Confidence Intervals for p.

Let p denote the probability of success in every single trial, i.e., w/ probability p, a randomly chosen person from your population has a particular trait (e.g., they will vote purple).

We plan to gather a sample of size n.
Then, X... count random variable,
i.e., we say this is the # of successes
$\Rightarrow$ Our the sampling distribution of the counts is $X \sim Binomial($# of trials $= n$, probab. of success $= p)$

Our unknown parameter of interest!

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

$$\Rightarrow \boxed{\hat{p} = \frac{X}{n}}$$

For "large" n, i.e., w/ $\boxed{np > 10}$ and $\boxed{n(1-p) > 10}$, we can use the normal approximation to the binomial, i.e.,

$$\boxed{X \text{ "}\sim\text{" Normal}\left(\text{mean} = np,\ \text{var} = \underline{np(1-p)}\right)}$$

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

Let $\underline{C}$ be our $\underline{\text{confidence level}}$.

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

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

observed sample proportion

UNIVERSITY OF TEXAS AT AUSTIN

Problem Set # 12

Confidence intervals: One-sample proportion.

Problem 12.1. 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 $\hat{p}$?

$\hat{p} = 18/100$

$$\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.18 \cdot 0.82}{100}} = \underline{\underline{0.03842.}}$$

Problem 12.2. 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?

$$z^* \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 1.96 \cdot \sqrt{\frac{(1/6)(5/6)}{60}} = \underline{\underline{0.0943}}$$

$\hat{p} = \frac{10}{60}$

Problem 12.3. 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?

$$\hat{p} = \frac{51}{85} = 0.6$$

$$\text{margin of error} = z^* \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} =$$

$$= 1.96 \sqrt{\frac{\left(\frac{51}{85}\right)\left(\frac{34}{85}\right)}{85}} =$$

$$= 1.96 \sqrt{\frac{(0.6)(0.4)}{85}} = 0.1041$$

$$\Rightarrow \quad \underline{\underline{p = 0.6 \pm 0.1041}}$$

Q: What is the smallest sample size necessary so that the margin of error is smaller than a given value m?

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

Problem: We don't have $\hat{p}$!

One option: Use a previous study's estimate instead; call it $\tilde{p}$.

Two option: The conservative choice for what to substitute for $\hat{p}$ is $1/2$ (because it maximizes the variance of the Bernoulli trial):

$$z^* \cdot \sqrt{\frac{(1/2)(1/2)}{n}} \leq m$$

$$(z^*)^2 \cdot \frac{1}{4n} \leq m^2$$

$$\boxed{\frac{(z^*)^2}{4 \cdot m^2} \leq n}$$

Problem 12.4. 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?

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

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

(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?

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

stderror = the same

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

std errors are the same.

Problem 12.5. 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 costruct 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?

Problem 12.6. (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.