Chapter 21

The Accuracy of Percentages

Confidence Intervals

Example

There are nearly 25,000 registered voters in the county. Use a simple random sample of size 400 to estimate the percentage of registered voters who favor the sale of municipal bonds for the construction of two new elementary schools.

1. Is this a chance process? What is the box model?

Yes

$\boxed{?\ 1s,\ ?\ 0s}$ → Draw 400 & consider the % 1s drawn.

2. It turns out that 55% of the voters in your sample favor the sale of bonds. What is your estimate for the entire population?

55%

3. How accurate is your estimate? Could another sample yield a different result?

Don't know yet. Yes

4. If you knew the box SD, could you find the SE for the percentage of 1's drawn?

Yes

5. Can you estimate the box SD?

Yes, use the sample SD

$$\text{Box SD} \approx \sqrt{(.55)\times(.45)} = .497$$

6. Use your estimate of the Box SD to find (approximately) the SE for the percentage of 1's drawn?

$$\text{SE for \% 1s} = \frac{\text{Box SD} \times \sqrt{400}}{400} \times 100\%$$

$$\approx \frac{(.497)\,20}{400} \times 100\% = 2.485\,\%$$

About 2.5%

7. Build a 95% confidence interval for the true population percentage in favor of bonds.

55% ± 2 SEs

55% ± 2 (2½ %)

55% ± 5%

(50%, 60%)

8. Build a 68% confidence interval for the true population percentage in favor of bonds.

55% ± 1 SE

55% ± 2.5%

(52.5%, 57.5%)

9. If each student in the class were to sample 400 voters and construct a 68% confidence interval, about how many of these intervals would actually contain the true population percentage in favor of bonds?

about 68%

The Bootstrap

When we do not know what is in the box, we estimate the SD of the box by using the SD of the sample.

Confidence Intervals

A 95% confidence interval for the population percentage is given by

Sample percentage ± 2 (SE for %)

The confidence interval is valid if the number of draws is large enough.

For a confidence interval with a different confidence level, normal tables are used to find the multiplier.

An 80% confidence interval is

Sample percentage ± 1.3 (SE for %)

A 90% confidence interval is

Sample percentage ± 1.65 (SE for %)

Example. A health inspector takes a random sample of 300 10-year-olds in a city, and finds that 73% of them have had chicken-pox. Find a 90% confidence interval for the percentage of 10-year-olds in the city who have had chicken-pox.

$\boxed{?\ 1s,\ ?\ 0s}$ → Draw 300 & consider the % 1s drawn.

Box SD ≈ sample SD $= \sqrt{(.73)(.27)} = .44$

sample % $\pm$ (1.65) SE for %

73% $\pm$ (1.65) SE for %

$$\text{SE for \%} = \frac{\text{Box SD} \times \sqrt{300}}{300} \times 100\%$$

$$\approx \frac{(.44)\sqrt{300}}{300} \times 100\% = 2.54\%$$

73% $\pm$ (1.65)(2.54%)

73% $\pm$ 4.2%

Example. Senator Smith wants to know if she should seek re-election. From a population of 1,000,000 voters, a simple random sample of size 2500 is taken. In the sample, 53% of the voters favor Senator Smith. As her statistical consultant, what should you tell her?

Let's find a 95% confidence interval.

$\boxed{?\ 1\text{s},\ ?\ 0\text{s}}$ → Draw 2500 & consider the % 1s drawn.

Box SD $\approx \sqrt{(.53)(.47)} = .499$

SE for % 1s $= \dfrac{\text{Box SD} \times \sqrt{2500}}{2500} \times 100\%$

$\approx \dfrac{(.499) \times 50}{2500} \times 100\% = .998\%$

53% ± 2 SEs

$53\% \pm 2\,(.998)\%$

$\boxed{53\% \pm 2\%}$