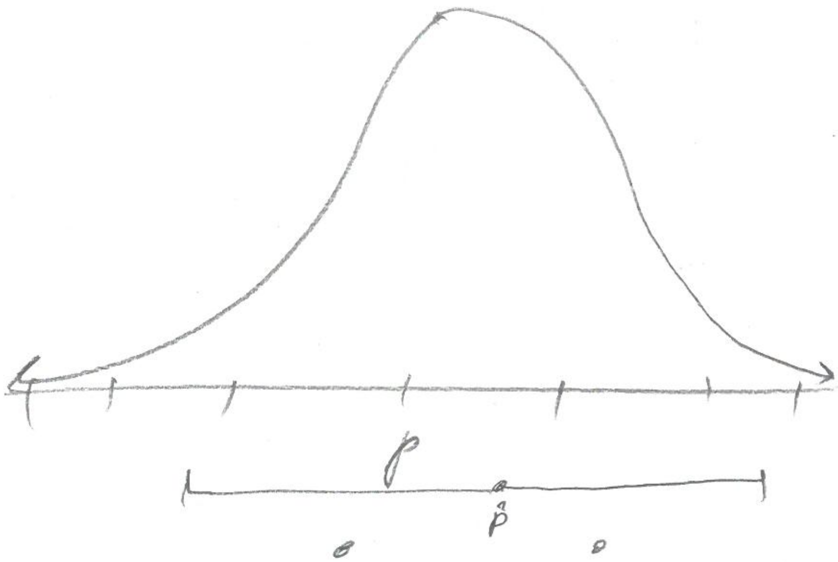
lec 22 12/4/17 Math 241



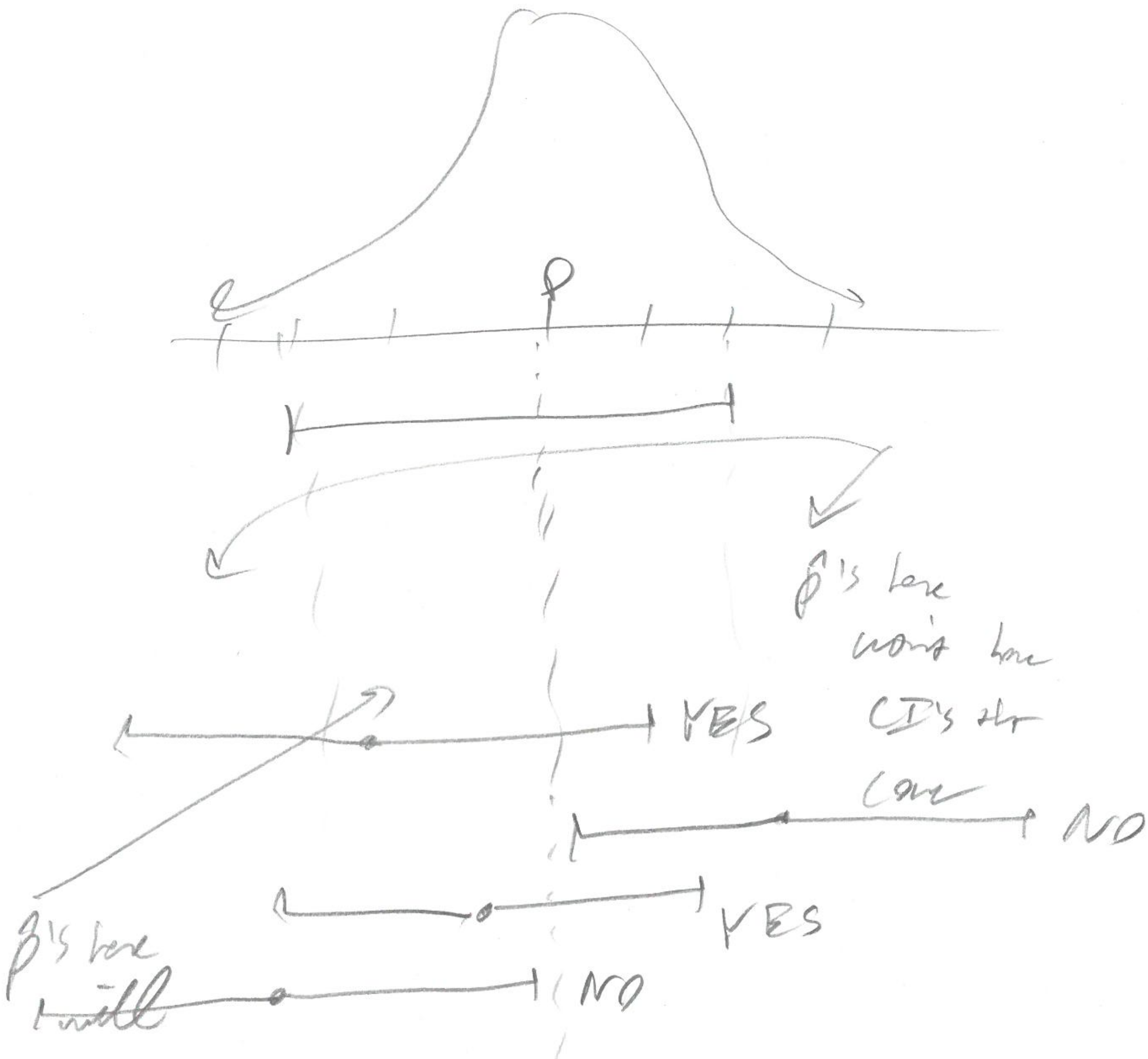
$$\underbrace{P\left(p \in \left(\hat p \pm z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}\right)\right)}_{=1-\alpha} \approx P\left(p \in \left(\hat p \pm z_{\frac{\alpha}{2}}\sqrt{\frac{\hat p(1-\hat p)}{n}}\right)\right) \approx 1-\alpha$$

problem

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

$1-\alpha$ coverage of p, (the parameter). What does this mean?

1. Before you do the procedure
2. If you repeat the procedure many times... → $\mathbb{1}_{p \in CI} = 1$ is deg! Deg(0) Deg(1)
3. $P\left(p \in \left[\hat p \pm z_{\frac{\alpha}{2}}\sqrt{\frac{\hat p(1-\hat p)}{n}}\right]\right)$
4. If you are a subjectivist... with the appropriate prior info $P(p \in CI_{p,1-\alpha}) = 1-\alpha$
   Take Math 341


β
YES
β̂'s here
won't have
CI's that
cover
NO
YES
β̂'s here
will
NO

Who likes mushrooms. Build CI...

Representative ?

---

hmm No?

Is the prop of babies born male ≠ 50%. ? YES?

Let's test if so...

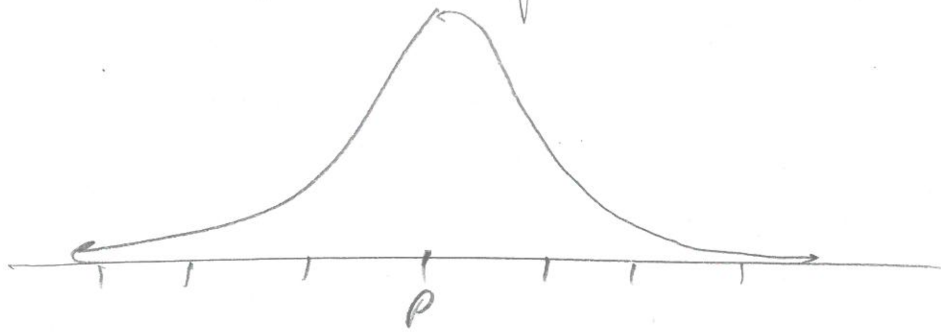
3 Goals of Statistical Inference

---

1. Pt est.
2. Conf In.
3. Testing

the likely simple theory, Occam's Razor is that model $p = 50\%$.

$\hat{p} \mid H_0$ true ($p=0.5$)

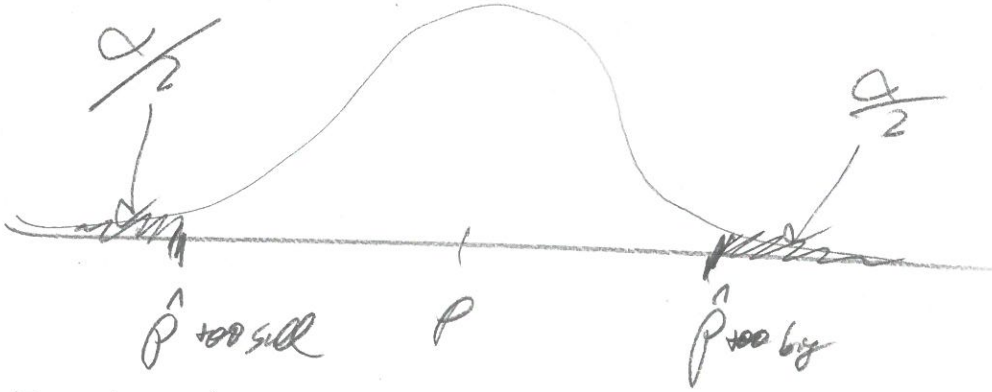


at some pt.. reject...

At what pt? Highway... too fast, too slow. You need to make a decision. Some states 65MPH, 70MPH, 75MPH...

Let $\alpha := P(\hat{p}$ "too rare")

Make one rejection region for $\hat{p}$'s "big" and one for $\hat{p}$'s "too small"



Find these $\hat{p}$'s

from previous class

$$1-\alpha = P\left(\hat{p} \in [\hat{p}_{small}, \hat{p}_{big}]\right) = P\left(\hat{p} \in \left[p \pm z_{\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n}}\right]\right)$$

$\Rightarrow \left[p \pm z_{\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n}}\right]$ is called the Retention Region

$\Rightarrow \left[p \pm z_{\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n}}\right]^c$ is called the Rejection Region

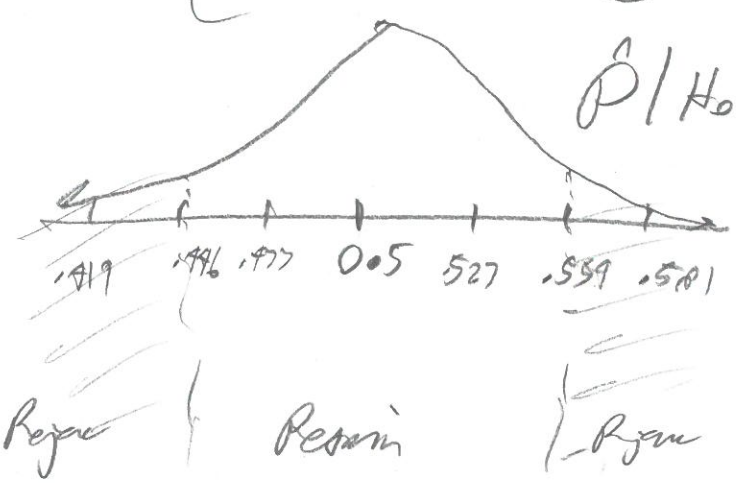
If $\hat{p} \in$ Retention Region $\Rightarrow$ Retain Ho. Not enough evidence to reject default model!

If $\hat{p} \notin$ Retention Region
i.e. $\hat{p} \in$ Rejection region $\Rightarrow$ Reject Ho. Thereby accepting H_a. There is enough evidence to reject default model.

e.g. $H_0: p = 0.5$, $H_a: p \neq 0.5$, $\alpha = 5\%$.
sample: $n = 345$

$\Rightarrow z_{\frac{\alpha}{2}} = z_{2.5\%} = 2$

$$\text{Retention region} = \left[0.5 \pm 2\underbrace{\sqrt{\frac{0.5(1-0.5)}{345}}}_{.0269}\right] = [.446, .554]$$



Actual data: 169 babies male $\Rightarrow \hat{p} = \frac{169}{345} = .48 \in$ Retention Reg

$\Rightarrow$ Retain Ho

Why do we need this??

Flip coin 100 times

H_0: coin is fair $\Rightarrow p_H = 0.5$

H_1: coin is not fair $\Rightarrow p_H \neq 0.5$

Exp 1: coin flips H 51x $\Rightarrow$ fair? Yes!

Exp 2: " " " " " " 98 x $\Rightarrow$ fair? No!

Exp 3: " " " " " " 61 x $\Rightarrow$ fair?

Not so clear!!

Run test at $\alpha = 5\%$

Retention region $= \left[0.5 \pm 2\sqrt{\frac{0.5(1-0.5)}{100}}\right] = [0.4, 0.6]$

$\hat{p} = \frac{61}{100} = 0.61 \notin$ Retention Region $\Rightarrow$ Reject H_0. Coin is not fair!

But what if $\alpha = 1\% \Rightarrow z_{\frac{\alpha}{2}} = 2.58$

Ret. Reg $= \left[0.5 \pm 2.58\sqrt{\frac{0.5(1-0.5)}{100}}\right] = [.371, .629]$

$\hat{p} = 0.61 \in$ Retention Region $\Rightarrow$ Retain H_0! Not enough evidence to say coin is unfair.

Choice of α matters! Must be selected beforehand! We will return to this...

M & M sample.

H_0: 20% are Blue! $p = 0.2$

H_a: Not so $p \neq 0.2$

$\alpha = 5\% \Rightarrow z_{\frac{\alpha}{2}} = 2$

$\frac{1.024}{\sqrt{n}}$

Retention Region $= \left[0.2 \pm 2 \sqrt{\frac{0.2(1-0.2)}{n}} \right]$ Do experiment!

Where does α come from and why should it matter?

You are making a random decision (due to random data). You could be wrong!

DECISION

TRUTH	Retain H_0	Reject H_0
H_0 true	✓	Type I error
H_0 false	Type II error	✓

P(Type I error) $= \alpha$. You choose this!

P(Type II error) ... math 242

P(Reject H_0 | H_0 false) = POWER! Not covered!