Constructing confidence intervals for population proportion p:

(i) Choose confidence level — the most common choices are 90%, 95% and 99%, with 95% being the most common of all.

(ii) Find the *critical value* associated with the chosen confidence level.

- For $0 < \alpha < 1$, the critical value $z_\alpha > 0$ is the number such that

$$P(Z > z_\alpha) = P(Z < -z_\alpha) = \alpha,$$

where $Z \sim N(0, 1)$ is the standard normal random variable.

- If the confidence level is $1 - \alpha$, then the critical value associated with this confidence level is $z_{\alpha/2}$.
  - For 90% confidence, $\alpha = 0.1$ and $z_{\alpha/2} = z_{0.05} = 1.645$.
  - For 95% confidence, $\alpha = 0.05$ and $z_{\alpha/2} = z_{0.025} = 1.96$.
  - For 99% confidence, $\alpha = 0.01$ and $z_{\alpha/2} = z_{0.005} = 2.575$.

(iii) Collect a simple random sample and find the sample proportion $\hat{p}$. Then

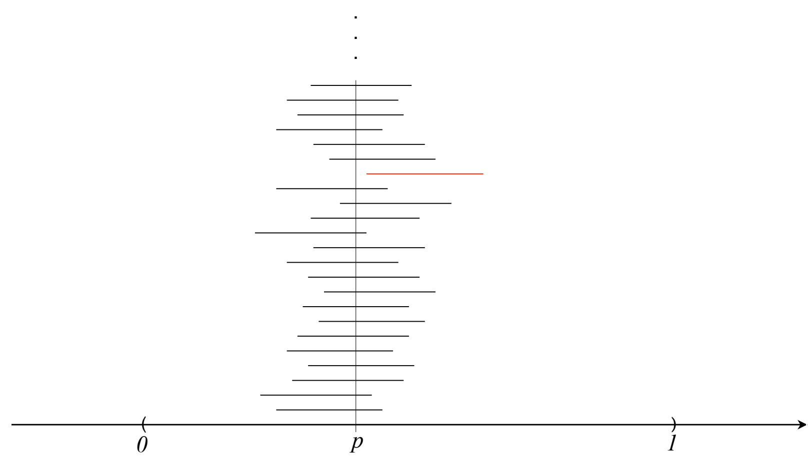
$$\hat{p} \pm E_\alpha = (\hat{p} - E_\alpha,\ \hat{p} + E_\alpha)$$

is a $(1 - \alpha) \times 100\%$ confidence interval for p, where $E_\alpha = z_{\alpha/2} \cdot \sqrt{\hat{p}\hat{q}/n}$ is the ***margin of error***, $\hat{q} = 1 - \hat{p}$, and n is the sample size.

Additional comments:

1. For such a confidence intervals to be accurate, the sample size must be big enough for the normal approximation to the binomial to be accurate. Specifically, we must have (at the very least) $n\hat{p} > 5$ and $n\hat{q} > 5$.

2. The confidence interval $\hat{p} \pm E_\alpha$ varies from sample to sample. When we say that we are 95% confident (for example) that the interval contains (the fixed) pop. parameter p, we mean that about 95 out of every 100 intervals constructed this way will contain the true population proportion (as illustrated below).



Confidence intervals for p: most of them capture p, but a small proportion (α) of them do not (in red).

Choosing sample size: to make the margin of error smaller (for the same confidence level).

The margin of error E decreases as the sample size increases because

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

If we want a specific margin of error E_0, we substitute E_0 for E in the previous equation and solve for n:

$$n = \frac{(z_{\alpha/2})^2 \cdot (\hat{p}\hat{q})}{(E_0)^2}.$$

This raises a technical problem. If we haven't collected a sample yet, we don't know what $\hat{p}$ (or $\hat{q}$) is. There are two solutions.

(i) If we have an estimate for $\hat{p}$ (say from a previous study), we can use that.

(ii) If we don't have any idea what $\hat{p}$ might be, ***or we want to be absolutely sure that n is big enough***, we use $\hat{p} = 0.5$.

Example. A simple random sample of 850 likely California voters is surveyed, and 442 of them say they will vote for Proposition 101 (banning infomercials on Sunday mornings) in the upcoming election.

(i) Find a 95% confidence interval for the proportion of all likely California voters who intend to vote for A.

- The critical value is $z_{0.025} = 1.96$.
- The sample proportion is $\hat{p} = \dfrac{442}{850} = 0.52$.
- The margin of error is $E = 1.96 \cdot \sqrt{(0.52)(0.48)/850} \approx 0.034$.

So the confidence interval is $0.52 \pm 0.034 = (0.486, 0.554)$.

(ii) The proposition requires that more than 50% of the voters approve it, and the people running the campaign for 101 want to get a clearer sense of whether it will pass. What sample size is necessary to ensure that a margin of error $E_0 \leq 0.015$?

Using the sample proportion from the first sample, we find the required sample size is

$$n = \frac{(z_{\alpha/2})^2 \cdot (\hat{p}\hat{q})}{(E_0)^2} = \frac{(1.96)^2 \cdot (0.52)(0.48)}{(0.015)^2} \approx 4262$$

(rounded up to the nearest integer).

(iii) In a *new* simple random sample of 4270 likely voters, 2212 of them indicate that they support Proposition 101. Construct a new 95% confidence interval for the proportion of all voters who support it.

- $\hat{p} = \dfrac{2212}{4270} \approx 0.518$
- $E = 1.96 \cdot \sqrt{(0.518)(0.482)/4270} = 0.0149875\ldots \approx 0.015$

The new confidence interval is $0.518 \pm 0.015 = (0.503, 0.533)$, painting a rosier picture for the supporters of Prop 101.