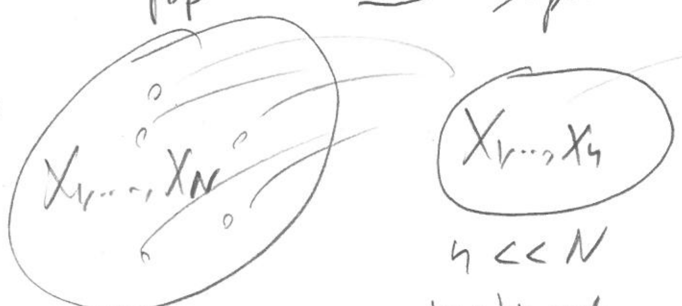
Lecture 26 5/7/15 Math 241

Review

Simple random

Pop $\supset$ Sample

$N \approx \infty$



realized .. $x_1, \dots, x_n$

$\hat{p} := \frac{x_1 + \dots + x_n}{n} = \frac{\#1\text{'s}}{n}$

$n << N$

but big enough for ...

$\bar{X} \sim N\left(\mu, \left(\frac{\sigma}{\sqrt{n}}\right)^2\right)$ to be true

In the case of Bernoulli pop., assume $X_1, \dots, X_n \overset{iid}{\sim} \text{Bern}(p)$

$\Rightarrow \hat{P} \sim N\left(p, \left(\sqrt{\frac{p(1-p)}{n}}\right)^2\right)$ special case of above.

Stat

Inference $\rightarrow$ Using the sample we want to say something about the pop. What do we care about here?

1. Point estimate: best guess? $\hat{p}$ for p
2. Interval estimate: range of guesses for p

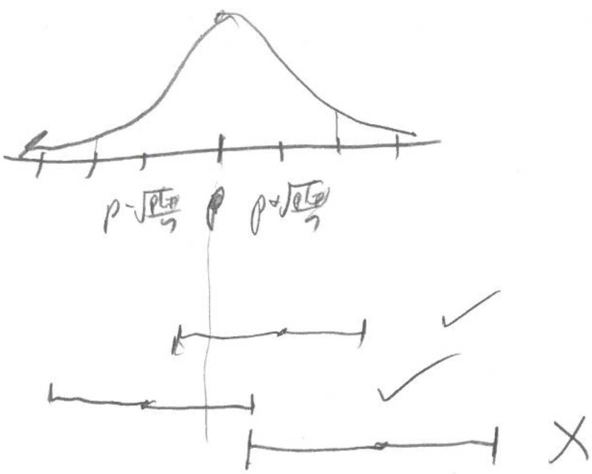
α

You get to choose how big the range is

e.g. $\alpha = 5\% \Rightarrow$ Prob not covering

$CI_{p,1-\alpha} := \left[\hat{p} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right] \Rightarrow P(\text{coverage}) = 1-\alpha$

What is coverage?



Coverage is
1. Long run chance of covering
2. A priori chance of covering
3. Prob (covering)

In the real world

You get an interval and don't know if you covered

~~Inference~~ Testing

Is a theory that we have true?

Theory: an idea about the population distr. In our case, an idea about p.

Theory: male and female humans are born in the same proportion.

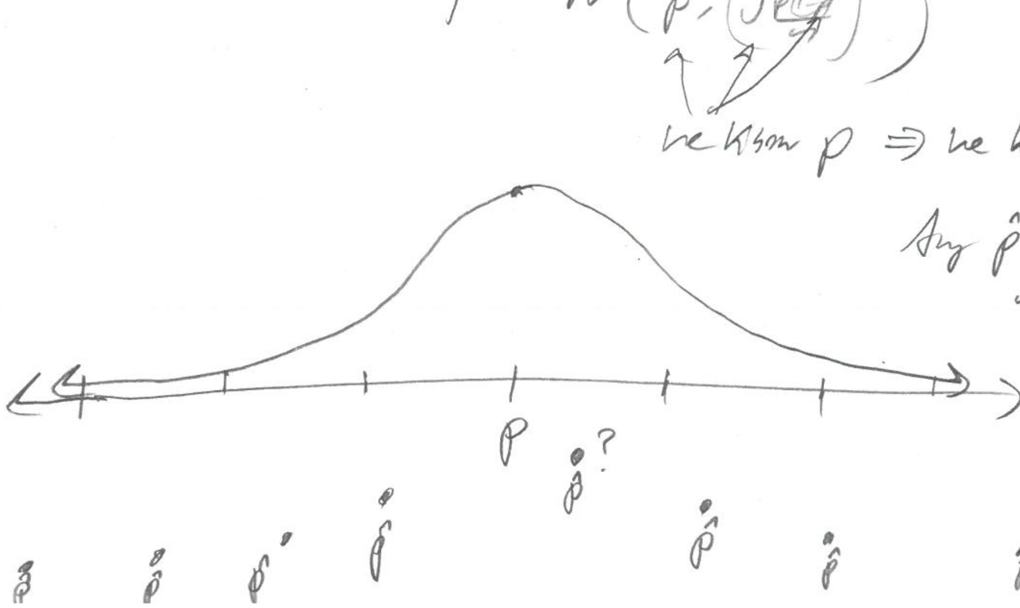
let p be $:= P(\text{MALE})$. $\Rightarrow$ $p = 0.5$

How to test theory? Get data acc'd to Pop Samp (SRS) paradigm...
and see if it would be expected. (Sample size is big enough for CLT)

If the theory is true
$$\hat{P} \sim N\left(p, \left(\sqrt{\frac{p(1-p)}{n}}\right)^2\right)$$

we know p $\Rightarrow$ we know this distr. exactly.

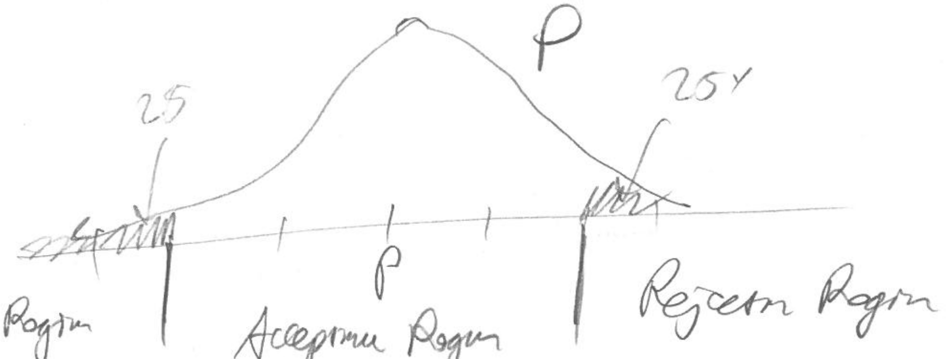
Any $\hat{p}$ can happen. But which ones are reasonable and unreasonable



At some point, you have to say it's "absurd" because it's too rare. What point do you stop?

Pick a pt $\alpha = P(\hat{p}$ being too rare$)$

If $\alpha = 5\%$ there is rareness too low (2.5%) and too high (2.5%)



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

If $\hat{p} \notin$ Acceptance Region $(\hat{p} \in$ Rejection$)$ $\Rightarrow$ Throw Out theory

What is acceptance region?

$$\alpha = 5\% \quad = \quad P(|Z| \le 2) = P\left(\left|\frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}}\right| \le 2\right) = P\left(|\hat{p}-p| \le 2\sqrt{\frac{p(1-p)}{n}}\right)$$

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

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

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

$|x+3| < 7$

$x+3 < 7$ & $-x-3 < 7$

$\Rightarrow x+3 > -7$

Generally Acceptance Region is

$\left[p \pm z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}\right]$ where α is your tolerance for error (just like in CI's)

Why do we need this? Can't we tell immediately if the theory is true or not just from $\hat{p}$? Example...

Flip a coin 100 times..

Why you unlikely to accept this

Scenario #1 51 heads. Is the coin fair? i.e. $p = 0.5 \neq \hat{p} = 0.51$

Scenario #2 98 heads. Is the coin fair? i.e. $p = 0.5 \neq \hat{p} = 0.98$

Intuitively too big!

Scenario #3 61 heads. Is coin fair? Not so clear!!!

Let's investigate..

Theory: coin fair $\Leftrightarrow p = 0.5$ ← we call this the "null hypothesis"

$H_0 : p = 0.5$ What's the alternative?

$H_a : p \neq 0.5$

we will talk more about this soon.

What are we comfortable with in terms of a false error? $\alpha = 5\%$

$$\Rightarrow \text{Acceptance region: } \left[0.5 \pm 2\underbrace{\overbrace{\sqrt{\frac{0.5(1-0.5)}{100}}}^{.10}}_{.05}\right] = [0.4, 0.6] \text{ just missed!}$$

Since $\hat{p} = 0.61 \notin$ Acceptance Region $\Rightarrow$ Reject H_0 / Accept H_a

This "procedure" is called a "2-sided 1-proportion Hypothesis Test"

Let's do some more...

Another example. Gender ratio. Is it even?

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

$H_0 : p = 0.5$

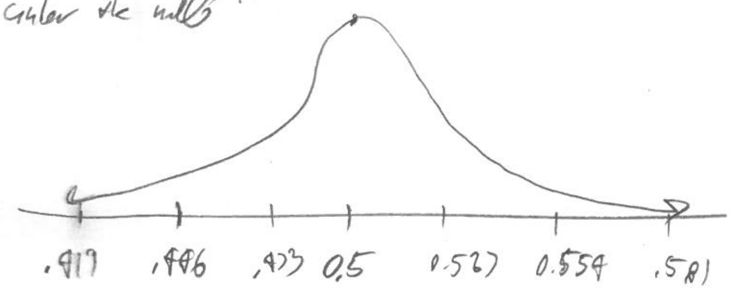
$H_a : p \neq 0.5$

$\alpha = 5\%$ n =

We sample $\underline{\underline{345}}$ children randomly from all hospital births on May 6, 2015

$$\underbrace{\hat{p} \sim N\left(0.5, \left(\sqrt{\frac{0.5(1-0.5)}{345}}\right)^2\right)}_{\text{Sampling proportion "under the null"}} = N(0.5, .0269^2)$$

<u>Leave on board</u>



At $\alpha = 5\%$. Acceptance Region $= [0.5 \pm 2 \cdot .0269] = [.446, .554]$

Do experiment? Count # of males: 169 $\Rightarrow \hat{p} = \frac{169}{345} = 0.49 \in$ AccRegn

$\Rightarrow$ Retain H_0

We do not have evidence to suggest the human sex ratio is not even.

Another example: M&M's (Pass Out)

If the factory is working well the prop of Blues should be $\frac{1}{5}$.

Test this:

$H_0: p = 0.2$

$H_a: p \neq 0.2$

$\alpha = 1\% \Rightarrow z_{\frac{\alpha}{2}} = 2.56$

I've decided you should memorize this one too

Count # of M&M's

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

$= \left[0.2 \pm \frac{1.024}{\sqrt{n}}\right]$

$= [\quad , \quad]$ don't calculate yet

Now do experiment: count # of blues: $\boxed{\text{EAT!}}$

Let's talk about α a little bit...

Our Decision

The truth (unknown)	H_0 true	H_a true
H_0 true	✓	Type I error
H_a true	Type II error	✓

What is Prob? α

We chose this!!!

Look at it on graph...

What's this? Failed to find the effect.. How to choose α?

Kept factory going when it's working	Shut down factory even though it's working
Kept factory going even though it's broken	Shut down factory when it's broken

$\alpha \downarrow \Rightarrow$ Type I error $\downarrow$, Type II error $\uparrow$

$\alpha \uparrow \Rightarrow$ " $\uparrow$, " $\downarrow$

How to decide??

P(Type II error) ... Mostly 282...

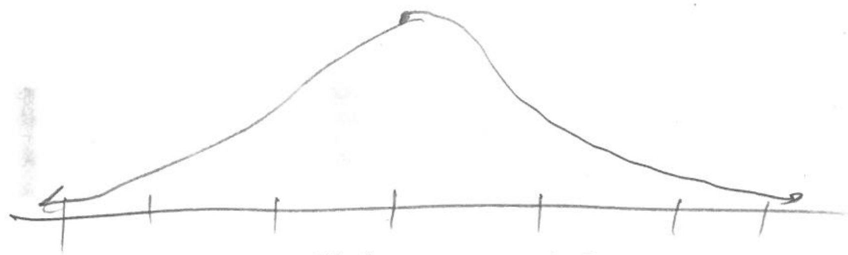
Another wrinkle...

Return ~~to~~ M/F ratio. $H_0: p = 0.5$ $H_a: p \neq 0.5$ $\alpha = 5\%$

Instead of 385, let's take a bigger sample ... in 2008

$n = 4{,}247{,}000$ children born of which $2{,}173{,}000$ male

$$\hat{p} \sim N\left(0.5, \sqrt{\frac{0.25}{n}}\right) = N(0.5, 0.00242^2)$$



Acc region = $[.49516, .50484]$

$$\hat{p} = \frac{2173000}{4247000} = 0.51165 \qquad \hat{p} \notin \text{Acc Region} \Rightarrow \text{Reject } H_0 / \text{Accept } H_a$$

this $\hat{p}$ before would've led to Retaining the H_0.

What happened? With a larger sample size, you have more "power" to detect small deviations from the null.
Here, you can distinguish between $p = 0.50$, $p = .51$,
Before you can do $p = 0.5$, $p = 0.6$

$$n \uparrow \Rightarrow P(\text{Type II error}) \downarrow$$