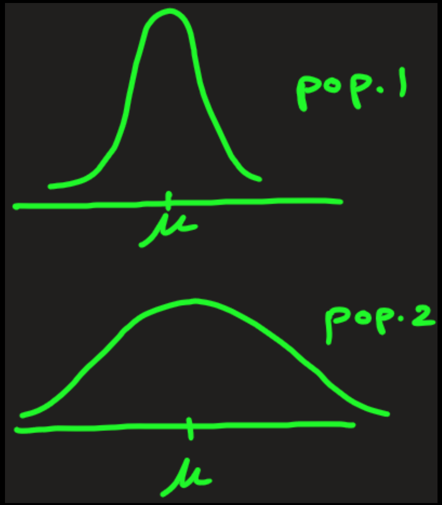
Sample Size

Problem: I have 2 sets of population with 15 data points in each:
Population1 = [3, 4, 2, 5, 2, 6, 3, 4, 3, 3, 2, 1, 3, 2, 2]
Population1 has standard deviation of 1.33.

Population2 = [-4.35, -4.35, -4.35, -4.35, -4.35, 10.35, 10.35, 10.35, 10.35, 10.35, 3.00, 3.00, 3.00, 3.00, 3.00]
Population2 has standard deviation of 6.



Both population have same mean. We have to determine that mean value.

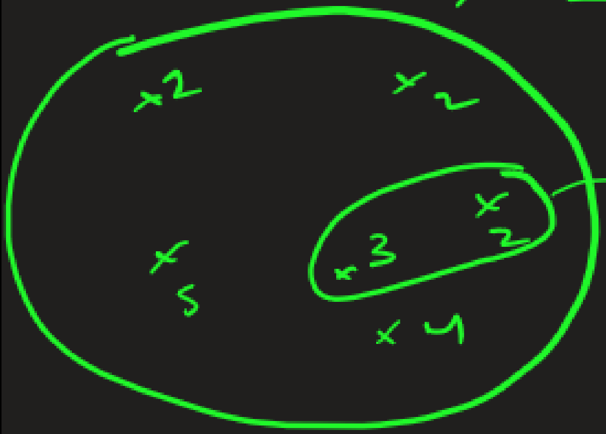
If I am allowed to take only 10 sample points to determine mean value, which population should I use to get better estimate of the mean?

Factors that Effect the Accuracy of Estimate

Small sample size gives more uncertainty i.e. less confident + less accurate

population $\mu = \frac{2+2+2+3+4+5}{6}$

$= 3$



sample mean $\bar{x} = \frac{3+2}{2}$

$= \underline{2.50}$

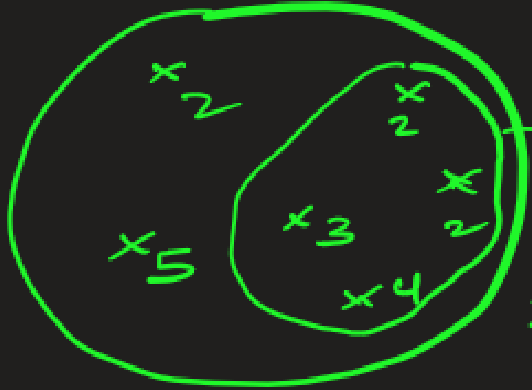
Small sample

($\bar{x}$ is less closer to μ)

Large sample size gives less uncertainty, i.e. more confident + more accurate

population μ

$= 3$



sample mean

$\bar{x} = \frac{2+2+3+4}{4}$

$= \underline{2.75}$

Large sample

($\bar{x}$ is more closer to μ)

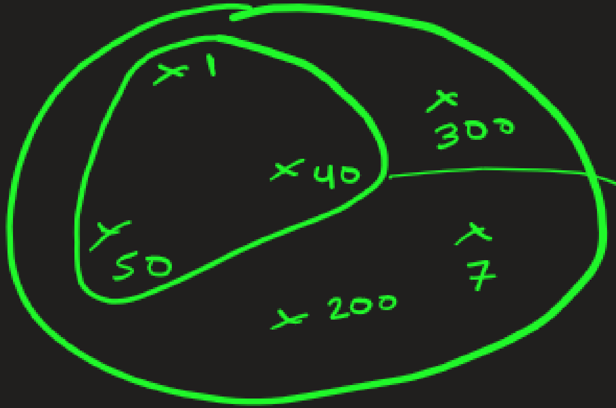
Factors that Effect the Accuracy of Estimate

High variability (i.e. high standard deviation σ) in the population **gives more uncertainty i.e. less confident + less accurate**

Low variability (i.e. low standard deviation σ) in the population **gives less uncertainty, i.e. more confident + more accurate**

pop. mean $\mu = \frac{\sum \text{values}}{\text{count}}$

$= 96.6$



sample mean

$\bar{x} = \frac{1+40+50}{3}$

$= 30.3$

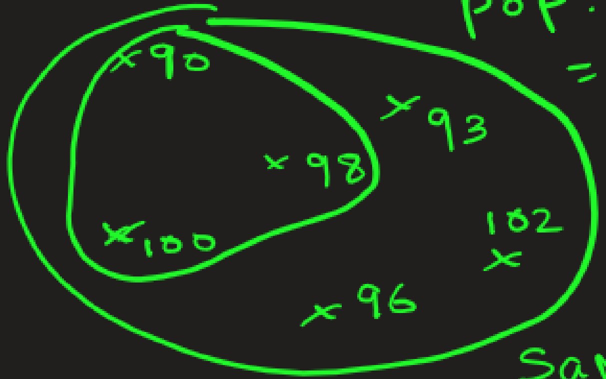
High variability

($\bar{x}$ is less closer to μ)

pop. mean

$= \frac{\sum \text{value}}{\text{count}}$

$= 96.6$



Sample mean

$\bar{x} = \frac{90+98+100}{3}$

$= 96$

Low Variability

($\bar{x}$ is more closer to μ)

How large a sample is necessary to make an accurate estimate ?

We can derive the formula for sample size n from margin of error E, population standard deviation σ and critical z-value as follows:

$$E = Z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

$$E\sqrt{n} = Z_{\alpha/2} \cdot \sigma$$

$$n = \left(\frac{Z_{\alpha/2} \cdot \sigma}{E}\right)^2$$

where,

E = desired **margin of error** (the maximum acceptable difference between sample mean and true mean)

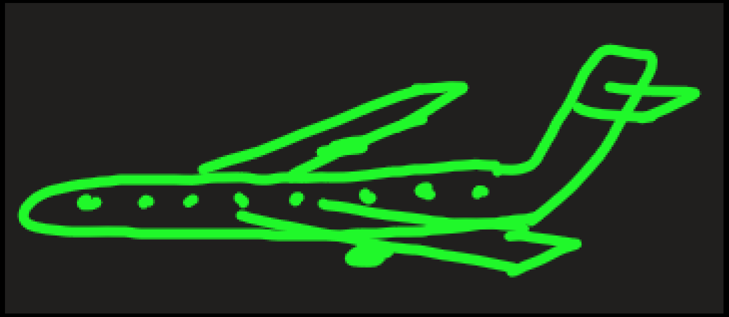
$z_{\alpha/2}$ = z-score corresponding to the confidence level (e.g. 1.645 for 90%, 1.96 for 95%, 2.58 for 99%)

The answer to above question depends on three factors:

- Margin of error (E): How much error are you willing to accept ? If you want less E, choose bigger sample size.
- Population standard deviation (σ): How much variation in the data exist ? If σ is high (high variability), then more samples are needed
- Degree of confidence (critical z-value) : What percentage of confidence level you want ? If you want high confidence, $z_{\alpha/2}$, you need more samples.

How large a sample is necessary to make an accurate estimate ?

Problem: An aerospace engineer is estimating the average fuel consumption rate of an aircraft engine during cruise conditions. For flight planning and performance certification, the engineer needs 99% confidence that the estimate is within 2 gallons per hour of the true mean consumption rate. Flight test data shows a standard deviation of 4.3 gallons per hour. How many minimum samples he needs?



Ans:

Given: $z_{\alpha/2}$ = 2.58 (for 99% confidence level) ,

σ = 4.3,

E = 2

$$n = \left(\frac{Z_{\alpha/2} \cdot \sigma}{E}\right)^2$$

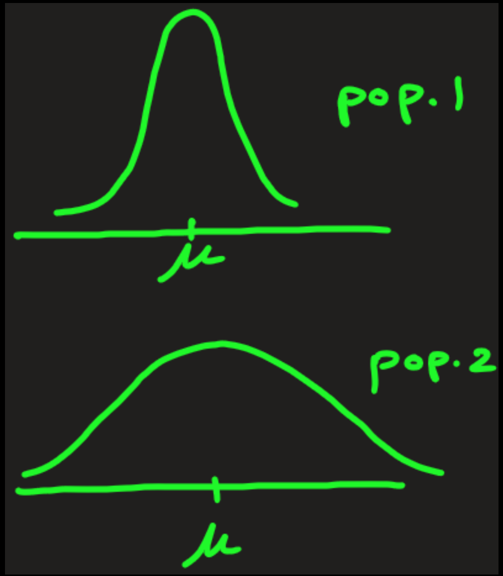
$$n = \left(\frac{2.58 \times 4.3}{2}\right)^2 = 32$$

To be 99% confident that the estimate is within 2 units of the true mean, the engineer needs a sample of **at least 32 measurements.**

Sample Size

Problem: I have 2 sets of population with 15 data points in each:
Population1 = [3, 4, 2, 5, 2, 6, 3, 4, 3, 3, 2, 1, 3, 2, 2]
Population1 has standard deviation of 1.33.

Population2 = [-4.35, -4.35, -4.35, -4.35, -4.35, 10.35, 10.35, 10.35, 10.35, 10.35, 3.00, 3.00, 3.00, 3.00, 3.00]
Population2 has standard deviation of 6.



Both population have same mean, and we have to determine that mean value.

If I am allowed to take only 10 sample points to determine mean value,
which population should I use to get better estimate of the mean?

Ans: From the relation,

$$E = Z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

we see that for E, margin of error, to be minimum we should select population with low standard deviation σ.
Select population1.

EXTRA

Ffnbksdfd
Fdslj;f

Fl;ds

EXTRA

Ffnbksdfd
Fdslj;f

Fl;ds