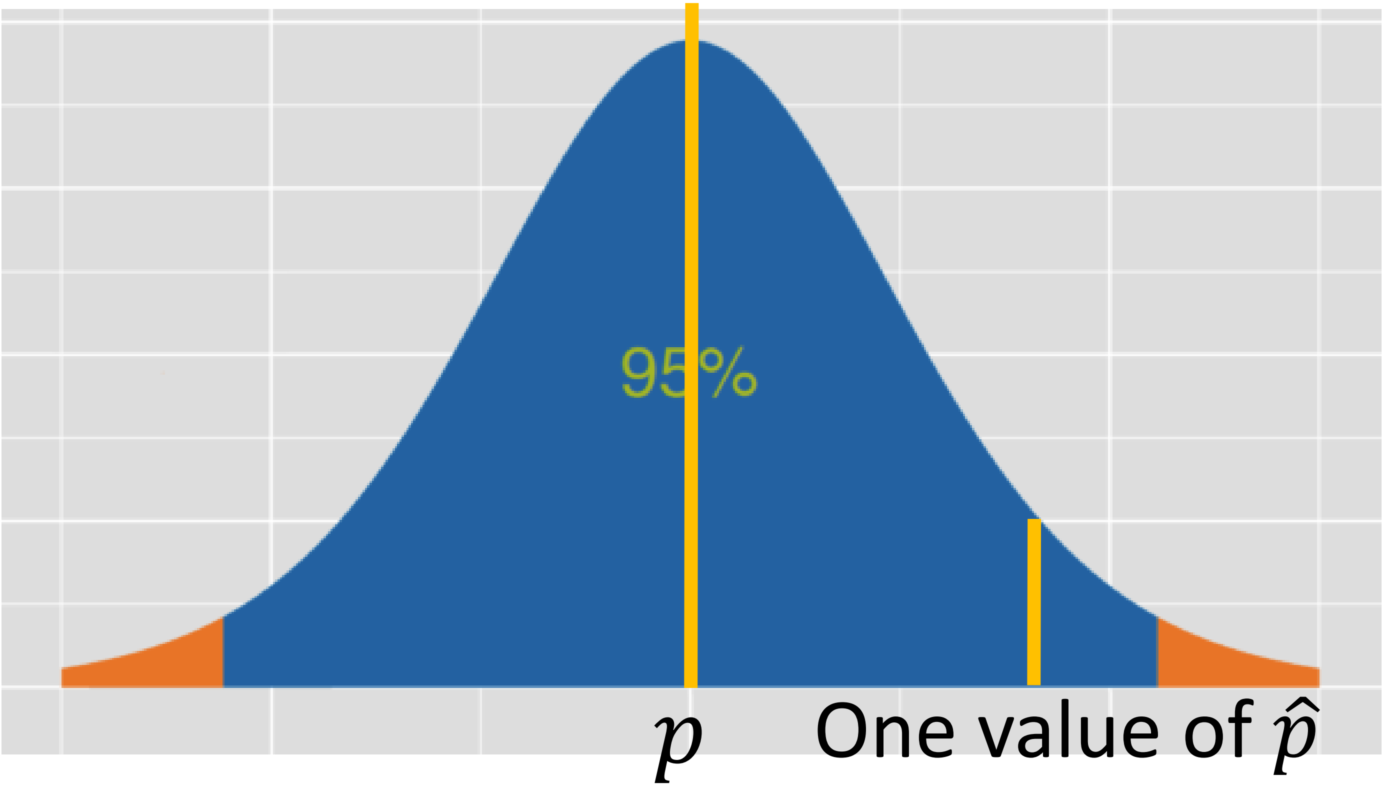
Confidence intervals (one-proportion z-intervals)

The idea

We have one value of of $\hat{p}$ and we want to find p.
We know the distribution of $\hat{p}$ is normal, with mean p and standard deviation $\sqrt{\frac{pq}{n}}$.

We also know 95% of values are within 2 SD of the mean.



So probably our $\hat{p}$ is within $2\sqrt{\frac{pq}{n}}$ of p.

But we don't know p, so we substitute $\hat{p}$ for p, to get the standard error SE= $\sqrt{\frac{\hat{p}\hat{q}}{n}}$.

Confidence intervals (one-proportion z-intervals)

The formula

$$\hat{p} \pm 1.96SE(\hat{p}) \quad \text{or} \quad \hat{p} \pm 1.96\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

is a **95% confidence interval** for the population proportion p.
More specifically, it's a **one-proportion** z-**interval**.
The distance $1.96SE$ from the middle to the end of the interval is called the **margin of error**.

Critical values

The $C\%$ confidence interval is $\hat{p} \pm z^{*}SE$, where z^{*} is the **critical value** that cuts off the middle $C\%$ of the area under the normal curve.

Level of confidence	Critical value
90%	1.65
95%	1.96
99%	2.57

Subtle points

- Although $z = 1.96$ cuts off the middle 95% of the distribution of sample proportions, the confidence interval doesn't cut off the middle 95%, since it's probably not in the middle. It's just the same width as the middle 95%.
- In fact, different confidence intervals have slightly different widths, since the formula for the standard error involves $\hat{p}$, which comes from the individual sample.
- Since 95% of the values of $\hat{p}$ are within 2 SD of the mean, 5% are not, so we expect 5% of our 95% confidence intervals to be wrong.