December 6, 2016

Statistical Inference

> ① Point Estimation $\Rightarrow$ best guess $= \hat{p}$

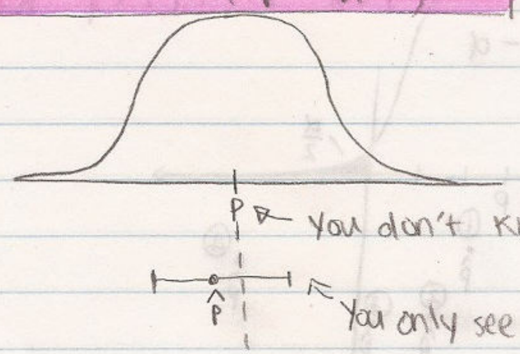
> ② Interval Estimation (range of value for p)

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

Coverage Probability
= How often you catch p in the interval.

* If you make the α too small, your interval will be too large. If you make the α too big, your interval will be too small = useless information.

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



You don't know p in the real world.

You only see this in the real world.

> Interpretation of Confidence Intervals

① If you do many experiments, then

$$\frac{\#\{p \in CI\}}{n} \longrightarrow 1-\alpha$$

Not Relevant: Doesn't matter because you are only doing one experiment.

② Before the experiment

$$P(p \in CI) = 1-\alpha$$

Not Relevant: Doesn't help you after you've done the experiment

③ You want to say $P(p \in CI) = 1-\alpha$ after the experiment.

Not Relevant: Not possible unless you believe in subjective probability and assume prior information on p. (Math 341)

Conclusion: Confidence Intervals don't really say much...

Statistical Inference (continued)

> ③ Testing theories about a parameter (Hypothesis Theory)

> Human Gender Ratio:

I believe $p := P(\text{male})$ is not 50%.

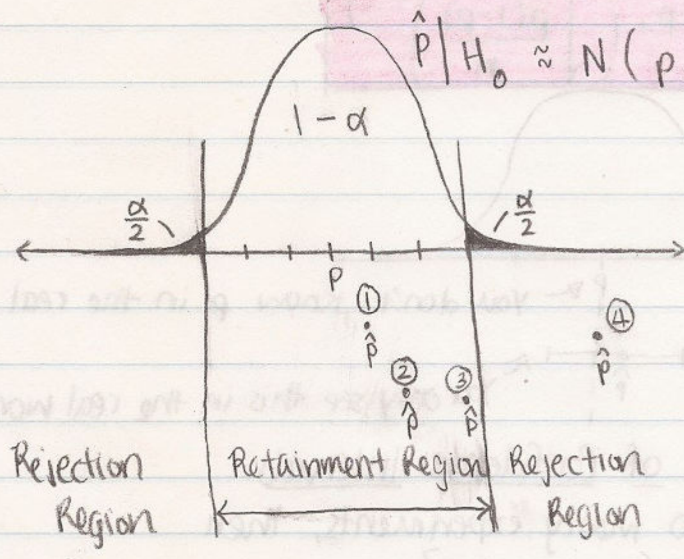
(The probability that a woman gives birth to a male is not 50%.)

Is that crazy? Yes!

> The default position is that $p = 0.5$. We call this the "null hypothesis," denoted as $H_0: p = 0.5$

> The alternate (crazy) hypothesis is denoted $H_a: p \neq 0.5$

$$\hat{p} \mid H_0 \approx N\left(p, \left(\sqrt{\frac{p(1-p)}{n}}\right)^2\right)$$



① and ② are reasonable. They are both close to p and are in the fat middle part of the bell curve.

③ is questionable. Some may say it is reasonable, some may say it's not.

④ is not reasonable. It is way too far from normal and too rare to happen.

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

$1-\alpha := P(\text{not too rare})$

$= P(\text{retain})$

$= P(\hat{p} \in [p \pm \text{margin}])$

$$1-\alpha := P\left(\hat{p} \in \left[p \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n}}\right]\right) \Big\} \text{ Retainment Region}$$

Now we do the experiment and calculate $\hat{p}$. If $\hat{p}$ is in the retainment region, what happened?

★ If $\hat{p} \in$ Retainment Region $\Rightarrow$ Retain H_0. We do not have enough evidence to reject the null hypothesis.

★ If $\hat{p} \notin$ Retainment Region $\Rightarrow$ Reject H_0, accept H_a. We have enough evidence to reject the null hypothesis.

> Human Gender Ratio Problem (continued):

$n = 345$ births $\qquad \alpha = 5\% \qquad Z_{\frac{\alpha}{2}} = 2$ *

$$\text{Retainment Region} = \left[0.5 \pm 2\sqrt{\frac{0.5(1-0.5)}{345}}\right] = [.446, .554]$$

Do the experiment and get 169 males.

$\hat{p} = \frac{169}{345} = 0.48 \in$ Retainment Region $\Rightarrow$ Retain H_0.

We do not have enough evidence to reject equal gender proportions.

> Why do we need Math 241?

Flip a coin 100 times.

Test a theory that the coin is unfair. "Fair" means $P(H) = 0.5$

Scenario I: You get 51 heads. $\Rightarrow \hat{p} = 0.51$.

Is this fair? Yes.

Scenario II: You get 98 heads $\Rightarrow \hat{p} = 0.98$

Is this fair? No. It's too far from 0.5.

Scenario III: You get 61 heads $\Rightarrow \hat{p} = 0.61$

Is this fair? We cannot tell by just looking.

We need Math 241 for this.

$n = 100$ $\alpha = 5\%$

$H_0 := p = 0.5$

$H_a := p \neq 0.5$

Retainment Region $= \left[0.5 \pm 2\sqrt{\dfrac{0.5(1-0.5)}{100}}\right] = [0.4, 0.6]$

$\hat{p} = 0.61 \notin$ Retainment Region $\Rightarrow$ Reject H_0.

We have enough evidence to reject the theory that the coin is fair.

> M&M's experiment

Mars (the candy company) claims that the proportion of blue is 20%. Let $p := P(\text{Blue})$.

$\alpha = 1\% \Rightarrow Z_{\frac{\alpha}{2}} = 2.84$

$H_0 := p = 0.2$

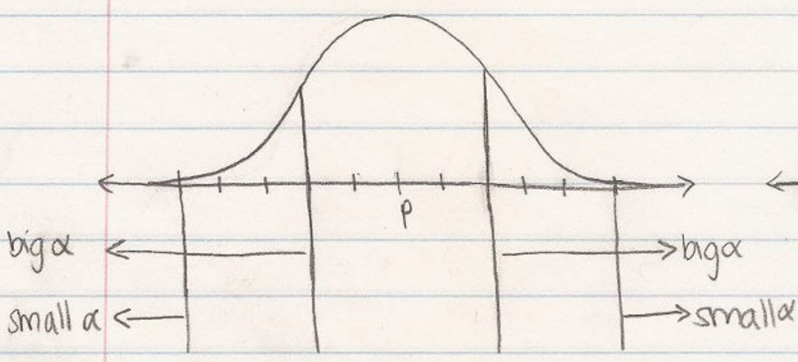
$H_a := p \neq 0.2$

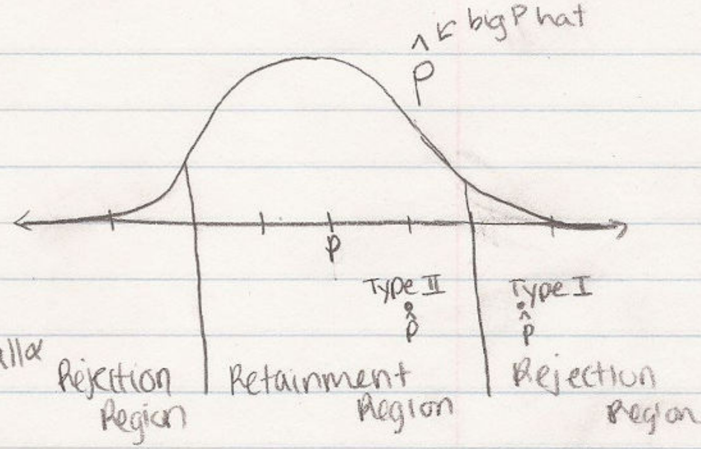
$n = 636$

Retainment Region $= \left[0.2 \pm 2.84\sqrt{\dfrac{0.2(1-0.2)}{636}}\right] = [0.155, 0.245]$

$\hat{p} = \dfrac{168}{636} = 0.264 \notin$ Ret. Region. $\Rightarrow$ Reject H_0. We have enough evidence to reject the claim made by the Mars candy company that the proportion of blue M&Ms is 20%

# of blue





Truth	Decision: Retain H_0	Decision: Reject H_0
H_0 True	✓	Type I Error
H_0 False	Type II Error	✓

So if α is big, there is a smaller retainment region. Big α allows $\hat{P}$ to produce a $\hat{p}$ that, even though is true, gets rejected = Type I. If α is small, there is a larger ret. region. Small α allows $\hat{P}$ to produce a $\hat{p}$ that is false, but gets included in ret. region.

$P(\text{Type I Error}) = P(\text{Reject } H_0 \mid H_0 \text{ True}) = \alpha$

You choose this!

$P(\text{Type II Error})$ ... not covered in this class, but

$P(\text{Reject } H_0 \mid H_0 \text{ False}) =$ POWER!

$\alpha \uparrow \Rightarrow P(\text{Type I Error}) \uparrow \Rightarrow P(\text{Type II Error}) \downarrow$

$\alpha \downarrow \Rightarrow P(\text{Type I Error}) \downarrow \Rightarrow P(\text{Type II Error}) \uparrow$

> Court Case

H_0 := Innocent

H_a := Guilty

Decision: Punish or not.

Type I Error: Punish innocent person

Type II Error: Guilty person goes free.