarm

H_a: fire Reject: sound the alarm

no fire, no alarm	no fire, false alarm
fire, no alarm	fire, alarm

Type I err: false alarm

Type II err: no alarm when there's a fire (people can die)

α should be high to minimize Type II err

⇒ You need to decide α... what you're comfortable with

And then understand that you may make a mistake!

Data is random. You make decisions based on random stuff, you may screw up!

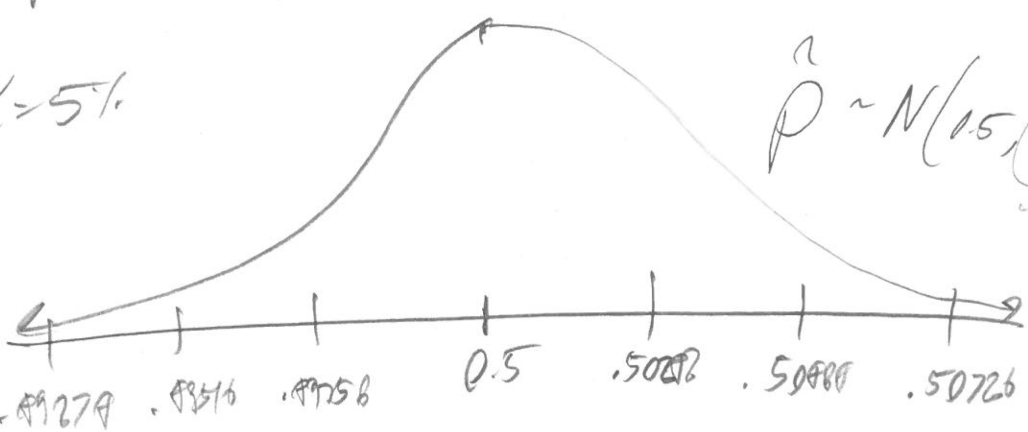
Return to M/F ratio

$H_0: p = 0.5$, $H_a: p \neq 0.5$, In 2008 $n = 4{,}272{,}000$, $= 2{,}173{,}000$

$\alpha = 5\%$

$\hat{p} \sim N\left(0.5, \sqrt{\frac{0.5 \cdot 0.5}{n}}^{\,2}\right)$

$.00242^2$



Retain Region = [.49516, .50484] $\hat{p} = \text{——} = 0.51165$ $\hat{p} \in$ Retain Region

⇒ Reject H_0

Why did I fail to reject before?

With higher n, much more clarity! Your vision is very clear!

You have more "power" to detect an effect

$n\uparrow$ P(Type II err) $\downarrow$ P(Type I err) $=$ (no difference)

If n gets larger and larger ... most H_0's in the real world will be rejected ... why?

$0.51165 \neq 0.5$

higher ... men are born more often ... why?

No one knows ....

Why does "retain" $\neq$ "accept"

Retain means I'll keep it until I find something better ...

Accept means it is the gospel truth which you hold onto forever

All of science works this way. We have H_0's for all of our scientific theories... and as soon as we see some evidence they are not true "beyond a reasonable doubt", we trash them.

Okay so why is retain ≠ accept?

$H_0: p = 0.5$, $H_a: p \neq 0.5$

$n = 345$