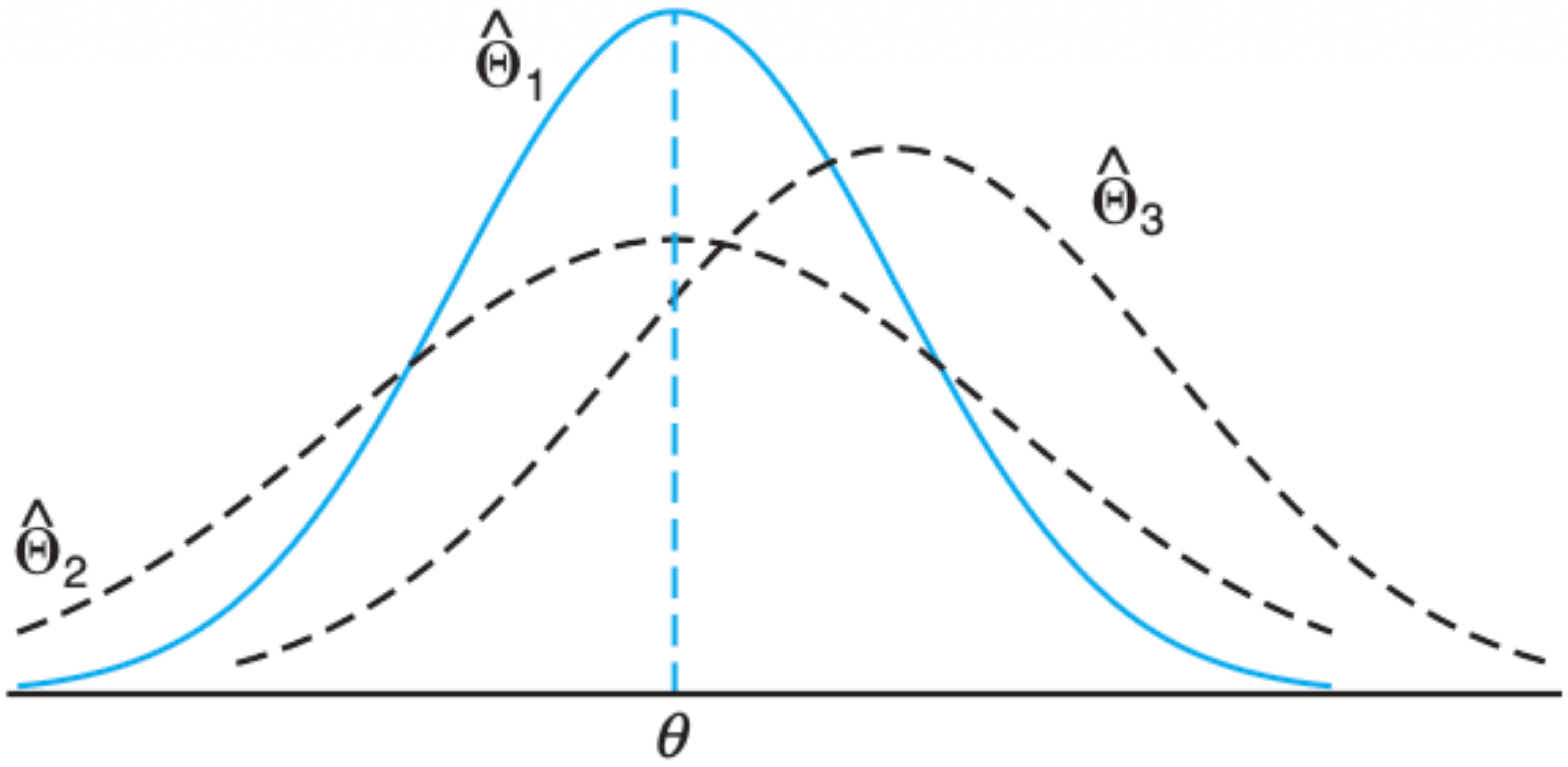
Quiz

We have 3 estimators $\hat{\Theta}_1$, $\hat{\Theta}_2$, $\hat{\Theta}_3$ for a parameter of interest θ.

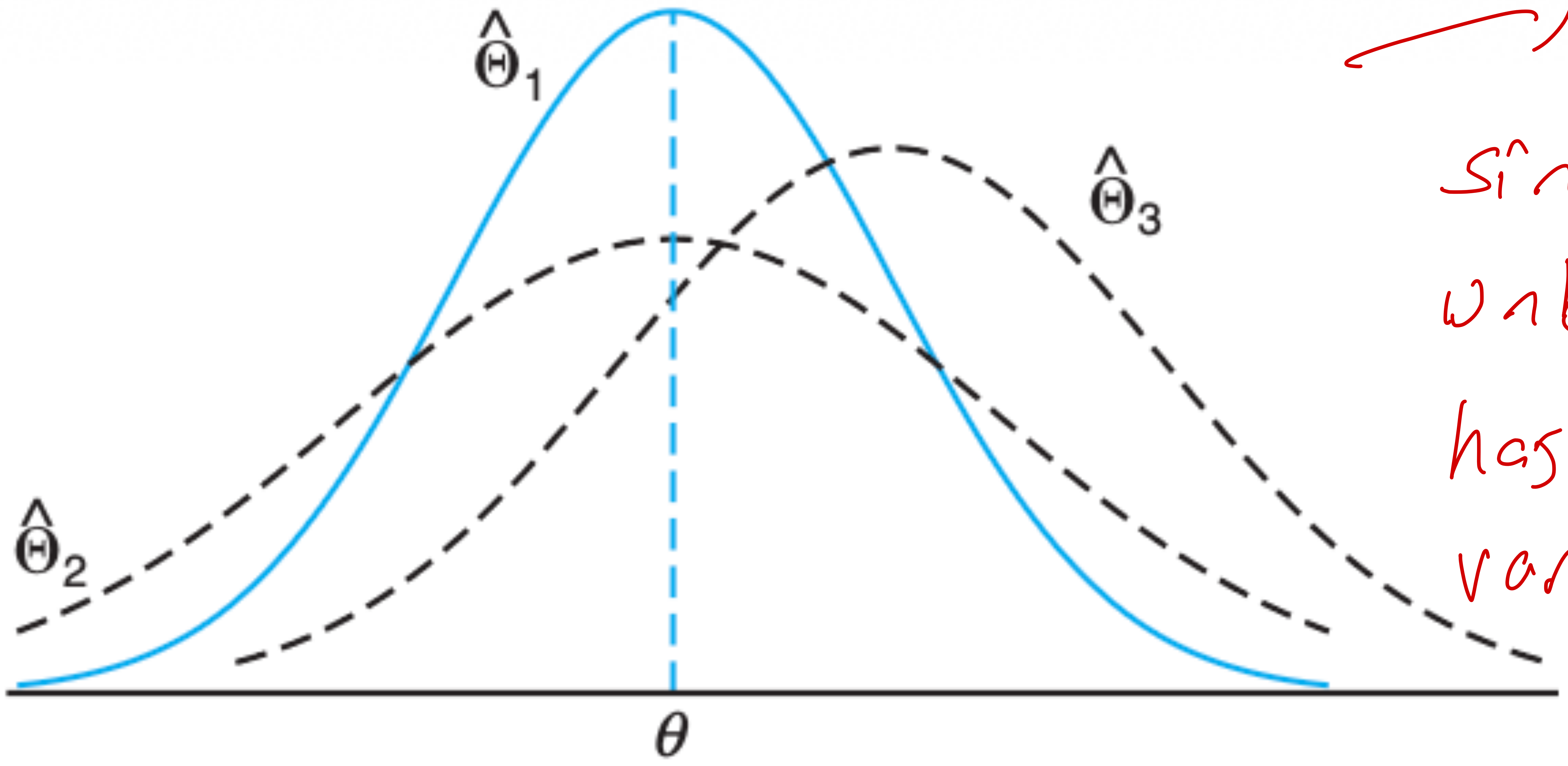


The figure shows where θ lies and the distributions of the estimators.

Which one would you use to estimate θ? Why?

Quiz

We have 3 estimators $\hat{\Theta}_1$, $\hat{\Theta}_2$, $\hat{\Theta}_3$ for a parameter of interest θ.



→ $\hat{\Theta}_1$ since it is unbiased and has smallest variance.

The figure shows where θ lies and the distributions of the estimators.

Which one would you use to estimate θ? Why?

Announcements

HW8 is due on next Monday, April 6.

Midterm 2 is on Monday, April 13.

A sample exam will be shared through D2L sometime next week.

No office hour today due to an overlap with another departmental activity.

Send email if you need to meet before Monday office hours.

HW6:

Exercise. Write Python code using scipy.stats and matplotlib to reproduce Figure 6.22.

Misunderstanding: Exercise 6.22 rather than Figure 6.22 in textbook. Partial credit.

Outline

- We did an introduction to confidence intervals for the mean:

With $100(1-\alpha)$ confidence μ is in $\bar{x} \pm z_{\alpha/2} s.e.(\bar{x})$

Today

- Error in estimating μ
- One-sided confidence intervals for the mean
- Confidence intervals for the mean with σ^2 unknown
  - t-distribution
- Prediction and tolerance intervals (if time permits)
- Confidence intervals for difference of two means (if time permits

Outline

➢ We did an introduction to confidence intervals for the mean:

With $100(1-\alpha)$ confidence μ is in $\bar{x} \pm z_{\alpha/2} s.e.(\bar{x})$

Today

➢ Error in estimating μ

➢ One-sided confidence intervals for the mean

➢ Confidence intervals for the mean with σ^2 unknown

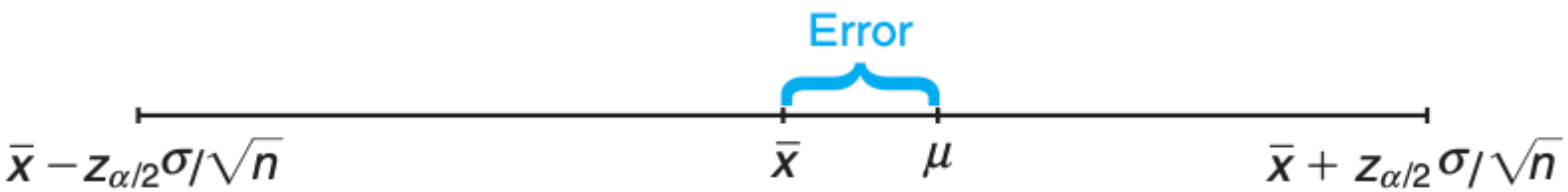
➢ t-distribution

➢ Prediction and tolerance intervals (if time permits)

➢ Confidence intervals for difference of two means (if time permits

Confidence Interval for the Mean

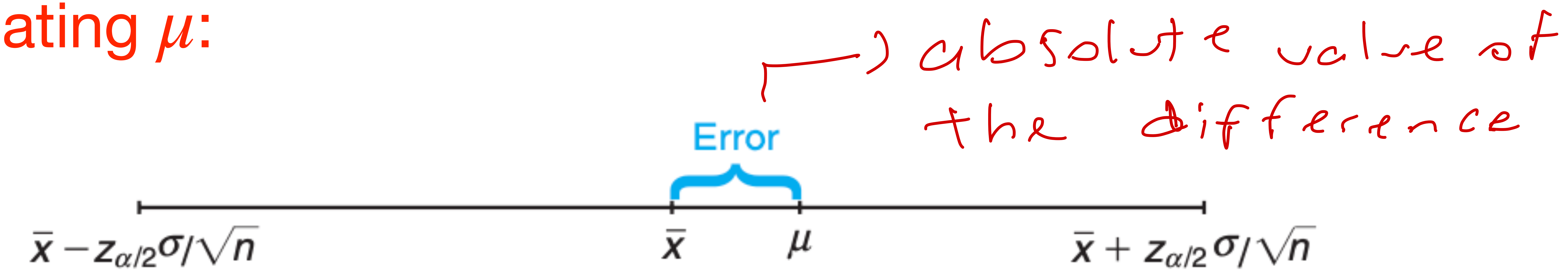
Error in Estimating μ:



Theorem 9.1: If $\bar{x}$ is used as an estimate of μ, we can be $100(1-\alpha)\%$ confident that the error will not exceed $z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$.

Confidence Interval for the Mean

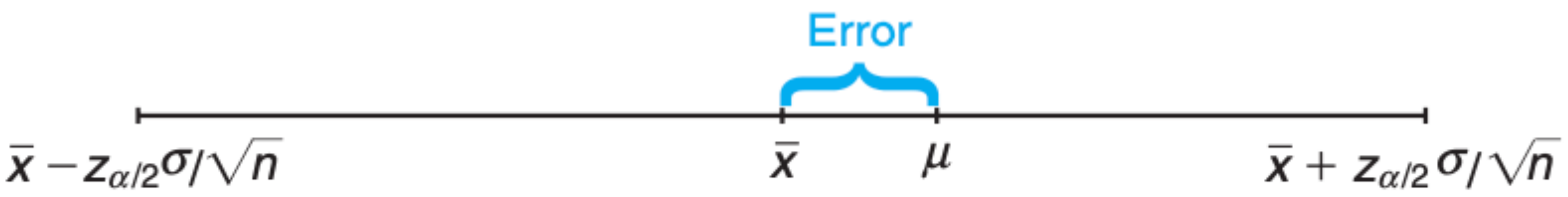
Error in Estimating μ:



Theorem 9.1: If $\bar{x}$ is used as an estimate of μ, we can be $100(1-\alpha)\%$ confident that the error will not exceed $z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$.

Confidence Interval for the Mean

Error in Estimating μ:



Theorem 9.1: If $\bar{x}$ is used as an estimate of μ, we can be $100(1-\alpha)\%$ confident that the error will not exceed $z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$.

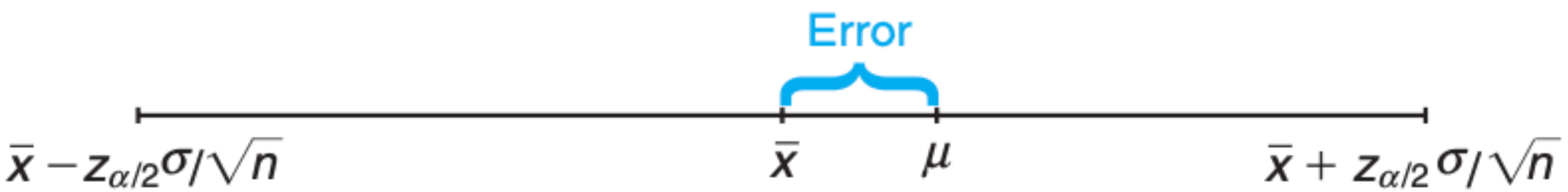
If we need to determine necessary sample size given error e:

Theorem 9.2: If $\bar{x}$ is used as an estimate of μ, we can be $100(1-\alpha)\%$ confident that the error will not exceed a specified amount e when the sample size is

$$n = \left(\frac{z_{\alpha/2}\sigma}{e}\right)^2.$$

Confidence Interval for the Mean

Error in Estimating μ:



Theorem 9.1: If $\bar{x}$ is used as an estimate of μ, we can be $100(1-\alpha)\%$ confident that the error will not exceed $z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$. = e and extract n

If we need to determine necessary sample size given error e:

Theorem 9.2: If $\bar{x}$ is used as an estimate of μ, we can be $100(1-\alpha)\%$ confident that the error will not exceed a specified amount e when the sample size is

$$n = \left(\frac{z_{\alpha/2}\sigma}{e}\right)^2.$$

Confidence Interval for the Mean

Example 9.3: How large a sample is required if we want to be 95% confident that our estimate of μ in Example 9.2 is off by less than 0.05?

Use $z_{0.025} = 1.96$.

Note $\sigma = 0.3$ in Example 9.2.

Confidence Interval for the Mean

Example 9.3: How large a sample is required if we want to be 95% confident that our estimate of μ in Example 9.2 is off by less than 0.05?

Use $z_{0.025} = 1.96$.

Note $\sigma = 0.3$ in Example 9.2.

Solution:

$$n = \left[\frac{(1.96)(0.3)}{0.05}\right]^2 = 138.3.$$

$\Rightarrow$ a sample size of 139 suffices.

One-Sided Confidence Intervals

One-Sided Confidence Bounds on μ, σ^2 Known

If $\bar{X}$ is the mean of a random sample of size n from a population with variance σ^2, the one-sided $100(1-\alpha)\%$ confidence bounds for μ are given by

$$\text{upper one-sided bound:} \quad \bar{x} + z_\alpha \sigma/\sqrt{n};$$

$$\text{lower one-sided bound:} \quad \bar{x} - z_\alpha \sigma/\sqrt{n}.$$

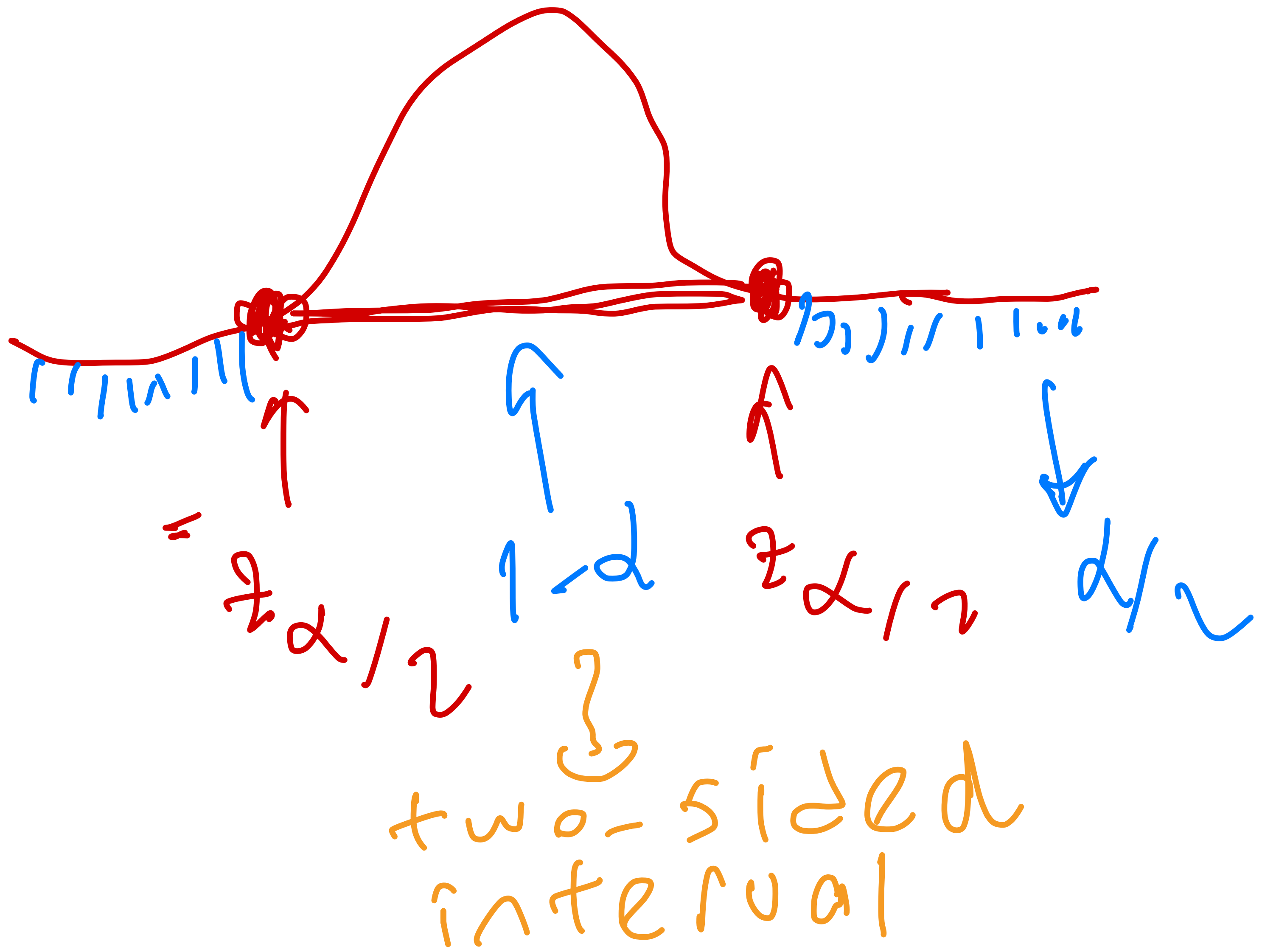
One-Sided Confidence Intervals

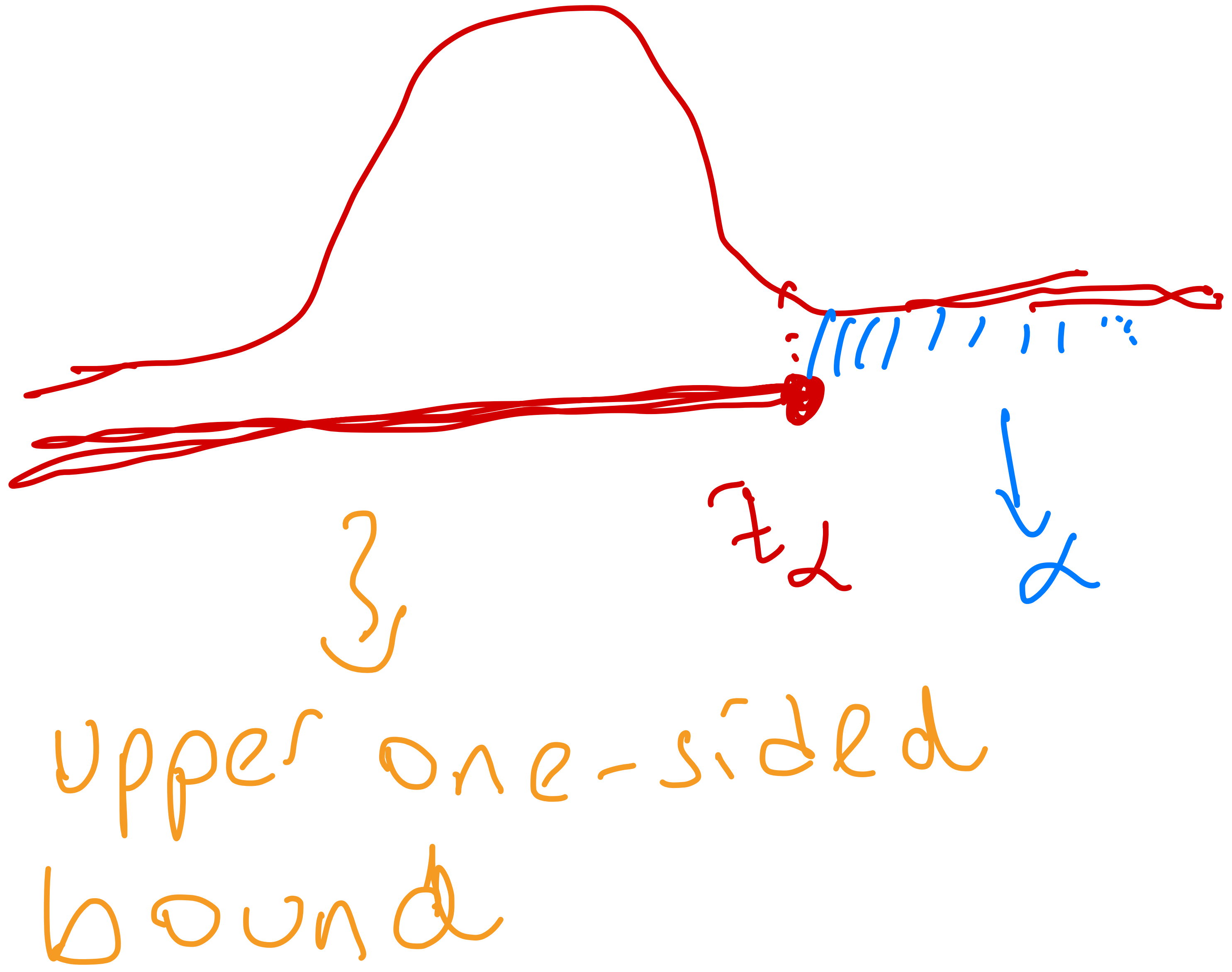
One-Sided Confidence Bounds on μ, σ^2 Known

If $\bar{X}$ is the mean of a random sample of size n from a population with variance σ^2, the one-sided $100(1-\alpha)\%$ confidence bounds for μ are given by

$$\text{upper one-sided bound:} \quad \bar{x} + z_\alpha \sigma/\sqrt{n};$$

$$\text{lower one-sided bound:} \quad \bar{x} - z_\alpha \sigma/\sqrt{n}.$$





One-Sided Confidence Intervals

Example 9.4: In a psychological testing experiment, 25 subjects are selected randomly and their reaction time, in seconds, to a particular stimulus is measured. Past experience suggests that the variance in reaction times to these types of stimuli is 4 sec^2 and that the distribution of reaction times is approximately normal. The average time for the subjects is 6.2 seconds. Give an upper 95% bound for the mean reaction time.

Use norm.ppf(0.95)=1.645 or norm.ppf(0.975)=1.96

Choose which one to use.

One-Sided Confidence Intervals

Example 9.4: In a psychological testing experiment, 25 subjects are selected randomly and their reaction time, in seconds, to a particular stimulus is measured. Past experience suggests that the variance in reaction times to these types of stimuli is 4 sec^2 and that the distribution of reaction times is approximately normal. The average time for the subjects is 6.2 seconds. Give an upper 95% bound for the mean reaction time.

Use norm.ppf(0.95)=1.645 or norm.ppf(0.975)=1.96

Choose which one to use.

n, σ^2, $\bar{x}$, $100(1-\alpha)$

$$\bar{x} + z_\alpha \cdot \frac{\sigma}{\sqrt{n}}$$

One-Sided Confidence Intervals

Example 9.4: In a psychological testing experiment, 25 subjects are selected randomly and their reaction time, in seconds, to a particular stimulus is measured. Past experience suggests that the variance in reaction times to these types of stimuli is 4 sec^2 and that the distribution of reaction times is approximately normal. The average time for the subjects is 6.2 seconds. Give an upper 95% bound for the mean reaction time.

Use norm.ppf(0.95)=1.645 or norm.ppf(0.975)=1.96

Choose which one to use.

Solution:

95 % upper one-sided bound:

$$\bar{x} + z_\alpha \sigma/\sqrt{n} \quad \Rightarrow \quad 6.2 + z_{0.05} 2/\sqrt{25} = 6.2 + 1.645 \times 2/5 = 6.2 + 0.658 = 6.858$$

We are 95 % confident that mean reaction time is less than 6.858.

Confidence Interval for the Mean When σ^2 Unknown

Corollary 8.1: Let $X_1, X_2, \ldots, X_n$ be independent random variables that are all normal with mean μ and standard deviation σ. Let

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad \text{and} \qquad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

Then the random variable $T = \frac{\bar{X}-\mu}{S/\sqrt{n}}$ has a t-distribution with $v = n - 1$ degrees of freedom.

From Section 8.6 of the textbook.

What is the t-distribution?

Confidence Interval for the Mean When σ^2 Unknown

Corollary 8.1: Let $X_1, X_2, \ldots, X_n$ be independent random variables that are all normal with mean μ and standard deviation σ. Let

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad \text{and} \qquad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

Then the random variable $T = \frac{\bar{X}-\mu}{S/\sqrt{n}}$ has a t-distribution with $v = n - 1$ degrees of freedom.

From Section 8.6 of the textbook.

$$z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \Rightarrow \bar{x} \mp z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

What is the t-distribution?

Confidence Interval for the Mean When σ^2 Unknown

Corollary 8.1: Let $X_1, X_2, \ldots, X_n$ be independent random variables that are all normal with mean μ and standard deviation σ. Let

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad \text{and} \qquad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

Then the random variable $T = \frac{\bar{X}-\mu}{S/\sqrt{n}}$ has a t-distribution with $v = n-1$ degrees of freedom.

From Section 8.6 of the textbook.

$\bar{x} \mp t_{\alpha/2} \frac{s}{\sqrt{n}}$

$z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \Rightarrow \bar{x} \mp z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

What is the t-distribution?

Confidence Interval for the Mean When σ^2 Unknown

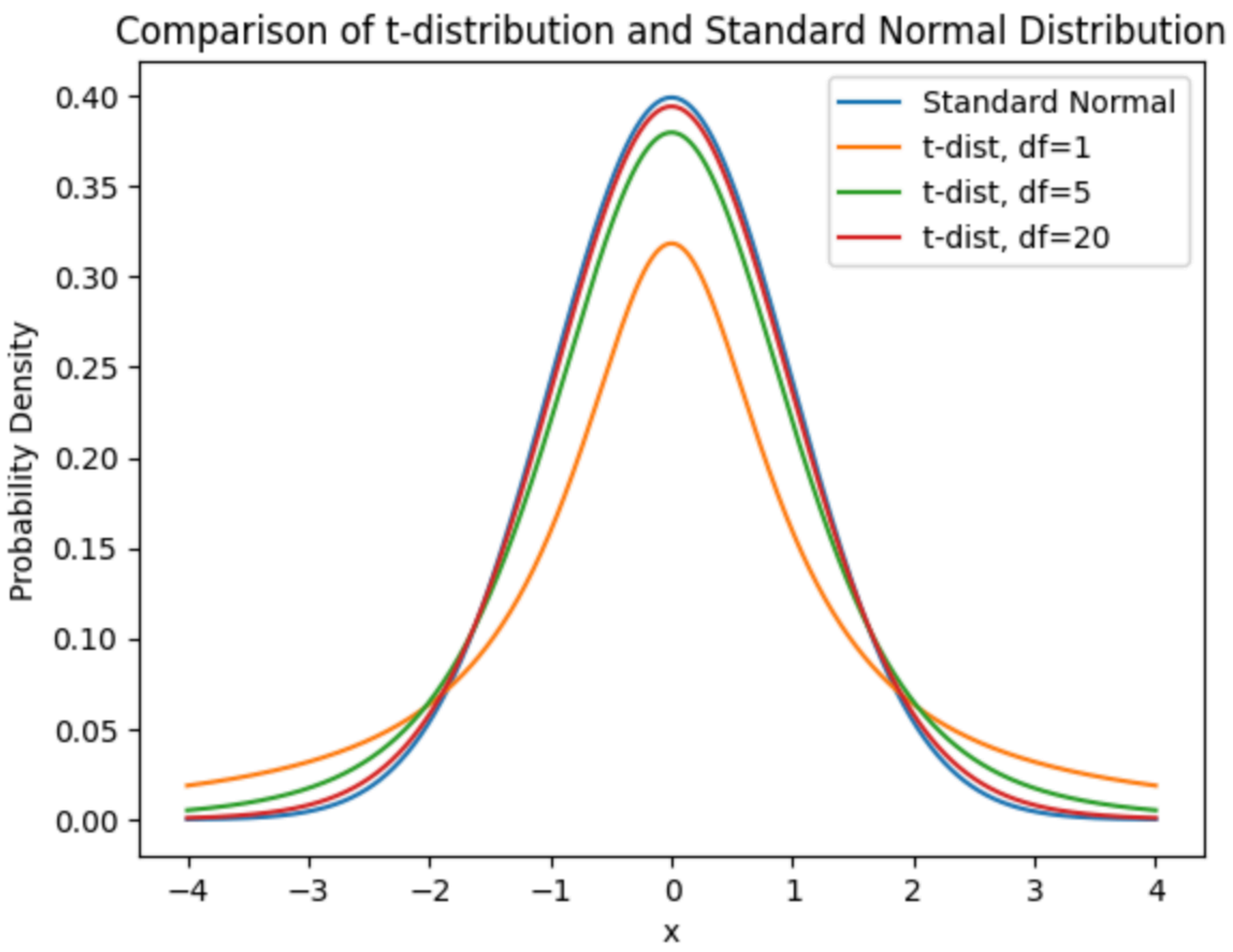
(Student's)
t distribution

William S Gosset
(aka *Student*)

While working at
Guinness

Confidence Interval for the Mean When σ^2 Unknown

(Student's)
t distribution



larger degrees of freedom ⟹ smaller variance

William S Gosset (aka *Student*)

While working at Guinness

Confidence Interval for the Mean When σ^2 Unknown

(Student's)
t distribution

```
from scipy.stats import t
t.cdf(1,20)
```

0.8353717114141455

the area to the left of 1 (i.e. $P(T \leq 1)$)

with degrees of freedom (dof) 20.

William S Gosset
(aka *Student*)

While working at Guinness

Confidence Interval for the Mean When σ^2 Unknown

How to compute the confidence interval?

Almost same as σ known case, except replace z with t and σ with s.

Confidence Interval on μ, σ^2 Unknown	If $\bar{x}$ and s are the mean and standard deviation of a random sample from a normal population with unknown variance σ^2, a $100(1-\alpha)\%$ confidence interval for μ is $$\bar{x} - t_{\alpha/2}\frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\alpha/2}\frac{s}{\sqrt{n}},$$ where $t_{\alpha/2}$ is the t-value with $v = n - 1$ degrees of freedom, leaving an area of $\alpha/2$ to the right.

Note that this is not a requirement for confidence interval for mean, when σ^2 is known.

Confidence Interval for the Mean When σ^2 Unknown

Example: Heights of men approximately normally distributed. Sample of size 9: 168, 176, 195, 182, 188, 150, 165, 170, 158. Find $95\,\%$ confidence interval for distribution mean. Use one of t.ppf(0.975,8)=2.306 or norm.ppf(0.975)=1.96

Solution:

$$\bar{x} \mp t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

Confidence Interval for the Mean When σ^2 Unknown

Example: Heights of men approximately normally distributed. Sample of size 9: 168, 176, 195, 182, 188, 150, 165, 170, 158. Find $95\ \%$ confidence interval for distribution mean. Use one of t.ppf(0.975,8)=2.306 or norm.ppf(0.975)=1.96

Solution:

$95\ \%$ confidence interval:

$$\bar{x} = \frac{1}{9}\sum_{i=1}^{9} X_i = 172.4,\ S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 = \frac{1}{8}\sum_{i=1}^{9}(X_i - 172.4)^2 = 206.03 \Rightarrow S \approx 14.35$$

Confidence Interval for the Mean When σ^2 Unknown

Example: Heights of men approximately normally distributed. Sample of size 9: 168, 176, 195, 182, 188, 150, 165, 170, 158. Find $95\ \%$ confidence interval for distribution mean. Use one of t.ppf(0.975,8)=2.306 or norm.ppf(0.975)=1.96

Solution:

$95\ \%$ confidence interval:

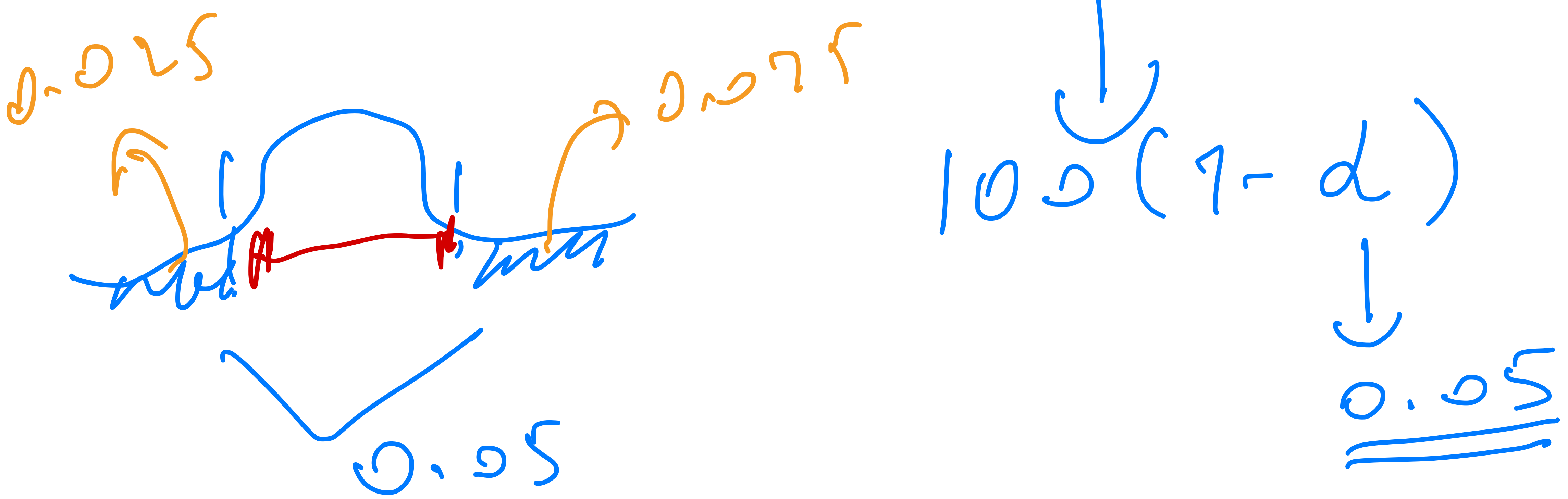
$$\bar{x} = \frac{1}{9}\sum_{i=1}^{9} X_i = 172.4,\ S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{8}\sum_{i=1}^{9}(X_i - 172.4)^2 = 206.03 \Rightarrow S \approx 14.35$$

$$\bar{x} \pm t_{\alpha/2} S/\sqrt{n} = 172.4 \pm t_{0.025} \times 14.35/\sqrt{9} = 172.4 \pm 2.306 \times 14.35/3 = 172.4 \pm 11.03$$

which implies $161.37 < \mu < 183.43$.

Confidence Interval for the Mean When σ^2 Unknown

Example: Heights of men approximately normally distributed. Sample of size 9: 168, 176, 195, 182, 188, 150, 165, 170, 158. Find $95\ \%$ confidence interval for distribution mean. Use one of t.ppf(0.975,8)=2.306 or norm.ppf(0.975)=1.96



Solution:

$95\ \%$ confidence interval:

$$\bar{x} = \frac{1}{9}\sum_{i=1}^{9} X_i = 172.4,\ S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{8}\sum_{i=1}^{9}(X_i - 172.4)^2 = 206.03 \Rightarrow S \approx 14.35$$

$$\bar{x} \pm t_{\alpha/2} S/\sqrt{n} = 172.4 \pm t_{0.025} \times 14.35/\sqrt{9} = 172.4 \pm 2.306 \times 14.35/3 = 172.4 \pm 11.03$$

which implies $161.37 < \mu < 183.43$.

Confidence Interval for the Mean When σ^2 Unknown

Example: Heights of men approximately normally distributed. Sample of size 9: 168, 176, 195, 182, 188, 150, 165, 170, 158. Find $95\ \%$ confidence interval for distribution mean. Use one of t.ppf(0.975,8)=2.306 or norm.ppf(0.975)=1.96

Solution:

$95\ \%$ confidence interval: $161.37 < \mu < 183.43$

Note:

If we were to incorrectly use $z_{0.025}$:

$$\bar{x} \pm z_{\alpha/2} S/\sqrt{n} = 172.4 \pm z_{0.025} \times 14.35/\sqrt{9} = 172.4 \pm 1.96 \times 14.35/3 = 172.4 \pm 9.38$$

$95\ \%$ confidence interval: $163.02 < \mu < 181.78$

Confidence Interval for the Mean When σ^2 Unknown

Example: Heights of men approximately normally distributed. Sample of size 9: 168, 176, 195, 182, 188, 150, 165, 170, 158. Find $95\,\%$ confidence interval for distribution mean. Use one of t.ppf(0.975,8)=2.306 or norm.ppf(0.975)=1.96

Solution:

$95\,\%$ confidence interval: $161.37 < \mu < 183.43$

Note:

If we were to incorrectly use $z_{0.025}$:

$$\bar{x} \pm z_{\alpha/2} S/\sqrt{n} = 172.4 \pm z_{0.025} \times 14.35/\sqrt{9} = 172.4 \pm 1.96 \times 14.35/3 = 172.4 \pm 9.38$$

$95\,\%$ confidence interval: $163.02 < \mu < 181.78$

we incorrectly get a narrower interval.