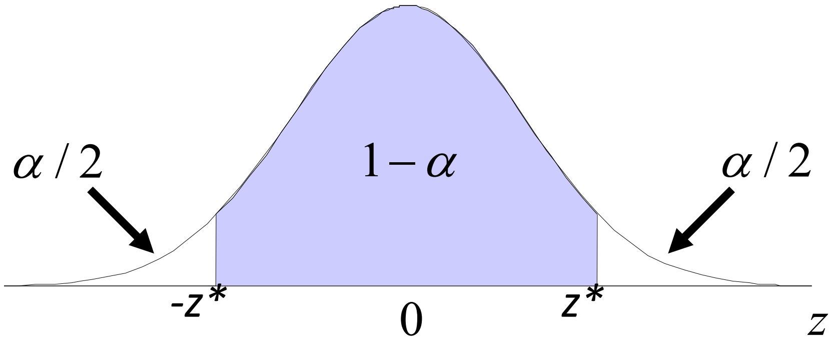
1-α Confidence Interval for μ

$$\bar{X} \pm \underbrace{z^* \frac{\sigma}{\sqrt{n}}}_{\text{Margin of error}}$$

- $\bar{X}$ is the *point estimate* of the population mean

- z^*: *critical value*, the z-value that cuts off a tail area $\alpha/2$
- $\bar{X}$ is assumed to be normally distributed



CASE : Water Quality

- As part of solving a legal dispute between a local environmental interest group and a company (accused of polluting the water in a nearby river), 30 water samples measuring the level of dissolved oxygen were obtained.
- The interest group believes that the level is below 4.5 [mg/l] whereas the company claims that it is above that. (The std.dev. is known to be approximately 1.25 [mg/l])
- Does the data support any of the two parties in their claim?

Water Quality Data

Statistical Inference: Is μ above or below 4.5 [mg/l]?

$$\bar{X} \pm z^* \frac{\sigma}{\sqrt{n}}$$

3.50	3.75	3.00	3.37	3.72	3.96
5.20	5.51	3.46	6.09	1.83	3.75
5.51	3.17	4.68	5.51	5.75	1.60
4.94	1.15	4.46	2.46	2.96	4.20
6.28	3.59	3.21	6.03	4.02	6.95

Level of Dissolved Oxygen [mg/l]

$\bar{x} = 4.12$ [mg/l]

$\sigma = 1.25$ [mg/l]

Reducing the Margin of Error

Three items that affect the Margin of Error:

$$m = z^* \frac{\sigma}{\sqrt{n}}$$

- Level of Confidence *C=1-α*
- Standard deviation σ
- Sample Size *n*

For a given confidence level and standard deviation, the sample size necessary to achieve a margin or error *m is:*

$$n = \left(\frac{z^* \sigma}{m}\right)^2$$

Round up to the nearest integer

Water Quality Data

How many observations would be needed to obtain a margin of error not to exceed 0.20?

$\sigma = 1.25$ [mg/l]
95% Confidence level

CASE: Improved Gas Mileage?

- An automobile manufacturer claims that their new model is more fuel efficient than their previous model.
- The gas mileage of the previous model was found by a consumers magazine to be approximately normal distributed with mean 24 [mpg].
- A sample of 30 of the new cars were tested. Is there evidence that the gas mileage is better for the new model?

Gas Mileage Data

25.3	25.1	29.6	24.6	26.0	26.0
26.3	23.6	26.0	25.4	26.1	23.8
25.1	24.1	25.8	26.4	23.4	24.8
22.6	26.6	25.1	26.6	28.0	23.3
23.8	25.4	26.2	25.1	25.3	21.5

Measured Gas Mileage in [mpg]

$\bar{x} = 25.23$ [mpg]
$s = 1.59$ [mpg]

Large Sample Inference Procedures when σ is unknown.

- For large samples (say $n \geq 30$), we may approximate the confidence interval the following approximation.

$$\sigma \approx s$$

Approximate 1-α Confidence Interval:

$$\bar{X} \pm z^* \frac{s}{\sqrt{n}}$$

Gas Mileage Data

Statistical Inference: Is μ above 24 [mpg]?

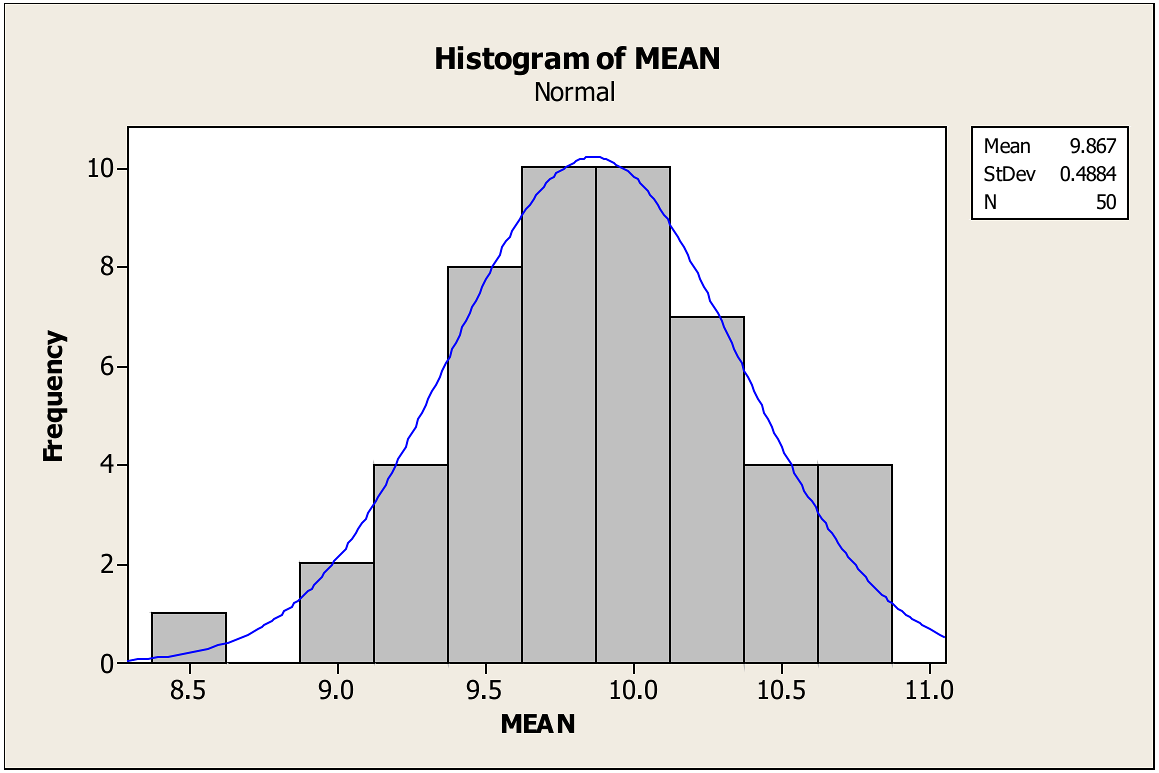
95% Confidence interval for μ:

25.3	25.1	29.6	24.6	26.0	26.0
26.3	23.6	26.0	25.4	26.1	23.8
25.1	24.1	25.8	26.4	23.4	24.8
22.6	26.6	25.1	26.6	28.0	23.3
23.8	25.4	26.2	25.1	25.3	21.5

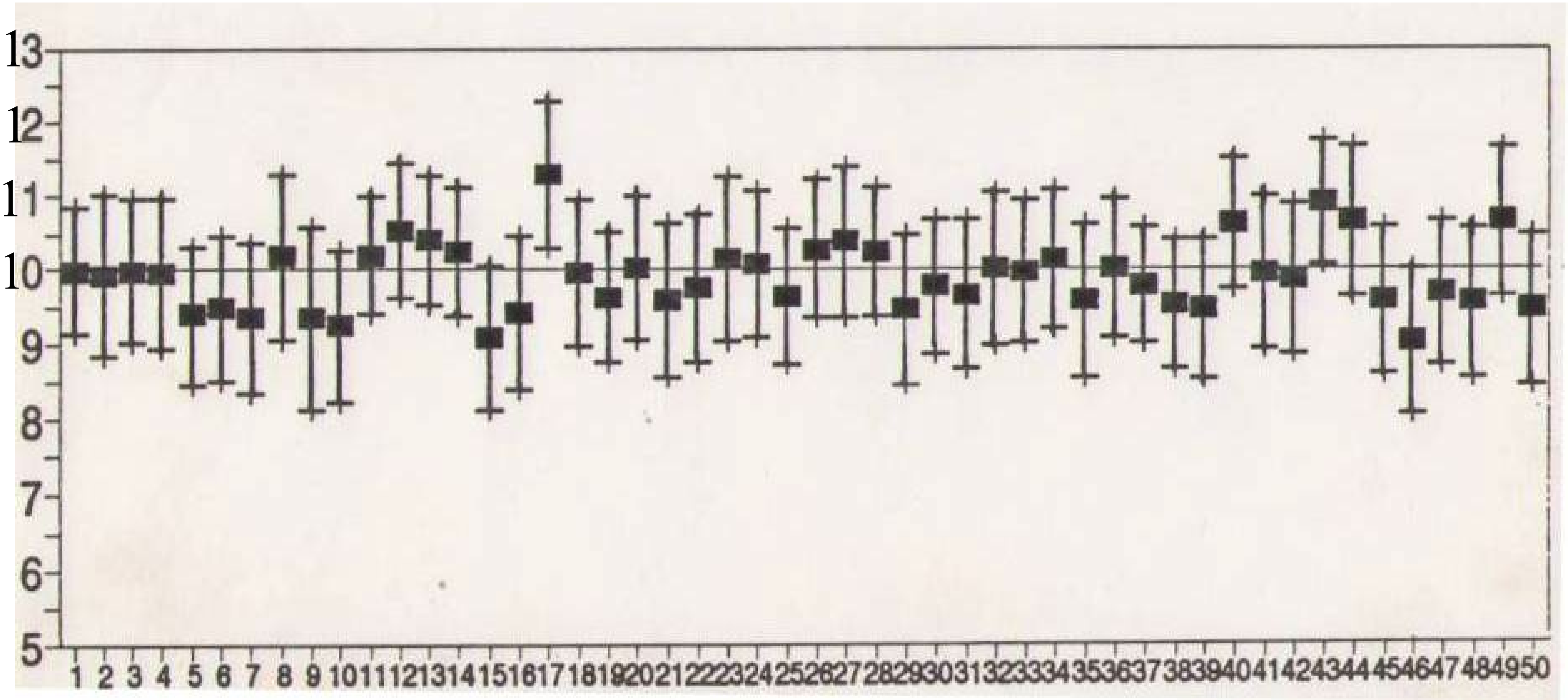
Measured Gas Mileage in [mpg]

$\bar{x} = 25.23$ [mpg]
$s = 1.59$ [mpg]

Confidence Intervals
Normal Samples with $\mu=10$ $\sigma=5$ $n=100$



Confidence Intervals
Normal Samples with μ=10 σ=5



DOs and DON'Ts for
Confidence Intervals

Each case below involves a statistical inference.

Determine whether the interpretation is correct. If not explain what is wrong.

- **Given:** *A 95% confidence interval for the mean student score on a particular college placement test, is found based on a random sample of size n = 100 to be [56,63].*
- **Interpretation:** *This means that we can expect 95% of the students that take this test to score between 56 and 63 points.*

WRONG! The confidence interval says something about where the mean of a set of observations is, not where individual observations will fall.

DOs and DON'Ts for
Confidence Intervals

Each case below involves a statistical inference.
Determine whether the interpretation is correct. If not explain what is wrong.

- **Given:** *A 90% confidence interval for the mean of a normal population is calculated based on a random sample of n = 10 observations.*
- **Interpretation:** *This means that we are 90% confident that this confidence interval will include the true (unknown) population mean.*

CORRECT! NOTE : The fact that the sample is small does not impact the truthfulness of the statement, as long as the correct method is used when calculating the confidence interval.

DOs and DON'Ts for

Confidence Intervals

Each case below involves a statistical inference.

Determine whether the interpretation is correct. If not explain what is wrong.

- **Given:** *99% confidence interval for an unknown population mean is found as [0.34, 0.67]*
- **Interpretation:** *This means that μ is between 0.34 and 0.67 99% of the time.*

WRONG! The mean is either between 0.34 and 0.67 or it is not. There is no probability associated with that fact. Note that the limits are the random variables. They vary from sample to sample. The mean (μ) is fixed. It will not change from sample to sample. (See b) for a correct interpretation)

DOs and DON'Ts for
Confidence Intervals

Each case below involves a statistical inference.

Determine whether the interpretation is correct. If not explain what is wrong.

- **Given:** *A 95% confidence interval for an unknown population mean is found as [-0.03, 0.14].*
- **Interpretation:** *This means that $\mu > 0$ since almost all the values are above 0.*

  WRONG! The confidence interval includes both positive and negative values. We can not conclude that μ is positive.