Confidence Intervals for p [cont'd]

M358K: 11/26/2018.

Let p denote the probab. of success in every trial. It's our parameter of interest!

$\hat{P}$ ... statistic of interest, i.e., the proportion of successes in our sample

$$\hat{P} = \frac{\#\text{ of successes}}{\text{sample size}}$$

n ... sample size

$\Rightarrow$ # of successes $\sim$ Binomial(# of trials $= n$, probab. of success $= p$)

For large n: we have the approximate dist'n of # of successes, i.e.,

# of successes "$\sim$" Normal(mean $= np$, var $= np(1-p)$)

$\Rightarrow$ $\hat{P}$ "$\sim$" Normal$\left(\text{mean} = p, \text{var} = \frac{p(1-p)}{n}\right)$

Confidence interval @ the confidence level C

$$\text{pt. estimate} \pm \underbrace{\text{margin of error}}_{z^* \cdot \text{std error}}$$

$$\hat{p} \pm z^* \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Q: What is the smallest sample size necessary so that the m.e. is smaller than a given value d ?

We want : $z^* \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq d$ ⍨

<u>Problem</u>: we don't have $\hat{p}$!

One option : use a prior pt. estimate for the sample proportion from a prior study!

<u>Call it</u> $\tilde{p}$.

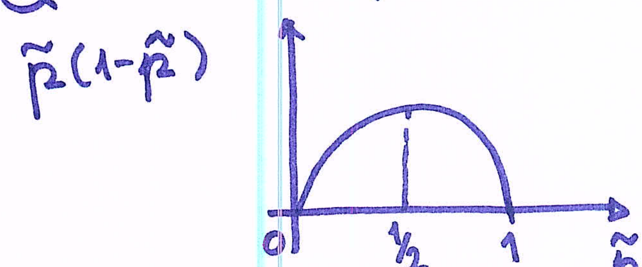
Then, our condition is

$$z^* \cdot \sqrt{\frac{\tilde{p}(1-\tilde{p})}{n}} \leq d$$

$$\frac{z^* \sqrt{\tilde{p}(1-\tilde{p})}}{d} \leq \sqrt{n} \quad /^2$$

$$\frac{(z^*)^2 \, \tilde{p}(1-\tilde{p})}{d^2} \leq n$$

Q : What if we don't have access to a $\tilde{p}$?



The conservative choice for $\tilde{p}$ is ½.

Our condition becomes:

$$\frac{(z^*)^2}{4d^2} \leq n$$

EXAMPLE 5.5.6

Suppose that a local TV station in a city wants to conduct a survey to estimate
ident's policies on economy within 3% error with 95% confidence.

(a) How many people should the station survey if they have no information ?
(b) Suppose they have an initial estimate that 70% of the people in the city
policies of the president. How many people should the station survey?

(a) We use $\tilde{p} = \frac{1}{2}$ as a conservative choice :

$$n \geq \frac{(z^*)^2}{4d^2} = \frac{(1.96)^2}{4(0.03)^2} = 1{,}067.1\ldots$$

z^* corresponding to $C = 0.95$ is 1.96

$$\boxed{n = 1068}$$

(b) $$n \geq \frac{(z^*)^2\,\tilde{p}(1-\tilde{p})}{d^2} = \frac{(1.96)^2 \cdot 0.7 \cdot 0.3}{(0.03)^2}$$

$$n \geq 896.37 \Rightarrow \boxed{n = 897}$$

UNIVERSITY OF TEXAS AT AUSTIN

Problem Set # 10

One-sample proportion inference.

Problem 10.1. A simple random sample of 100 bags of tortilla chips produced by Company X is selected every hour for quality control. In the current sample, 18 bags had more chips (measured in weight) than the labeled quantity. The quality control inspector wishes to use this information to calculate a 90% confidence interval for the true proportion of bags of tortilla chips that contain more than the label states. What is the value of the standard error of $\hat{p}$?

$$\text{std error} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.18 \cdot 0.82}{100}} = 0.03842$$

Just for fun: $z^* = 1.645 \Rightarrow$ m.e. $= 1.645 \cdot 0.03842 = 0.063$

90%-Confidence interval: 0.18 ± 0.063

Problem 10.2. A simple random sample of 60 blood donors is taken to estimate the proportion of donors with type A blood with a 95% confidence interval. In the sample, there are 10 people with type A blood. What is the margin of error for this confidence interval?

$$\text{std error} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{1/6\,(5/6)}{60}}$$

$z^* = 1.96$

m.e. $= z^* \times$ std error $= 0.0943$

Problem 10.3. A simple random sample of 85 students is taken from a large university on the West Coast to estimate the proportion of students whose parents bought a car for them when they left for college. When interviewed, 51 students in the sample responded that their parents bought them a car. What is a 95% confidence interval for p, the population proportion of students whose parents bought a car for them when they left for college?

$$\hat{p} = \frac{51}{85} = \ldots$$

$$\Rightarrow \text{std error} = \sqrt{\frac{\frac{51}{85} \cdot \frac{34}{85}}{85}} = \frac{1}{85}\sqrt{\frac{51 \cdot 34}{85}}$$

$z^* = 1.96$

m.e. $= z^* \cdot$ std error $= 0.1041$

$$p = \hat{p} \pm \text{m.e.} = 0.6 \pm 0.1041$$

Problem 10.4. A simple random sample of 450 residents in the state of New York is taken to estimate the proportion of people who live within 1 mile of a hazardous waste site. It was found that 135 of the residents in the sample live within 1 mile of a hazardous waste site.

(1) What are the values of the sample proportion of people who live within 1 mile of a hazardous waste site and its standard error?

$$\hat{p} = \frac{135}{450} = 0.3$$

$$\text{std error} = \sqrt{\frac{0.3 \cdot 0.7}{450}} = 0.0216$$

(2) What are the values of the sample proportion of people who live outside of the 1 mile radius around a hazardous waste site and its standard error?

$$\hat{p}' = 0.7$$

$$\text{std error} = \sqrt{\frac{0.3 \cdot 0.7}{450}}$$ ↑ the same :)

(3) Do you notice something interesting about the above?

The width of the conf. interval for successes & failures is the same.

Problem 10.5. Sample size

The *Information Technology Department* at a large university wishes to estimate the proportion p of students living in the dormitories who own a computer. They want to construct a 90% confidence interval. What is the minimum required sample size the IT Department should use to estimate the proportion p with a margin of error no larger than 2 percentage points?

$d = 0.02$ $z^* = 1.645$

$$n \geq \frac{(z^*)^2}{4d^2} = \frac{(1.645)^2}{4 \cdot (0.02)^2} = 1{,}691.27$$

$\Rightarrow$ min. sample size is 1,692.

Problem 10.6. (5 points)

You want to design a study to estimate the proportion of people who strongly oppose to have a state lottery. You will use a 99% confidence interval and you would like the margin of error of the interval to be 0.05 or less. What is the minimal sample size required?

a. 666
b. 543
c. 385
d. Not enough information is provided.
e. None of the above.
