Using Margin of Error to calculate sample size

Learning Targets:

- Margin of Error:
  - Understand what a margin of error is.
  - Understand what factors contribute to the size of the margin of error.
  - Understand why the margin of error is important when designing an experiment.
  - Know how to calculate a sample size based on a given margin of error.

Margin of Error

- The Margin of error is the name given for the plus and minus part in the confidence interval.
- The 95% confidence interval for the population mean is
  [sample mean $\pm 1.96 \times \sigma/\sqrt{n}$]
- **The margin of error is $1.96 \times \sigma/\sqrt{n}$.**
- For example, if the 95% confidence interval for the mean rent of an apartment in Dallas is [980 $\pm 1.96 \times 88.54$] the margin of error for the mean rent **1.96×88.54**.

Understanding the Margin of Error

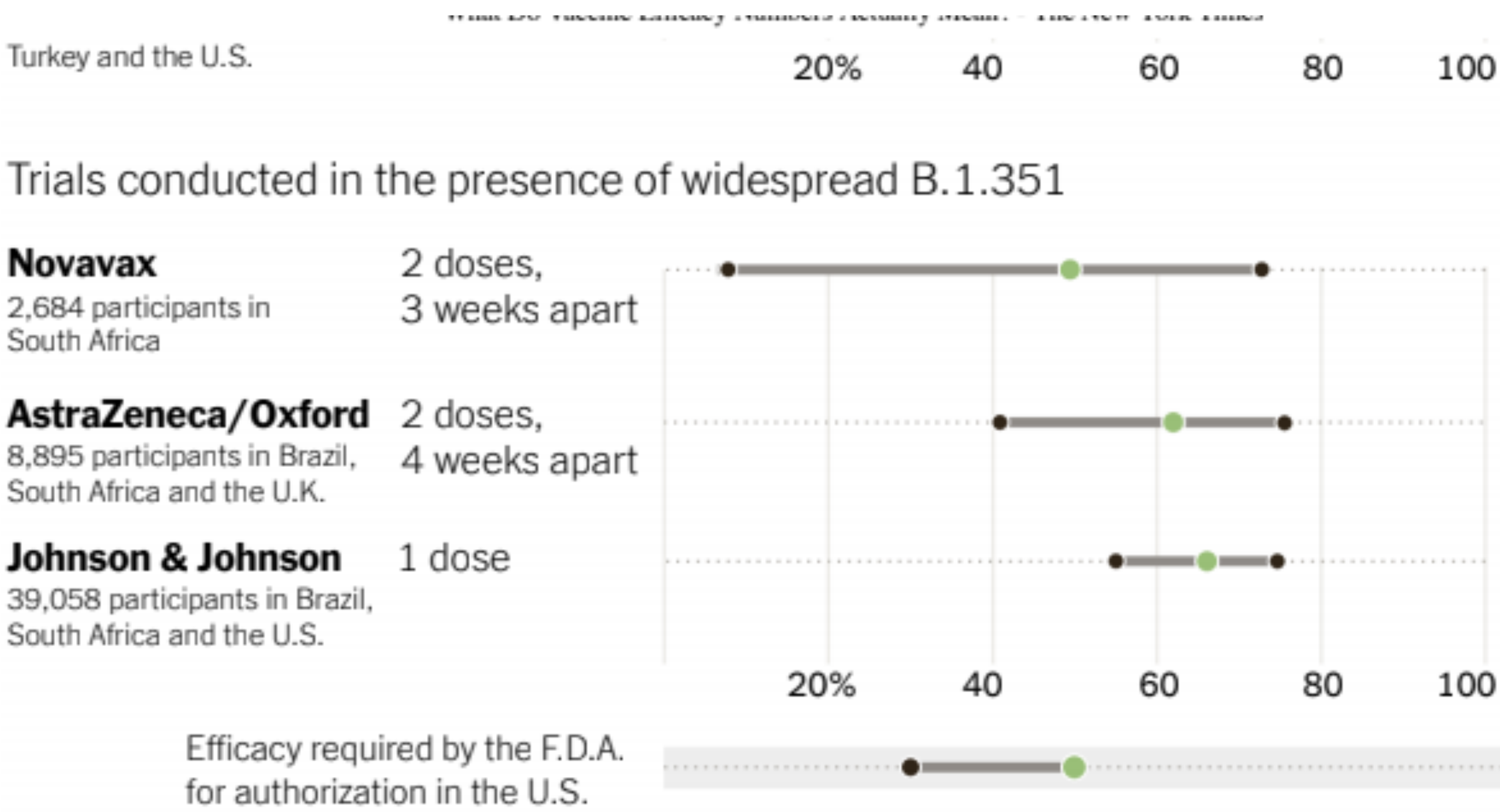
- We are given the 95% confidence interval for the mean to be [10,20].
- The **sample mean** is in the middle of the interval which is ______________.
- The interval can be written as [15-5,15+5]
- The margin of error is ________.
- The margin of error will be ___________ as the sample size grows.

More practice

- We are given the 80% confidence interval [100,140]
- The sample mean is the middle of the interval which is _____.
- The interval can be written as [120-20,120+20]
- The margin of error is _________.
- If the sample size is 5, we find σ by solving the equation

$$20 = 1.28 \times \frac{\sigma}{\sqrt{5}}$$

Confidence intervals for vaccines



These are the confidence intervals for the efficacy of different vaccines.

Discussion

- The point estimates (based on the sample) for the Johnson and Johnson is better than Novavax, but the confidence intervals different story.
- The confidence intervals explain there the population efficacy lies.
- As all the confidence intervals overlap it is impossible to distinguish between the three vaccines.
- Notice that the confidence interval for the Novavax vaccines is far wider than the confidence interval for the Johnson and Johnson.
- The Johnson and Johnson vaccines is more informative of the location of the population parameter.
- This is because the sample sizes is substantially greater.
- Indeed the confidence interval for the Johnson and Johnson appears to be about 1/3 of the size of the Novavax.
- But the sample size for the Johnson and Johnson 10 times larger!

Margin of errors in the Media

- You read in a newspaper that

  The proportion of the public that supports same-sex marriage is somewhere between 55%±15%.

- Where did this interval come from?
- It is impossible to interview the entire nation. Instead a survey was done. The proportion in the survey who supported same-sex marriage was 55%, the uncertainty (standard error) for this estimator and the 95% confidence interval for the population proportion is [55-15,55+15]% = [40%,70%].

Good length for a confidence interval

- [55-15,55+15]% = [40%,70%] is an extremely large interval, it is so wide, that it is really not that informative about the opinion of the public.
- As we will see on a later slide, the reason it is so wide is that the sample size is *too small*. A larger sample size is required to make the interval narrower and better locate the national proportion.
- A more informative Margin of Error is 3%, [55-3,55+3]% conveys more information.
- To **reduce** the Margin of Error we need to use a **larger** sample size.

How large a sample size?

- The margin of error is a measure of reliability. For a given confidence level, the smaller the margin error the more precisely we can pinpoint the true mean.
- Suppose we want the margin or error to be equal to some value, then solve for n in the equation
- MoE = $1.96 \times \sigma/\sqrt{n}$ (the Margin of Error and the standard deviation σ are given): **n = $(1.96 \times \sigma/MoE)^2$**
- **This equation is given in the cheat sheet.**

MoE: Calculation practice

- An company is reaching out to customers who bought a pedometer watch. It wants to construct a 80% confidence interval for the mean rating of its pedometer watch.
- **It is known that the standard deviation** for ratings of the pedometer watch is about 1.2.
- The company would like the margin of error to be 0.2. How many people should they contact to obtain this margin of error?
- We know that the margin of error for an 80% CI is

$$\mathrm{MoE} = 1.28 \times \frac{\sigma}{\sqrt{n}}$$

MoE: Calculation practice

$$\text{MoE} = 1.28 \times \frac{\sigma}{\sqrt{n}}$$

- Solving the above $1.28 \times 1.2/\sqrt{n} = 0.2$,
- $\sqrt{n}=(1.28 \times 1.2/0.2) = 7.68$. Solving this gives n = 59.
- This means we need to question at least 59 people, such that the 80% confidence interval has margin of error 0.2.

MoE: When the standard deviation is unknown?

- In the previous example we assumed the standard deviation was known. In general before we collect the data, we will not have much information about the standard deviation. It will be unknown.
- However, we can make an educated guess on the range of values that the standard deviation takes.
- For example, the standard deviation for human heights is probably between 2-5 inches.
- Based on this information we can can find the sample size whose Margin of Error is **at most** a certain length.

MoE: Calculation practice with unknown σ

- **Question** How large a sample size do we require such that the margin of error for a 95% confidence interval for the mean of height of a human is **maximum 0.25 inch**, given that we do know that σ lies somewhere between **2-5 inches**.
- **IMPORTANT:** The larger σ, the larger the Margin of Error:

$$MoE = 1.96\frac{\sigma}{\sqrt{n}}$$

- To be sure that the MoE is less than 0.25 we must always choose the **largest** standard deviation in the range of possible values.

- **Justification:** We use the formula is $n = (1.96 \times \sigma/0.25)^2$.
  - If we use σ=2, then the sample size we should choose is $n=(1.96 \times 2/0.25)^2 =$ **246**.
  - If we use σ=3.5, then the sample size we should choose is $n=(1.96 \times 3.5/0.25)^2 =$ **753**
  - If we use σ=5, then the sample size we should choose is $n=(1.96 \times 5/0.25)^2 =$ **1537**.
  - For standard deviations between 2 and 5, the sample size should be between 246 – 1537.
- But we do not know σ so which sample size to choose?

- If we use σ=3.5, then the sample size we should choose is $n=(1.96 \times 3.5/0.25)^2$ = **753**
- However, suppose the true standard deviation turns out to be σ=4.5.
- The MoE using a sample size 753 turns out to be

$$MoE = 1.96 \frac{4.5}{\sqrt{753}} = 0.32$$

- 0.32 is larger than the desired 0.25!
- **Moral** Use n = 1537 to be sure the MoE is less than 0.25.

- **Moral of story** Always err on the cautious side and use the largest standard deviation in the given range.

$$n = \left(\frac{1.96 \times \sigma_{MAX}}{\text{MoE}} \right)^2$$

- To be sure that the MoE is **maximum 0.25**, in the margin of error calculation always use the ___________ standard deviation.
- The largest standard deviation will always give the ___________ possible sample size.

Question Time

- A health agency wants to construct a 95% confidence interval for the mean height of a 5 year old. The standard deviation of a 5 year old is known to be somewhere between 1 to 3 inches. What is the minimum sample size required to ensure the margin of error is less than 0.5 inches?

(A) 16 (B) 44 (C) 62 (D) 139 (E) 156

MoE: Calculation practice

- **Question:** A 95% confidence interval for the mean length of parrots beaks is [4,10] = [7-**3**,7+**3**] inches.
- The old MoE is **3** inches
- The sample size is 20. By what **factor** (by factor one can means double, triple etc) should one increase the sample size such that the margin of error is reduced to 1?

Answer: The margin of error (using formula) is

$$3 = 1.96 \times \frac{\sigma}{\sqrt{20}}$$

- **Option 1** You can solve for σ and then calculate the sample size using the formula.

- **Option 2:** We want to decrease the MoE, such that MoE = 1. New margin of error is **one third** of the old one:

$$3 = 1.96 \times \frac{\sigma}{\sqrt{20}}$$

- The new MoE decreases the old MoE to a third of its original size. The maths is below

$$\text{new MoE} = 1 = \frac{3}{3} = \frac{1}{3} \times \text{old MoE} = \frac{1}{3} \times 1.96 \times \frac{\sigma}{\sqrt{20}}$$

- Maths reminder

$$\frac{1}{3\sqrt{20}} = \frac{1}{\sqrt{9}\sqrt{20}} = \frac{1}{\sqrt{9 \times 20}} = \frac{1}{\sqrt{180}}$$

- Using the maths above:

$$\text{new MoE} = 1 = 1.96 \times \frac{\sigma}{\sqrt{180}}$$

- We need to increase the sample size from 20 to 180 (interview 180 people!), this is a 9 fold increase to reduce the margin of error from 3 to 1.
- This is a **substantial increase** in the sample size.

MoE: Calculation practice (tricky)

- **Question:**
- A 95% confidence interval for the mean length of parrots beaks is [4,10] = [7-**3**,7+**3**] inches. It is based on a sample of size n (but n is not given). By what factor should the sample size increase such that the margin of error is reduced 1?
- **Answer:** This looks like an impossible question because we don't have any obvious information on the standard deviation or sample size. But as in the previous question can breakdown the problem into steps:

- The margin of error is **3** = $1.96 \times \sigma/\sqrt{n}$.

- We cannot solve for σ because n is not given. So we have to use the same trick given on the previous slide. To reduce by a margin of error by a third:

$$\text{new MoE} = 1 = \frac{3}{3} = \frac{\text{old MoE}}{3} = \frac{1}{3} \times 1.96 \times \frac{\sigma}{\sqrt{n}}$$

- Again we use some maths (to take take 3 into the square root):

$$\frac{1}{3\sqrt{n}} = \frac{1}{\sqrt{9}\sqrt{n}} = \frac{1}{\sqrt{9n}}$$

- Using the maths above:

$$\text{new MoE} = 1 = \frac{3}{3} = \frac{1}{3} \times 1.96 \times \frac{\sigma}{\sqrt{9n}}$$

- We need to increase the sample size from n to 9n this is a 9 fold increase to reduce the margin of error from 3 to 1.
- This is **exactly** what we did previously when n = 20.

Calculation 2 (more practice)

- **Example** If a sample size of 20 give the 95% confidence interval for the mean to be [2,10], how large a sample size is required to reduce the margin of error to 1/2 (0.5)?
- **Solution** Since the confidence interval is [2,10] =[6-4,6+4].
- The Margin of Error is 4.

$$\text{MoE} = 4 = 1.96 \times \frac{\sigma}{\sqrt{20}}$$

$$\text{MoE} = 4 = 1.96 \times \frac{\sigma}{\sqrt{20}}$$

$$\text{new MoE} = 0.5 = \frac{4}{8} = \frac{4}{8} \times 1.96 \times \frac{\sigma}{\sqrt{20}} = 1.96 \times \frac{\sigma}{\sqrt{8^2 \times 20}}$$

This means increasing sample size from 20 to 1280.

We see that to decrease the margin of error from 4 to ½ (by an eighth) we need to increase the sample size by factor 64!