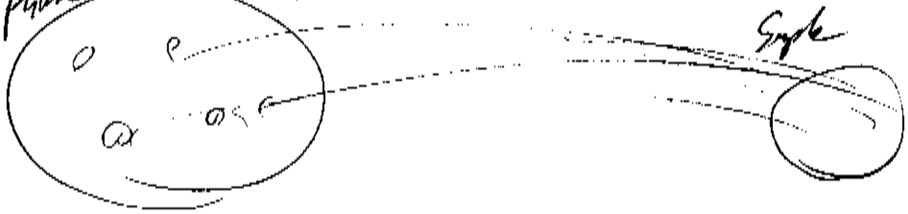
Lecture 21 Dec 9, 2014

Review: $X_1, \ldots, X_N \overset{iid}{\sim}$ Bernoulli(p) $\qquad X_1', \ldots, X_n' \overset{iid}{\sim}$ Bernoulli(p)

Population $\qquad\qquad$ Sample



$N \approx \infty$

$n < \infty$ but big enough for the CLT to kick in

Due to CLT

$$\bar{X} = \frac{X_1 + \ldots + X_n}{n} \sim N\left(\mu, \left(\frac{\sigma}{\sqrt{n}}\right)^2\right)$$

Sample statistic

$$\hat{p} = \frac{\sum x_i}{n} \in [0,1]$$

In the Bernoulli case

$$\bar{X} = \hat{p} = \frac{X_1 + \ldots + X_n}{n} \sim N\left(p, \left(\sqrt{\frac{p(1-p)}{n}}\right)^2\right)$$

↑ Don't know it!

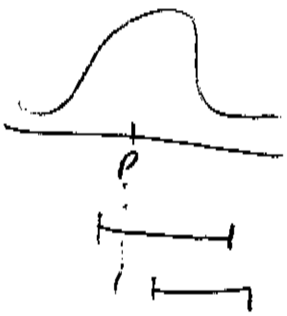
Goal: Infer p. What's best point estimate? $\hat{p}$

Put in a corner, your ~~best~~ guess is your sample stat.

What if we allow for a confidence interval?

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

erased



$1-\alpha = P($includes / covers$)$

true for any individual — ? No one knows

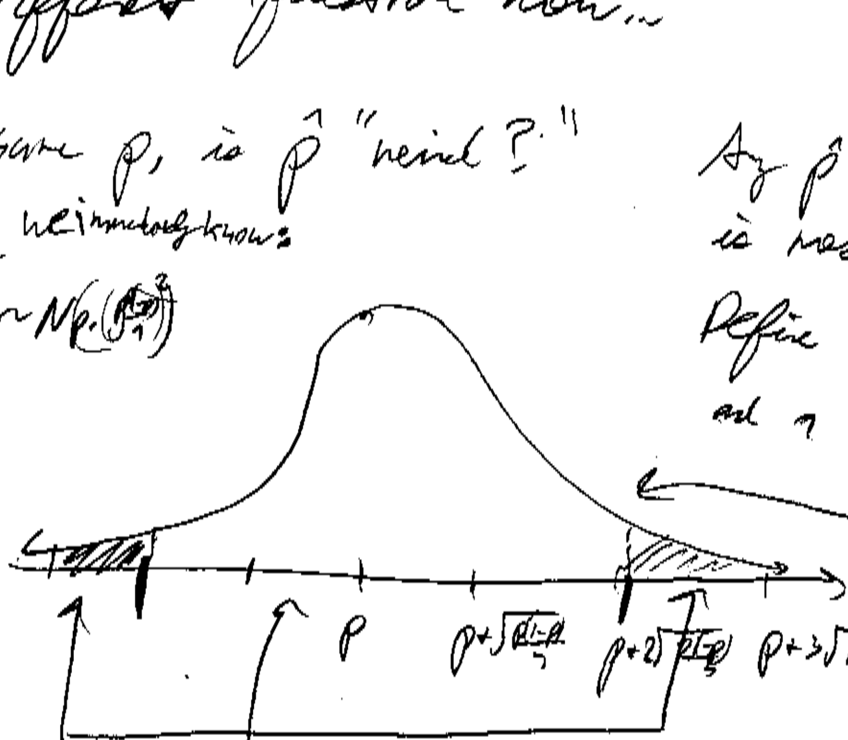
Different question now...

Assume p, is $\hat{p}$ "weird?"

↓ we immediately know:

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

Any $\hat{p}$ can happen. But what is most likely to happen?
Define an acceptance region and a rejection region.



Now, we don't erase. This we assume is the truth!

Rejection region has size → $\alpha = 5\%$. Extreme events either too low or too high.

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

Acceptance region

Acceptance: I think p is reasonable given this $\hat{p}$. $\hat{p} \in \left[p \pm z_{\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n}}\right]$ no longer rely on assumption of $\frac{\hat{p}(1-\hat{p})}{n} \approx \frac{p(1-p)}{n}$

Rejection: I think p is not reasonable given this $\hat{p}$. $\hat{p} \notin \left[p \pm z_{\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n}}\right]$ we know it explicitly

What are these regions for $p = 0.24$, $\alpha = 0.05$ (just from Internet report)?

$$\left[p \pm z_{\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n}}\right] = \left[0.24 \pm 2 \sqrt{\frac{0.24(1-0.24)}{598}}\right] = [.2050, .2750]$$

" " $\alpha = 0.01$? $\subset$ ← Subset (obviously)

$$= \left[0.24 \pm 2.56 \sqrt{\frac{0.24(1-0.24)}{598}}\right] = [.1951, .2849]$$

Conclusion: $\hat{p} \notin$ Acceptance region for $p = 0.24$ at $\alpha = 0.05$

$\hat{p} \in$ Acceptance region for $p = 0.24$ at $\alpha = 0.01$

Need more logical framework... "<u>Hypothesis Testing</u>"

Assume $p = 0.24$. Call this the "null hypothesis" denoted H_0.

We use $\hat{p}$ to see if we have enough evidence to "overthrow" H_0.

If it's overthrown, $p \neq 0.24$ i.e. the "alternative hypothesis" denoted H_a.

$$H_0: p = 0.24$$

$$H_a: p \neq 0.24$$

Construct "Acceptance region" at level α. $\rightarrow$ retain $\neq$ accept

retain means "could be"

accept means "is."

This word is chosen carefully!!

If $p \in$ Acceptance region $\Rightarrow$ "retain H_0"

If $p \notin$ Acceptance region $\Rightarrow$ "reject H_0 and thereby accept H_a"

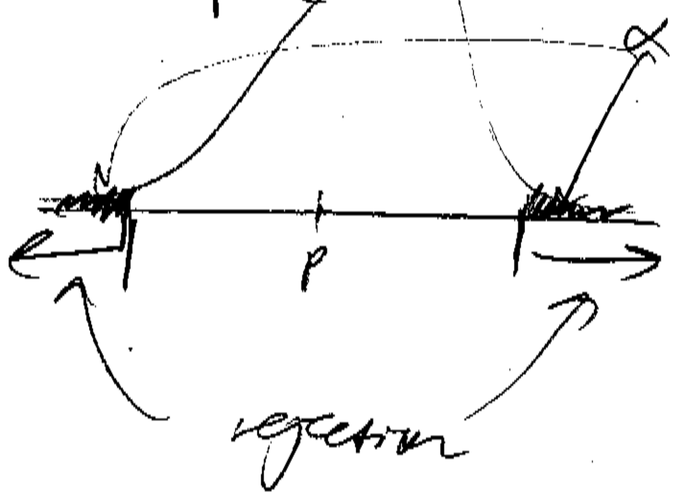
What do we do about α? What is at stake exactly?

... don't know this!
probab. of the case...

	the truth: H_0 true	H_0 false / H_a true
our decision: retain H_0	✓	Type II error
reject H_0 / accept H_a	Type I error	✓

→ Why is this?

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



Data is realization of r.v.'s! Data is random! You make decisions from random data, you may be wrong!!!! No way out!

Why is $P(\hat{p} \in \text{Rejection Region}) = \alpha = P(\text{Type I error})$ by construction!

You are free to choose your $P(\text{Type I error})$!

If you make it too big, you'll reject all the time (false alarms).
" " small, you may not detect abnormalities.

eg H_0: no fire
H_a: fire   What should α be? Small or large?

Type II error: not rejecting when you should. $P(\text{Type II error})$? Complicated... not covered here...

Why do we care? Imagine you want to see if a coin is fair. Hypothesis: $H_0: p = 0.5$, $H_a: p \neq 0.5$

Situation I

Flip coin 100 times. # heads = 52

$n = 100$, $\hat{p} = \frac{\text{Succ}}{n} = \frac{52}{100} = 0.52$

Is $0.52 \stackrel{?}{=} 0.50$ No! Point estimates $\neq$ parameters. $\hat{p} \sim N\left(p, \left(\frac{p(1-p)}{n}\right)^2\right)$

But do you think this is unexpected? NO $\Rightarrow$ Retain H_0, coin is probably fair

Situation II

Flip coin 100 times # heads = 98 $\Rightarrow \hat{p} = 0.98$

Do you think this is unexpected? YES $\Rightarrow$ Reject H_0! Coin is not fair.

Situation III

Flip Coin 100 times # heads = 60 $\Rightarrow \hat{p} = 0.6$

Do you think this is unexpected? ???? Not so easy...

This is why we need hypothesis testing... which comes from CLT... r.v.'s ... prob... everything!!

Gender Birth Data

$P(\text{born male}) \overset{?}{=}$ Assume .50

Is this really true?

Do a SRS. of births across the world? Only found data in USA.

so now $P(\text{born male in USA}) = 0.50$. No inference to whole world!!

$H_0: p = 0.5 \Longrightarrow \hat{p} \sim N\left(p, \left(\sqrt{\frac{p(1-p)}{n}}\right)^2\right)$ this is the "Null Distribution"

$H_a: p \neq 0.5$

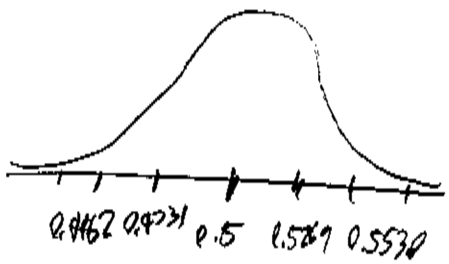
know we will sample $n = 345$ children

$SE(\hat{p}) = \sqrt{\frac{p(1-p)}{n}} = 0.0269$

$\alpha = 5\%$

↑

chance of Type I error I selected.



Acceptance Region: $[0.4462, 0.5538]$

with $\alpha = 0.05$

at $H_0: p = 0.5$

Now do experiment, 169 males

$\hat{p} = \frac{169}{345} = 0.48 \in$ Acceptance Region $\Rightarrow$ Retain H_0. Conclude "Births are most likely even ~~many~~ between the two genders"!

Now sample more. In 2008 $n = 4{,}247{,}000$ babies born in America.

$H_0: p = 0.5$, $H_a: p \neq 0.5$, $\alpha = 0.05$

$$\Rightarrow \hat{p} \sim N\left(p, \left(\sqrt{\tfrac{p(1-p)}{n}}\right)^2\right) = N\left(0.5,\ 0.000242^2\right)$$

$$\text{Acceptance Region} = [p \pm 2SE[\hat{p}]] = [0.4995140,\ 0.5004852]$$

Do experiment

$$\Sigma x_i = 2{,}173{,}000 \Rightarrow \hat{p} = \frac{2{,}173{,}000}{4{,}247{,}000} = 0.5116553$$

$\hat{p} \notin$ Acceptance Region $\Rightarrow$ Reject H_0. Human sex ratio is not 1:1!

Paradoxically, Nobody knows why!!!

March 24 complete!!

"p-value" $\quad p_{val} := P\left(\begin{matrix}\text{this}\\\text{Data or}\\\text{more "extreme"}\\\text{data as is}\\\text{observed}\end{matrix} \,\middle|\, H_0\right)$ $\quad P\left(\frac{\hat{p} - 0.5}{0.000242} > \frac{0.5116553 - 0.5}{0.000242}\right)$

$= P(Z > 2109) = $ ABSOLUTE ZERO

$\Rightarrow$ this could never happen ever due to chance if H_0 is true...

more extreme in this case is being this different from 0.5

$$\text{diff} = |\hat{p} - p| = |0.5116553 - 0.5| = 0.0116553$$

more extreme: $(-\infty, 0.5 - \text{diff}] \cup [0.5 + \text{diff}, \infty)$

$$P(\hat{p} \in \ \ | H_0) = P\left(\hat{p} \in \ \ \middle|\ \hat{p} \sim N\left(p, \left(\sqrt{\tfrac{p(1-p)}{n}}\right)^2\right)\right) = 2P(\hat{p} > 0.5116553)$$

Back to M & m's

H_0: $p = 0.24$

H_a: $p \neq 0.24$

$$\begin{aligned}
\text{pval} &:= P(\text{this data or more extreme} \mid H_0) \\
&= 2\,P(\hat{p} \le 0.1353) \\
&= 2\,P\left(\frac{\hat{p} - 0.24}{\sqrt{\frac{0.24(1-0.24)}{594}}} \le \frac{0.1353 - 0.24}{\sqrt{\frac{0.24(1-0.24)}{594}}} \right) \\
&= 2\,P(Z \le -2.5509) \\
&= 2 \cdot 0.00539 = 0.0108 = 1.08\%
\end{aligned}$$

Another interpretation: $\alpha = 1.08\%$ would fail to reject.