$z^* = 1.96$

c) Find a 95% confidence interval for p if the survey randomly selected 1500 workers and 879 said that they take their lunch to work with them.

$n = 1500$, $n\hat{p} = 879 \geq 10$

$\hat{p} = \frac{879}{1500} = 0.586$ $n(1-\hat{p}) = 621 \geq 10$

$1 - \hat{p} = \frac{621}{1500} = 0.414$

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.586 \pm 1.96 \sqrt{\frac{(0.586)(0.414)}{1500}}$$

$$= \underbrace{(0.5611, 0.6109)}_{\text{narrower than } n = 1000}$$

Example: A company estimates that 77% of adults in a certain country use coupons, with a margin of error of 2.45%. This estimate is based upon a survey of a random sample of 800 adults. What confidence level did they use?

$\hat{p} = 0.77$ $ME = 0.0245$ $n = 800$

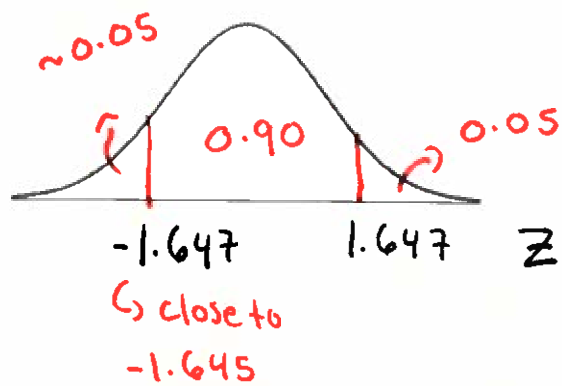
$$z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = ME$$

$$z^* \sqrt{\frac{0.77(0.23)}{800}} = 0.0245$$

$$z^* (0.014879) = 0.0245$$

$$z^* = 1.647$$

$\therefore$ approx. 90% confidence level was used.



Choosing your Sample Size

Sometimes we want to know what sample size to choose in order to have a specific margin of error with a specific confidence level.

To do this, we take the formula for the margin of error

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

and solve for n.

For example, if we want a 95% level of confidence with a margin of error of 0.03, we would solve

$$0.03 = 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

for n.

There is one problem: since we have not taken a sample, we do not have a value for $\hat{p}$. We can either use

- a value of $\hat{p}$ based on experience (such as a previous study), or
- $\hat{p} = 0.5$ → gives largest n

Thus, the formula for calculating the sample size needed to give a margin of error ME and a confidence level of $100(1-\alpha)\%$ with corresponding critical value z^* is:

ME ↓ n ↑

confidence ↑ n ↑

$$n = \hat{p}(1-\hat{p})\left(\frac{z^*}{ME}\right)^2$$

ME ↑ n ↓

where $\hat{p}$ is either guessed based on experience or $\hat{p} = 0.5$.

Note: When you use this formula, always round up (to an integer) at the end to find n!

Example: Public health officials want to know the proportion p of a population that require corrective lenses for their vision. The officials want to estimate this proportion within 5 percentage points with 98% confidence. How large a sample should they take if

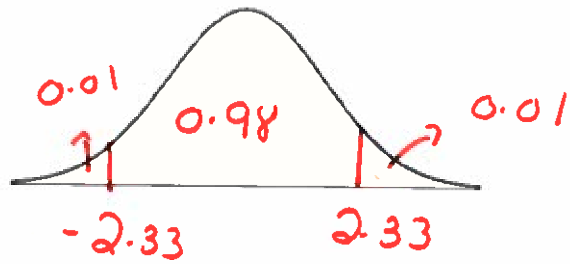
$ME = 0.05$

$1 - \alpha = 0.98$

$\alpha = 1 - 0.98 = 0.02$

$\alpha/2 = 0.01$

$z^* = 2.33$



a) there is no knowledge about what value of $\hat{p}$ to use? $\hat{p} = 0.5$

$$n = \hat{p}(1-\hat{p})\left(\frac{z^*}{ME}\right)^2$$

$$= (0.5)(0.5)\left(\frac{2.33}{0.05}\right)^2$$

$$= 542.89$$

$\therefore$ They should sample 543 people.

b) a previous study found $\hat{p} = 0.3$?

$$n = \hat{p}(1-\hat{p})\left(\frac{z^*}{ME}\right)^2$$

$$= (0.3)(0.7)\left(\frac{2.33}{0.05}\right)^2$$

$$= 456.0276$$

$\therefore$ They should sample 457 people.

Large Sample Confidence Intervals

An **estimator** is a statistic intended to estimate a parameter.

If we have a parameter θ with an estimator $\hat{\theta}$ whose sampling distribution is Normal with mean $\mu_{\hat{\theta}} = \theta$ and standard deviation $\sigma_{\hat{\theta}}$, then a confidence interval for θ has the form

$$\hat{\theta} \pm z^* \sigma_{\hat{\theta}}$$

If we do not know the value of $\sigma_{\hat{\theta}}$, but can obtain a reasonably accurate estimate from a formula and a sufficiently large sample, then a large sample confidence interval for θ is

$$\hat{\theta} \pm z^* SE(\hat{\theta})$$

where $SE(\hat{\theta})$ is the estimate for $\sigma_{\hat{\theta}}$ (standard error).

In general, a confidence interval has the form

$$\text{point estimate} \ \pm \ \text{margin of error}$$

$$= \text{point estimate} \ \pm \ \Big(\text{critical value} \times \ \text{standard error of the estimate}\Big)$$

A confidence interval is a balance between certainty and precision. A good confidence interval has two characteristics:

- it is as narrow as possible. The narrower the interval, the more precisely you have located the parameter. A very wide interval lacks precision.

  ↳ too narrow may mean lower confidence.

- it has a large confidence level. However, increasing the confident level increases the margin of error, which makes the confidence interval wider.