§7.1 Estimating a population Proportion

Goal: Create a Confidence Interval for a population proportion.

Requirements

1. Good Sample SRS
   Simple Random Sample
   Statistically Reliable Sample - No Bias

2. Binomial
   i) $n = \#$ trials ii) $p =$ prob. Success
   iii) Independent iv) Two Outcomes

3. $n\hat{p} > 5$, $n\hat{q} > 5$ so Normal is a good approximation to Binomial.

$x = np =$ Ex. # of Successes = # yes

$nq =$ # of failures = # NO

$\hat{p} = \frac{x}{n}$ = Sample prop. of yes $\hat{q} = 1 - \hat{p}$ = Sample prop. of NO

Spring Break $X = 11$ = # of people going out of Tow[n]

$n = 24$

Point estimate $= \hat{p} = \frac{11}{24} = .46 = .458 =$ Sample Statistic

Population = All SRJC Students

What proportion of SRJC Students are going out of town?

Best estimate is .46

How good is that estimate?

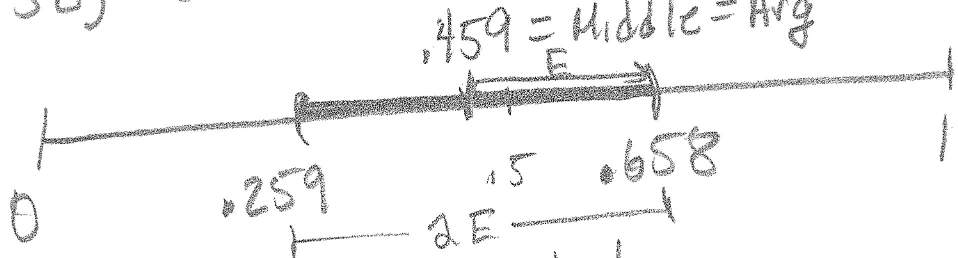
A 1-Prop Z Interval or a one proportion Confidence Interval is an interval of Values for which we are CL = 95% (is the Confidence level) Confident Contains the True proportion of SRJC students who will leave town.

$CI = (.259, .658) = (a, b)$

the True proportion is between 26% and 66% will leave Town

There are 3 ways to write a Confidence Interval

① $(.259, .658)$ Interval Notation



② $.259 < p < .658$ Inequality

$\uparrow$ Population proportion

③ $\hat{p} \pm E = .458 \pm E$

E = the Margin of Error

$$\hat{p} = \frac{LB + UB}{2} = \frac{.259 + .658}{2} = .459$$

$$E = UB - \hat{p} = .658 - .459 = .199$$

$$E = \frac{UB - LB}{2} = .199$$

As n increases, E decreases.

How do we find E? if given $n = 24$ $x = 11$

CLT for Proportions Gives $\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}$

The Standard Error for Sample $= \sqrt{\frac{\hat{p}\hat{q}}{n}}$

95% CI $\Rightarrow$ $Z_{.05/2}$ = invnorm$(1 - .05/2, 0, 1)$
$= 1.96 = Z_{.025}$

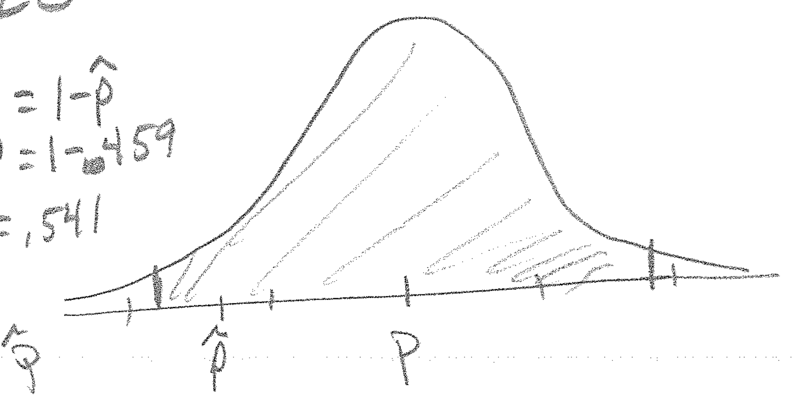
"Z sub Alpha over 2"

$\alpha = 1 - CL = 1 - .95 = .05$

$\alpha/2 = \frac{.05}{2} = .025$

$$\boxed{E = Z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}}$$

$\hat{q} = 1 - \hat{p}$
$\hat{q} = 1 - .459$
$\hat{q} = .541$



$$E = 1.96\sqrt{\frac{.459 \cdot .541}{24}}$$

$\hat{p} = \frac{x}{n} = \frac{11}{24}$

$$\boxed{E = .199}$$

§6.5 The Central Limit Theorem for Proportions

7.2

① Distribution of Sample proportions is approximately Normal

$n\hat{p} \geq 9 \quad n\hat{q} \geq 9$

More than 9 successes and 9 failures

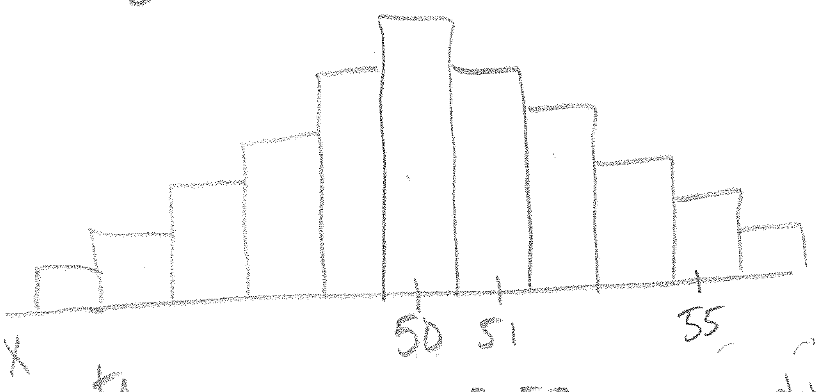
② $\mu_{\hat{p}} = p$

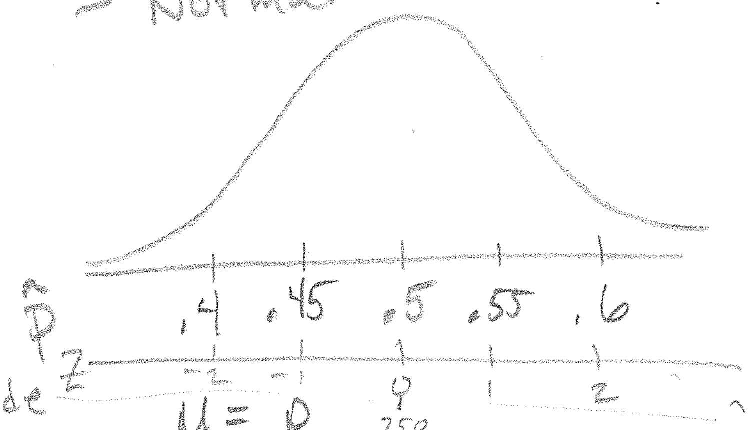
③ $\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}$

$X = n\hat{p}$

$\hat{p} = \frac{x}{n}$

Binomial → Normal





$\mu = np = 100 \cdot .5 = 50$

$\sigma = \sqrt{npq} = \sqrt{100 \cdot .5 \cdot .5}$

$\sigma = 5$

divide by n →

$\mu = p$

$\sigma_{\hat{p}} = \frac{\sqrt{npq}}{n} = \sqrt{\frac{pq}{n}} = .05$

$\sigma_{\hat{p}} = \sqrt{\frac{.5 \cdot .5}{100}} = .05$

Ex If a coin is Tossed 100 times what proportion of the time do we get more than 54 heads

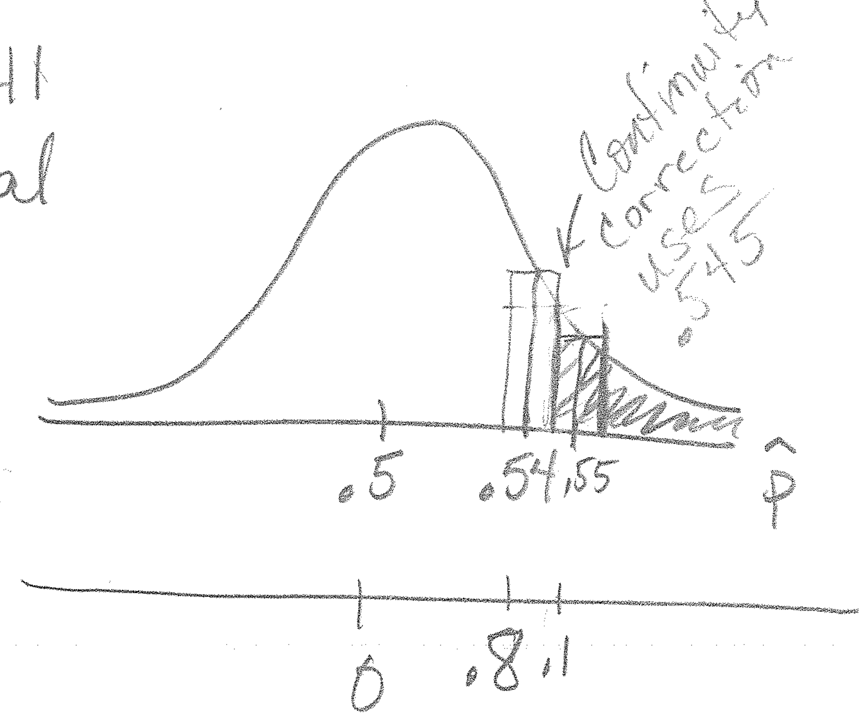
$$P(X > 54) = 1 - P(X \le 54)$$

$$= 1 - \text{binomialcdf}(n, p, r)$$

$$= 1 - \text{binomialcdf}(100, .5, 54)$$

$$= .1841$$

Now with Normal



$$Z = \frac{x - \mu}{\sigma}$$

$$Z = \frac{\hat{p} - .5}{\sqrt{\frac{pq}{n}}}$$

$$Z = \frac{.55 - .5}{\sqrt{\frac{.5 \cdot .5}{100}}} = 1$$

.545

$= .05$

$$P(\hat{p} \ge .55) = \text{normalcdf}(.55, 999, .5, .05)$$

$$= .1587$$

with continuity correction

$$P(\hat{p} \ge .545) = \text{Norcdf}(.545, 999, .5, .05)$$

$$= .18406009$$

§7.1b Estimates of a Population Proportion

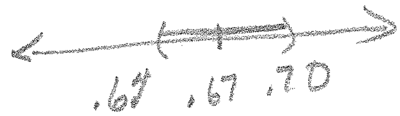
a Sample

Ex 1 The Press Democrate reports that 67% of voters plan to vote in the next election with a margin of error of ± 3% with 95% confidence

a) Write confidence interval all 3 ways.

$\hat{p} = .67 \qquad E = .03$

$\hat{p} \pm E = .67 \pm .03$



$(\hat{p} - E, \hat{p} + E) = (.64, .70)$

$\hat{p} - E < p < \hat{p} + E \qquad .64 < p < .70$

b) Write what this means in words?

Inference

We are 95% confident that the true proportion of Santa Rosans who plan to vote is between .64 and .70.

c) How large was their sample?

$E = z_{\alpha/2} \cdot \sqrt{\dfrac{\hat{p}\hat{q}}{n}}$

c) How large was sample size n?

$$\boxed{E = Z_{\alpha/2}\sqrt{\frac{\hat{p}\,\hat{q}}{n}}}$$

Solve for n

$$\frac{E}{Z_{\alpha/2}} = \sqrt{\frac{\hat{p}\,\hat{q}}{n}}$$

$$\frac{E^2}{(Z_{\alpha/2})^2} = \frac{\hat{p}\,\hat{q}}{n}$$

$$E^2 \cdot n = \hat{p}\,\hat{q}\cdot(Z_{\alpha/2})^2$$

★ $$\boxed{n = \frac{Z_{\alpha/2}^2 \cdot \hat{p} \cdot \hat{q}}{E^2}}$$

Given $\boxed{\hat{p} = .67}$

$\boxed{E = .03}$ $\alpha = 1 - CL$

$CL = .95$ $\alpha = .05$, $\alpha/2 = .025$

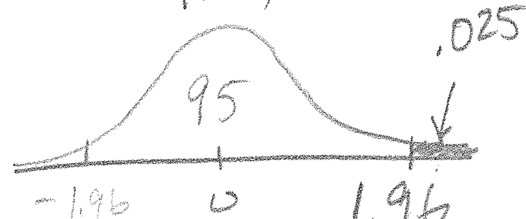
$\hat{q} = 1 - \hat{p} = 1 - .67$

$\boxed{\hat{q} = .33}$

$$n = \frac{1.96^2 \cdot .67 \cdot .33}{(.03)^2}$$

$Z_{\alpha/2} = Z_{.025} = \text{invnorm}(1 - .025$

$= 1.96$



$n = 943.7$

$\boxed{n = 944}$

$\alpha =$

Ex If we want to estimate the proportion of homes with LED light bulbs How ~~many~~ large a sample is Needed to have a margin of error less than 5% for a 95% CI?

$$n = \frac{Z_{\alpha/2}^2 \; \hat{p} \cdot \hat{q}}{E^2} \quad \leftarrow \hat{p} \text{ is Not given}$$

Note: $\hat{p} \cdot \hat{q} \le \frac{1}{4}$

$A = \hat{p} \cdot \hat{q} = \hat{p}(1-\hat{p})$

$A = -\hat{p}^2 + p$

$\hat{p} = \frac{-b}{2a} = \frac{-1}{2(-1)} = \frac{1}{2}$

$$n = \frac{Z_{\alpha/2}^2 \cdot .25}{E^2} = \frac{1.96^2 \cdot .25}{(.05)^2} = 384.16$$

Round up $\boxed{n = 385 \text{ people}}$

Ex In a Study $111/240 = \hat{p}$ Randomly Selected homes had LED's in Santa Rosa. Make a 93% Confidence interval for the proportion of SR residence with LED's.

(1) Requirements SRS ✓ $n\hat{p} = 240 \cdot \frac{111}{240} = 111 = x$ # of Successes > 5

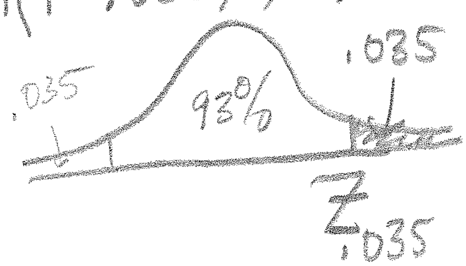
$n\hat{q} = 129 > 5$

(2)

$$E = Z_{\alpha/2}\sqrt{\frac{\hat{p}\,\hat{q}}{n}} \leftarrow \text{proportions}$$

(3) $\boxed{\hat{p} = \frac{111}{240} = .4625}$ $\boxed{\hat{q} = 1 - \hat{p} = .5375}$ $\boxed{n = 240}$

$$Z_{\alpha/2} = Z_{.07/2} = Z_{.035} = \text{invnorm}(1 - .035, 0, 1) = 1.81$$

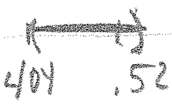


$CL = .93$ $\alpha = 1 - CL = .07$ $\frac{\alpha}{2} = .035$

$\boxed{Z_{\alpha/2} = 1.81}$

Plug them in $E = 1.81\sqrt{\frac{.4625 \cdot .5375}{240}} = .0583$

(4) $\hat{p} \pm E = .4625 \pm .0583 =$



$(\hat{p} - E, \hat{p} + E) = (.4042, .5208)$

(5) Interpret We are 93% Confident that "the proportion of homes in SR with LED's" is between .4042 & .5208

^ Describe Population

$.4042 < p < .5208$

Quiz 8 7.2 Math 15 Jones Name________________________________

A statistics professor asked her students whether or not they were registered to vote. In a sample of 50 of her students (randomly sampled from her 700 students), 30 said they were registered to vote.

1) (20 Points) Find a 98% confidence interval for the true proportion of the professor's students who were registered to vote. (Make sure to check any necessary conditions and to state a conclusion in the context of the problem.)

a) What is the sample and the population in this problem.

b) What point estimate of the population proportion p does this survey give?_______

c) What is the level of confidence for this interval?________ Find α_______

d) What is the critical value use it's symbol

e) Find the margin of error?

Show formula and all values you use to find this by hand. CV:_________ E=___________

f) Find the confidence interval.

____

2) Assuming that this class is representative of the population of all statistics studetents. What is the probability that the true proportion of the statistics students who were registered to vote is in your confidence interval?

3) According to a Gallup poll, about 73% of 18– to 29–year–olds said that they were registered to vote. Does the 73% figure from Gallup seem reasonable for the professor's students? Explain.

4) If the professor only knew the information from the Gallup poll and wanted to estimate the percentage of her students who were registered to vote to within ±3% with 98% confidence, how many students should she sample?

Steps for finding a Confidence Interval

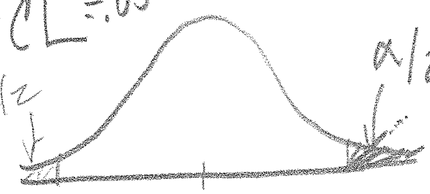
① Find $\hat{p} = \frac{x}{n}$ = point estimate

PE

② Find critical value. Given a Confidence Level, CL like 95% the $\alpha = 1 - CL = .05$

CV

CV: $Z_{\alpha/2} = \text{invnorm}(1 - \alpha/2, 0, 1) =$



$Z_{\alpha/2} = \text{invnorm}(1 - .05/2, 0, 1) = 1.96$

③ Find Margin of Error, E

E

$$E = Z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}\,\hat{q}}{n}} \cong 2 \cdot \sigma_{\hat{p}} =$$

④ Find Confidence Interval

CI $\hat{p} \pm E$, $\hat{p} - E < p < \hat{p} + E$, $(\hat{p} - E, \hat{p} + E)$

⑤ Sentence We are 95% confident that the proportion of people who will vote is between .64 and .70.

Procedure For Constructing a Confidence Interval for p.

① Verify that requirements are Satisfied
① $np > 5$ and $nq > 5$, and ② SRS and ③ binomial

② Find Correct formula on formula sheet for CI for one proportion.

$$E = z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}\hat{q}}{n}} \qquad CI: \hat{p} \pm E$$

③ Find $\hat{p}$, $\hat{q}$, n, and $z_{\alpha/2}$

$$\hat{p} = \frac{x}{n} = \frac{\text{\# of Successes}}{\text{Sample Size}} = \text{Sample proportion}$$

$$\hat{q} = 1 - \hat{p} = \frac{n-x}{n} = \frac{\text{\# of failures}}{\text{Sample Size}} = \text{Sample prop. of failures}$$

$z_{\alpha/2}$ = Invnorm$(1-\alpha/2, 0, 1)$ is the z-score that corresponds to an area of $\alpha/2$ in right tail of Z-distributi

④ Find E and the CI $\hat{p} \pm E$ or $(\hat{p} - E, \hat{p} + E)$ or $\hat{p} - E < p < \hat{p} + E$

round to 3 significant digits

Check with TI

⑤ Interpret Confidence Interval.

The Confidence Interval

for the population proportion p is given by

$$\hat{p} - E < p < \hat{p} + E \quad \text{where} \quad E = z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

Other Expressions

Interval Notation

$\hat{p} \pm E$, $(\hat{p} - E, \hat{p} + E)$

Round to 3 significant digits. (when all calculations are done.)

Example) Create a 95% Confidence Interval for the proportion of the cities middle schoolers who have been drunk.

1) Find CI for one proportion Formula on Formula Card

$$E = z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

2) Determine values for each of the unknowns needed

$\hat{p} = \frac{21}{110} \approx .19$ $\hat{q} = \frac{89}{110} \approx .81$ $n = 110$

$\alpha = 1 - \text{Confidence level} = 1 - .95 = .05$

$\alpha/2 = .025$

$z_{\alpha/2} = 1.96$ From Table

<u>Ex</u> A study Involves 634 randomly Selected deaths. 32 were caused by accidents Construct a 98% confidence interval for the true prop of all deaths Caused by accidents.

$\boxed{n=634,\ X=32 \quad CL=.98 \quad \alpha=.02}$

a) Population = All deaths
Sample = the 634 death in this study.

b) Find $Z_{\alpha/2} = \text{invnorm}\,(1-.02/2) = \boxed{2.33 = Z_{\alpha/2}}$

c) Find $E = Z_{\alpha/2}\sqrt{\dfrac{\hat{p}\,\hat{q}}{n}} = 2.33\cdot\sqrt{\dfrac{.0505\cdot .9495}{634}} = .0202$ $\boxed{E = .0202}$

d) Find point Estimate $= \hat{p} = \dfrac{x}{n} = \dfrac{32}{634} =$
For proportion of death caused by accidents $\boxed{\hat{p} = .0505}$

$\hat{q} = 1-\hat{p} = 1-.0505 = .9495$

d) Find CI $\quad \hat{p} - E < p < \hat{p} + E$

$.0505 - .0202 < p < .0505 + .0202$

$\boxed{.0303 < p < .0707}$

e) Interperate.
We are <u>98%</u> confident that the True proportion <u>of death that were Caused by accidents</u> is between <u>.0303</u> and <u>.0707</u>.