ECON 0150 | Economic Data Analysis

The economist's data analysis pipeline.

Part 4.1 | Extending The t-Test

One-Sample t-Test: Example (using sample stats)

How surprising would it be for the average wait time to be 10 minutes?

One-sample t-test:

- $H_0 : \mu = 10$
- $H_1 : \mu \neq 10$
- *n: 29*
- *t-stat:* $t = 2.401$
- *p-value: 0.0230*

```python
# Imports
import numpy as np
from scipy import stats
```

```python
# Sample Data
sample_mu = 10.864
pop_mu = 10    # null hypothesis
std_dev = 1.971
n = 30
```

```python
# Calculate t-statistic
t_stat = (sample_mu - pop_mu) / (std_dev / np.sqrt(n))
```

```python
# Calculate p-value
p_value = 2 * (1 - stats.t.cdf(abs(t_stat), df=n-1))
```

> so there is a 2.3% chance of seeing data this extreme given a true mean of 10

One-Sample t-Test: Example (using sample data)

How surprising would it be for the average wait time to be 10 minutes?

> *we can perform this very simply from raw data*

```
# Imports
import scipy.stats as stats
import numpy as np
```

> *load the observed data*

```
# Sample Data
data = [11.39, 13.55, 10.54, 12.45, 9.  , 12.34, 7.74, 10.97, 9.77, 10.03, 9.52,
    14.09, 8.01, 13.4 , 13.46, 9.43, 10.39, 10.32, 9.36, 7.42, 9.95, 13.72,
    8.14, 12.71, 12.43, 12.83, 11.46, 13.12, 10.12, 8.27]
```

> *define the null hypothesis*

```
# Null Hypothesis
null = 10
```

> *perform the test*

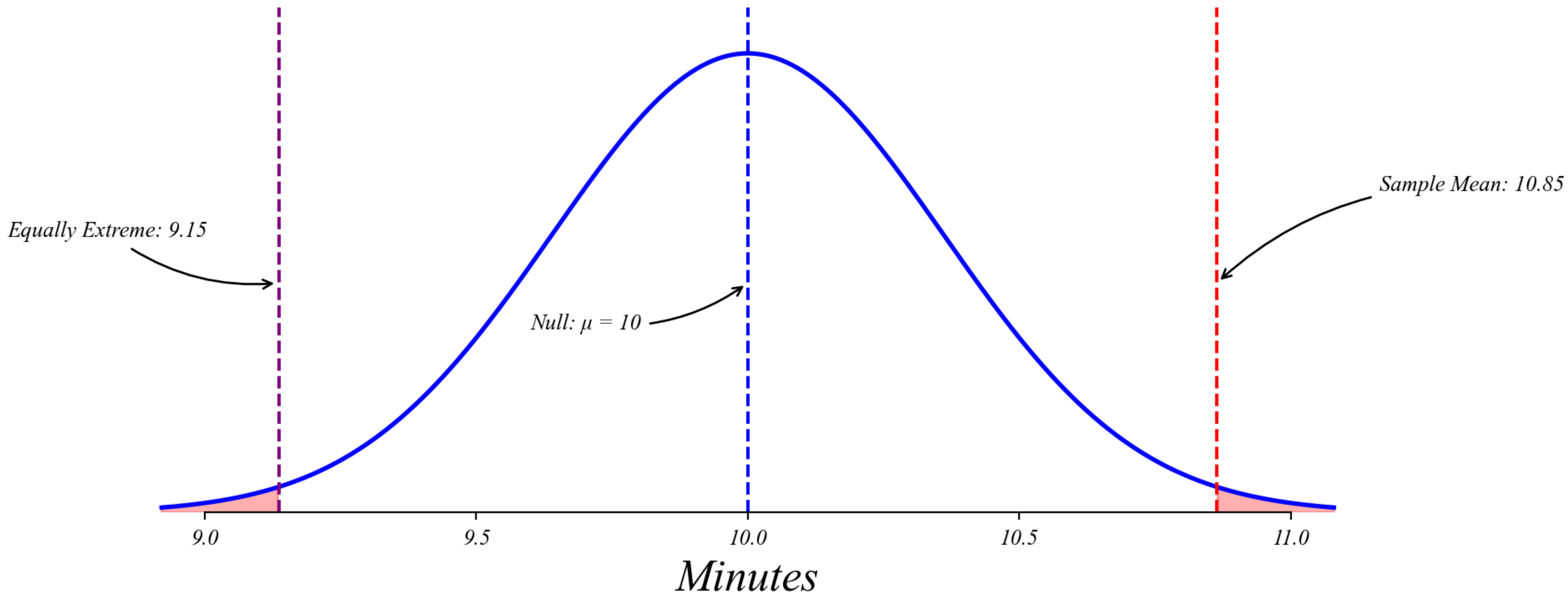
```
# Calculate t-statistic and p-value
t_stat, p_value = stats.ttest_1samp(data, null)
```

> *so there is a 2.3% chance of seeing data this extreme given a true mean of 10*

One-Sample t-Test: Example

How surprising would it be for the average wait time to be 10 minutes?

> *so there is a 2.3% chance of seeing data this extreme given a true mean of 10*



> *but why check both tails?*

> *why also consider 9.15, which is equally far from 10 but on the opposite side?*

One-Tail vs Two-Tail

… there are at least two key reasons for a two-tail test.

1. ***Scientific Integrity**: Since we're testing the hypothesis "$\mu = 10$" against "$\mu \neq 10$", we should be equally open to evidence in either direction.*

2. ***Statistical Reasoning**: If the true mean is 10, our sampling distribution is centered there, and random samples could fall on either side.*

t-Test: Differences in Means

Instead of asking whether $\mu = 10$ lets ask whether $\mu_1 = \mu_2$

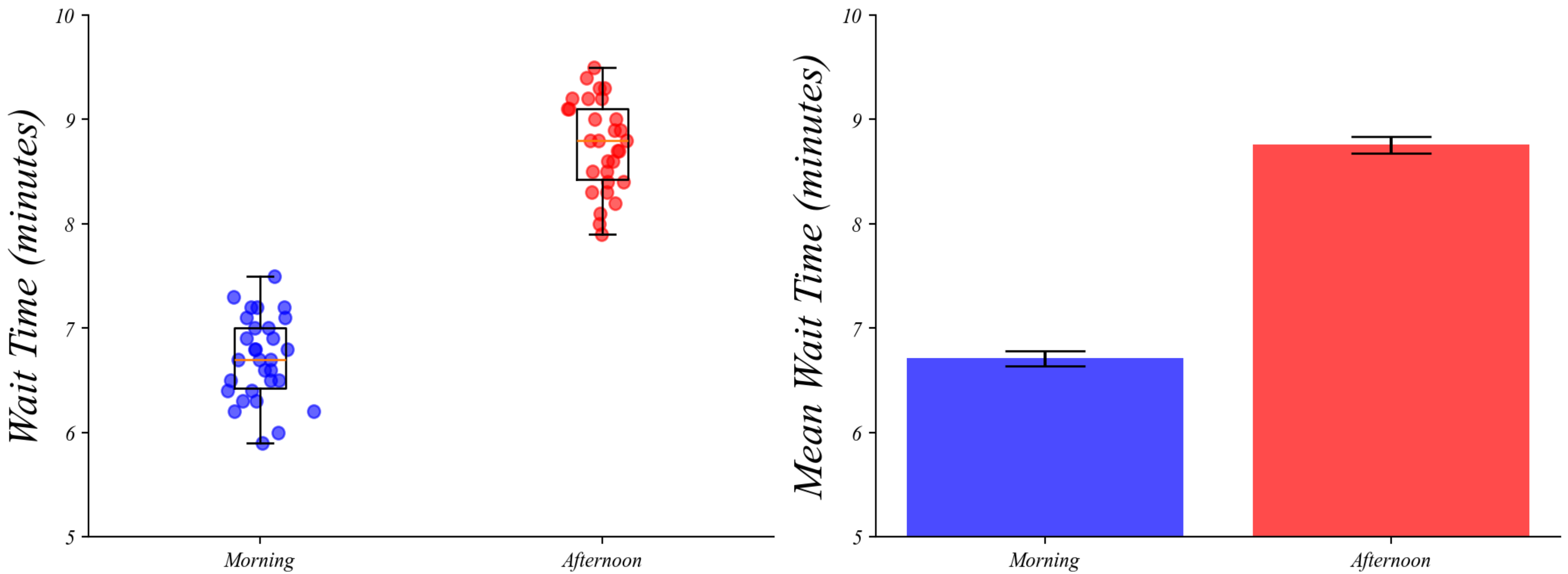
> *the core question: "are two means different from each other?"*

Common scenarios:

- *Wait times at different times of day*
- *Appointment lengths with different doctors*
- *Wages across different groups*
- *Treatment effects in experimental settings*

Two-Sample t-Test: Example

Are wait times different in the morning vs the afternoon?



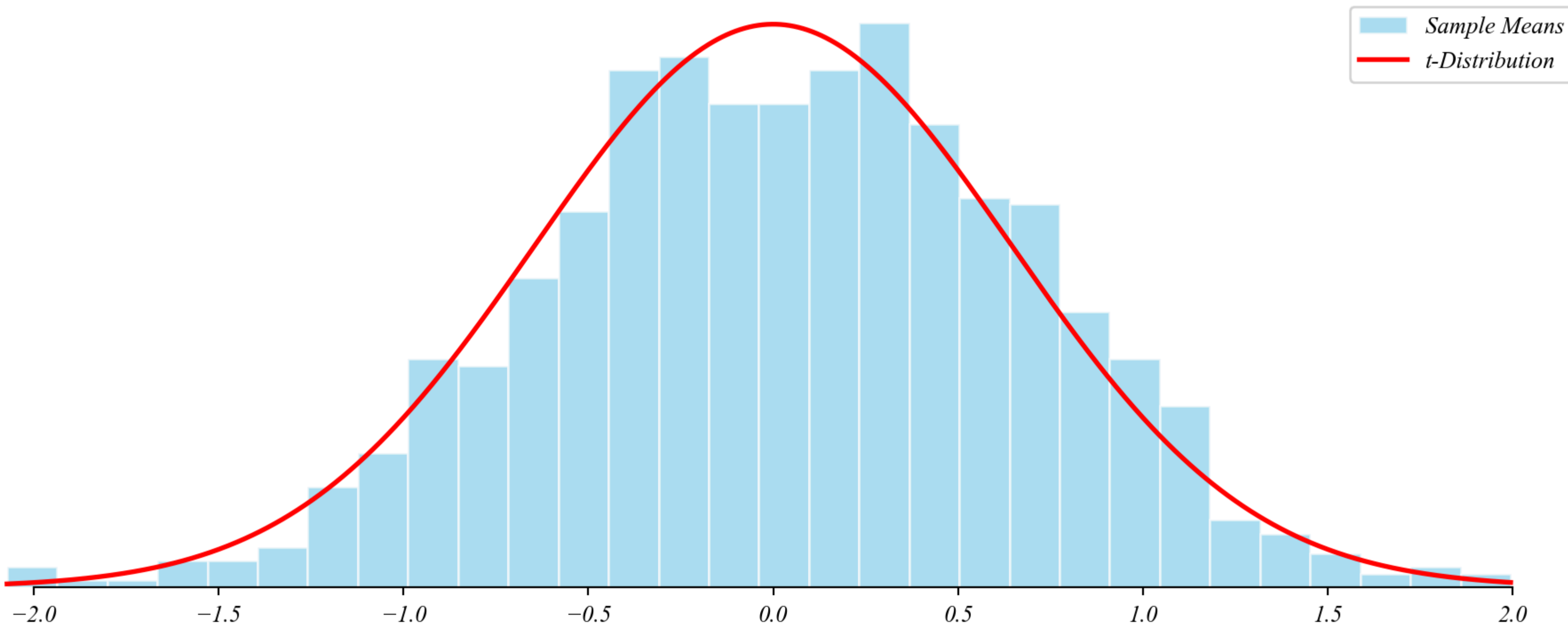
> is there a way to turn this into a one-sample t-Test that we're familiar with?

> what if we took the differences between sample menas?

Two-Sample t-Test: Example

If the true means are equal, the difference in sample means would center on zero.

> *I ran many sample means from two distributions with equal means*



> *the distribution of differences centers on zero*

> *but what is the standard error of the difference between two samples?*

The Math: Combining Standard Errors

What is the standard error of the difference in sample means?

For a single sample mean:

$$SE_{\bar{x}} = \sqrt{\frac{s^2}{n}}$$

For the difference in sample means:

$$SE_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

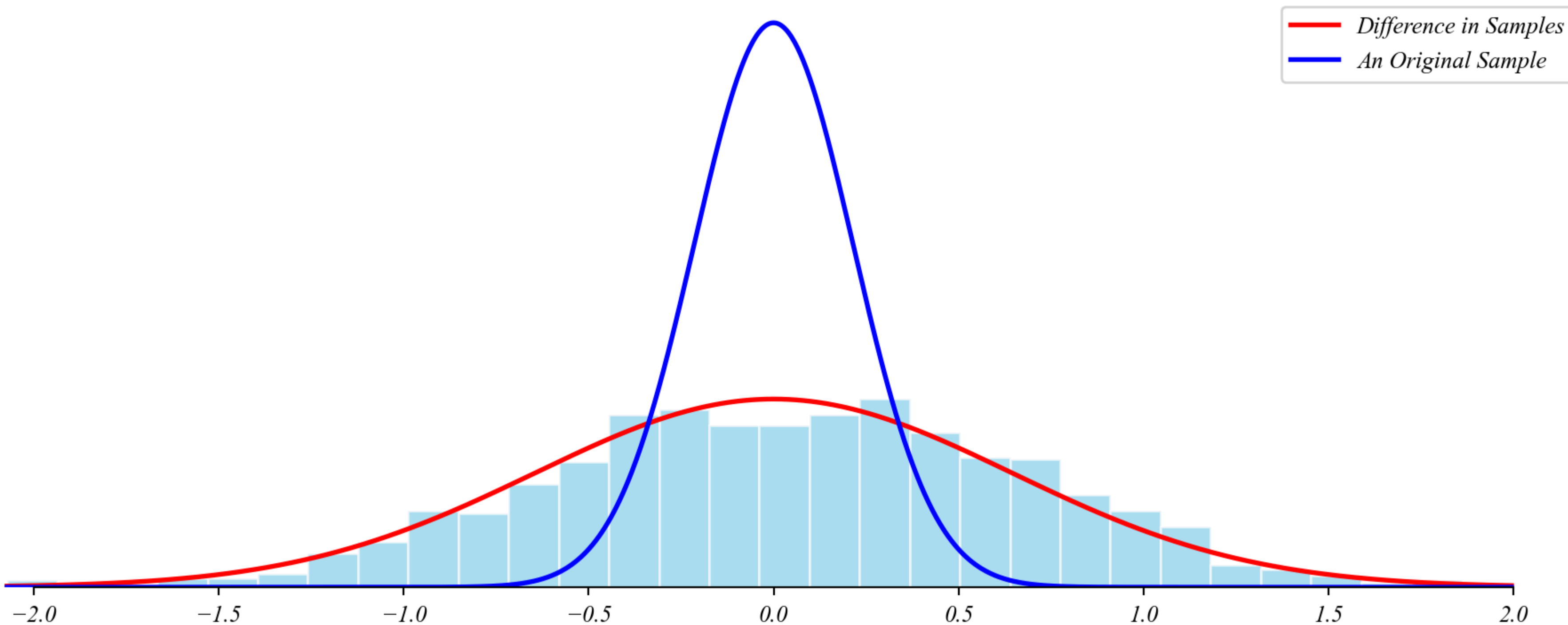
> *two sources of sampling variation combine, but not additively*

> *if variances and sample sizes are equal, then* $SE_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{2s^2}{n}} = \frac{s\sqrt{2}}{\sqrt{n}}$

Two-Sample t-Test: Example

If the true means are equal, the difference in sample means would center on zero.

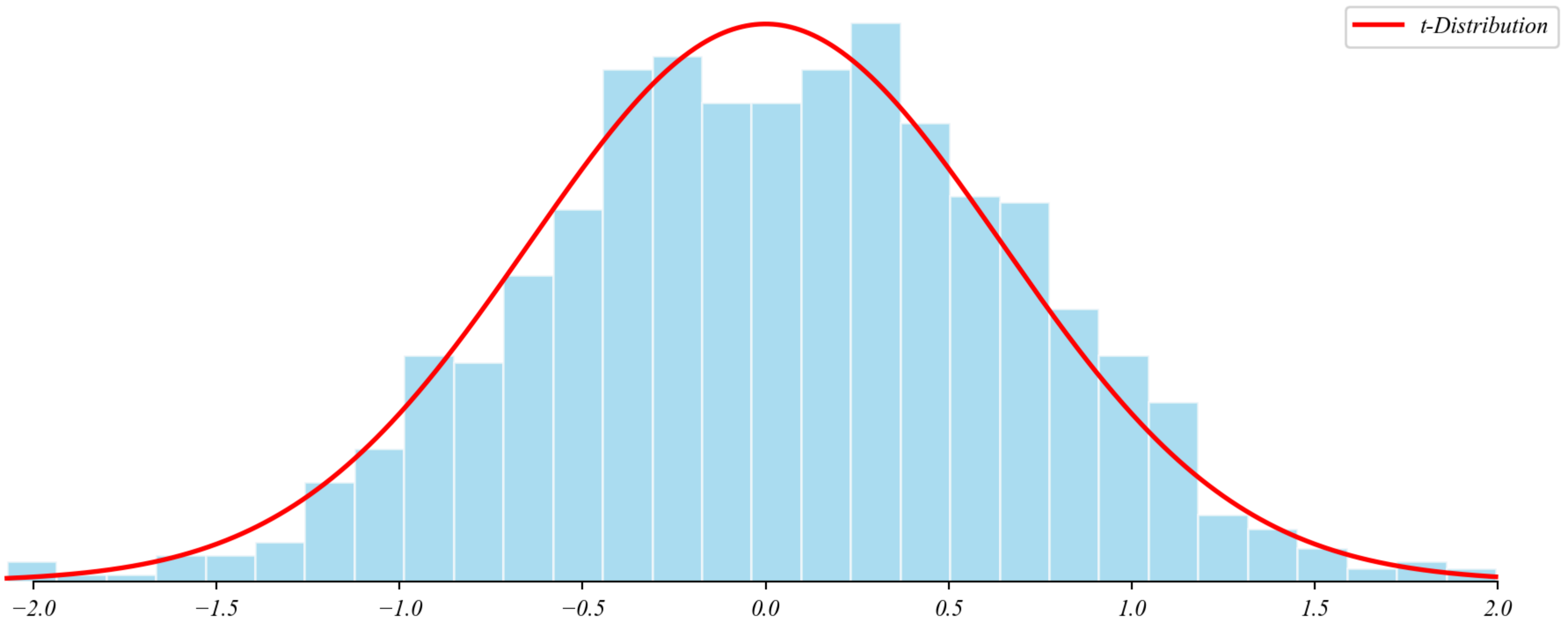
> *so the distribution of the differences will be wider than either sample*



Two-Sample t-Test: Example

If the true means are equal, the difference in sample means would center on zero.

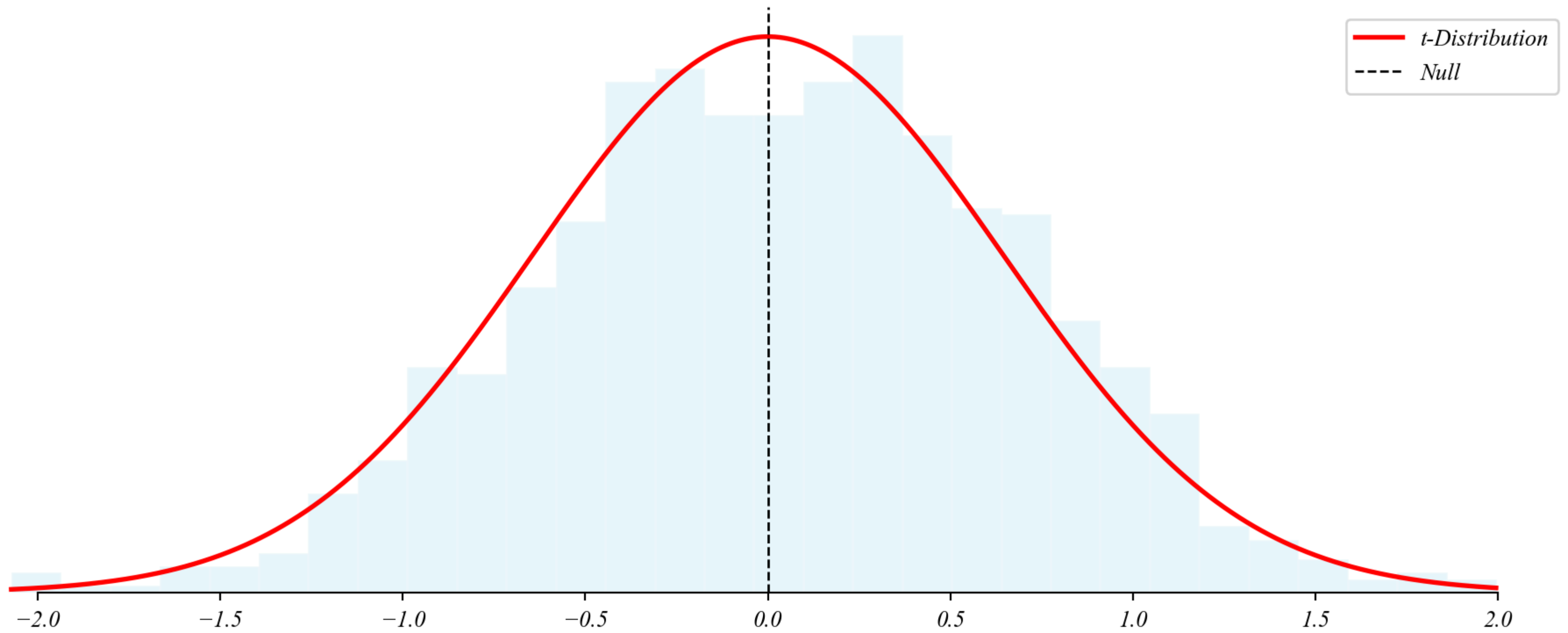
> *the null hypothesis comparing two means is typically:* $\mu_1 - \mu_2 = 0$



Two-Sample t-Test: Example

If the true means are equal, the difference in sample means would center on zero.

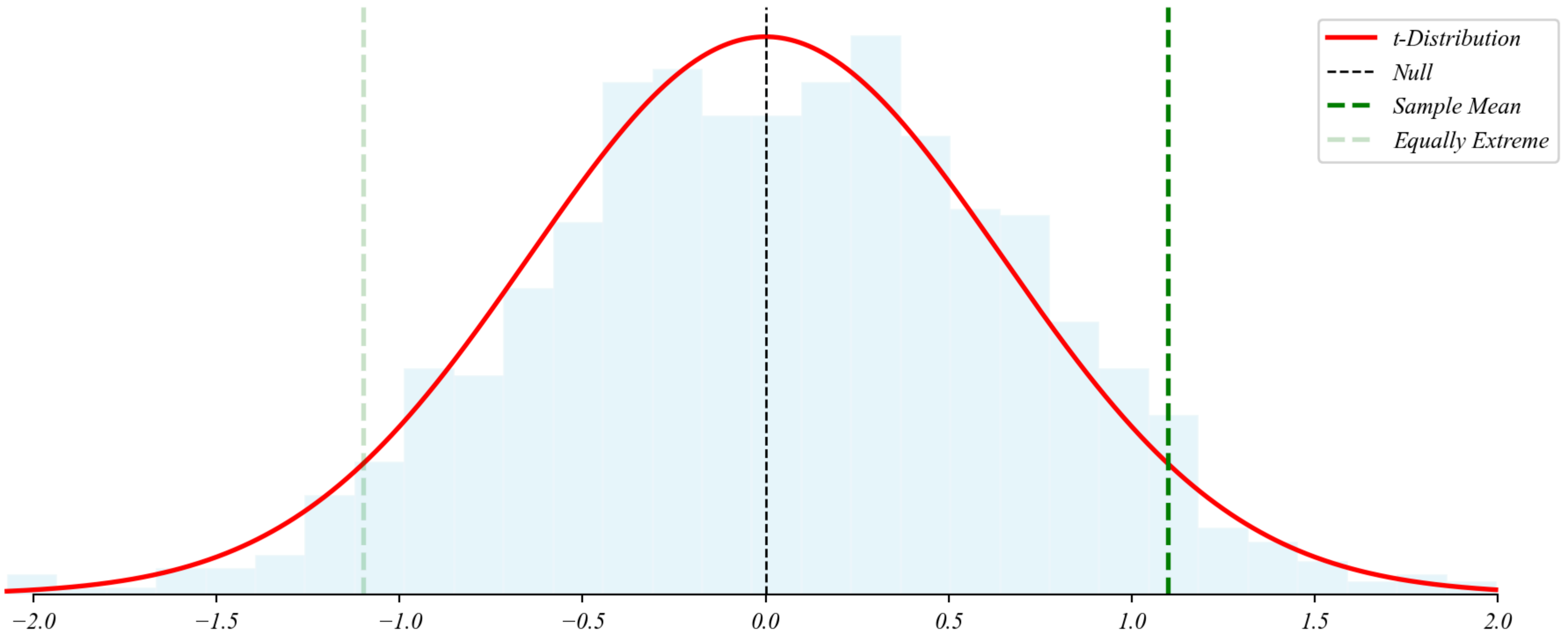
> *the null hypothesis comparing two means is typically:* $\mu_1 - \mu_2 = 0$



Two-Sample t-Test: Example

If the true means are equal, the difference in sample means would center on zero.

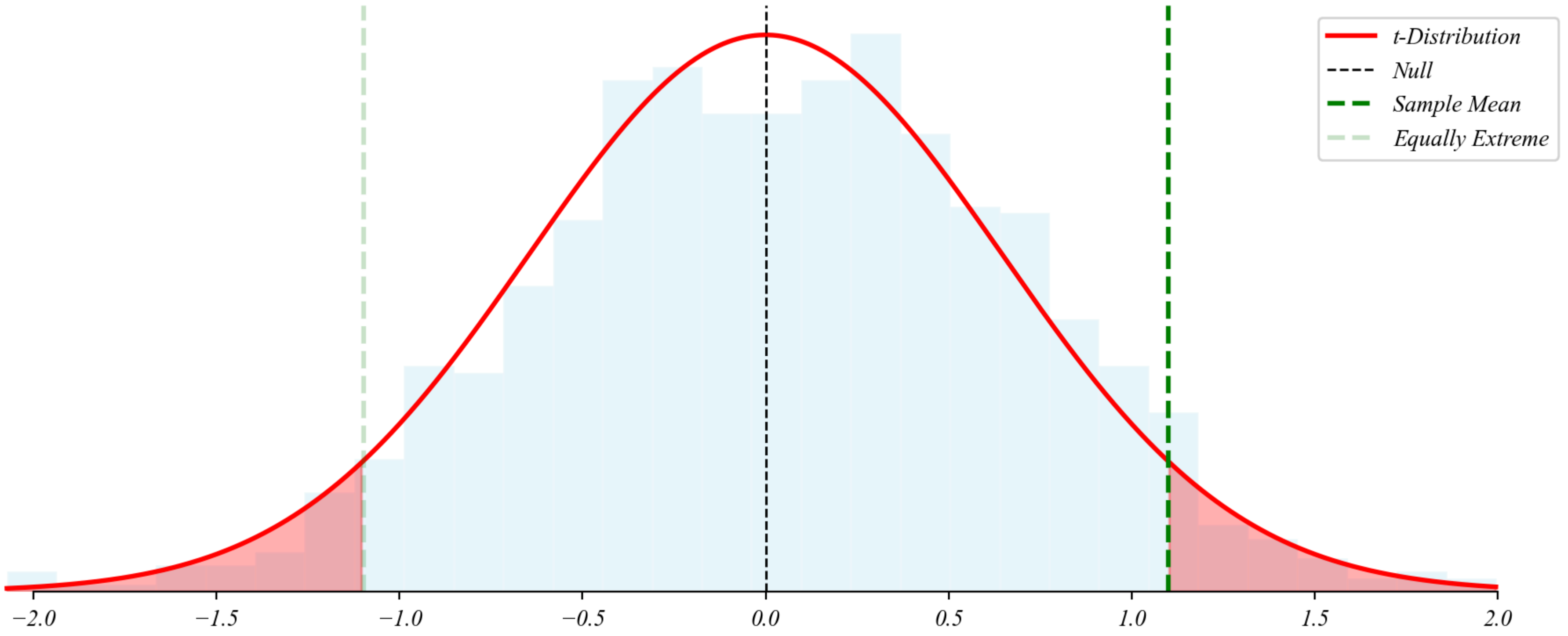
> *if the observed difference is* $\bar{x}_1 - \bar{x}_2 = 1.1 \ldots$



Two-Sample t-Test: Example

If the true means are equal, the difference in sample means would center on zero.

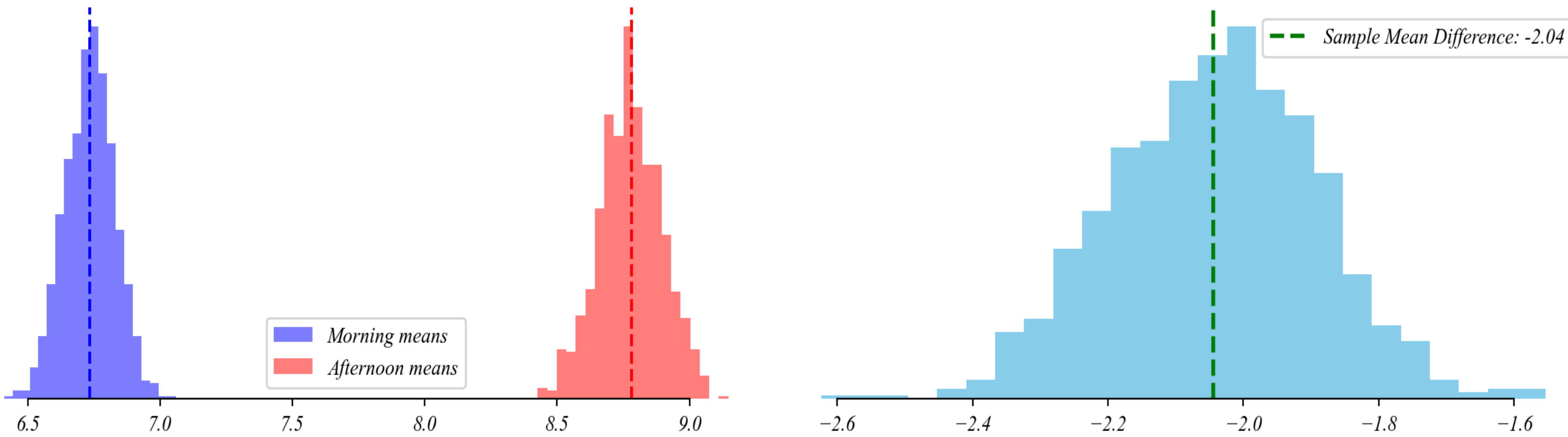
> *if the observed difference is* $\bar{x}_1 - \bar{x}_2 = 1.1\ldots$



> *like before, a p-value quantifies how surprising would be a big difference*

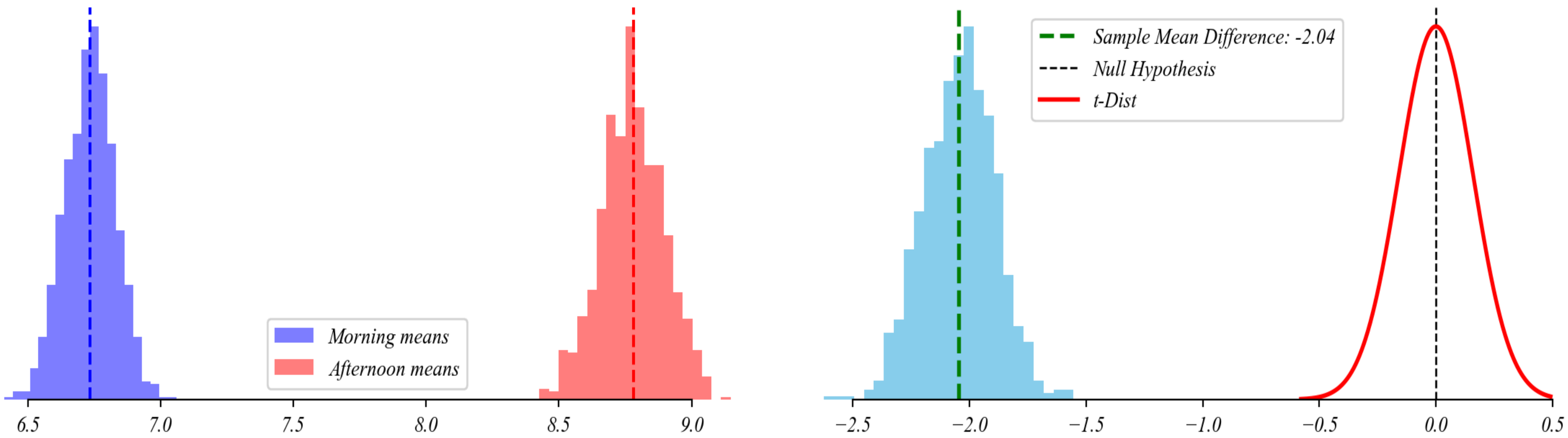
Two-Sample t-Test: Example

If the true means are not equal, the difference in sample means would not center on zero.



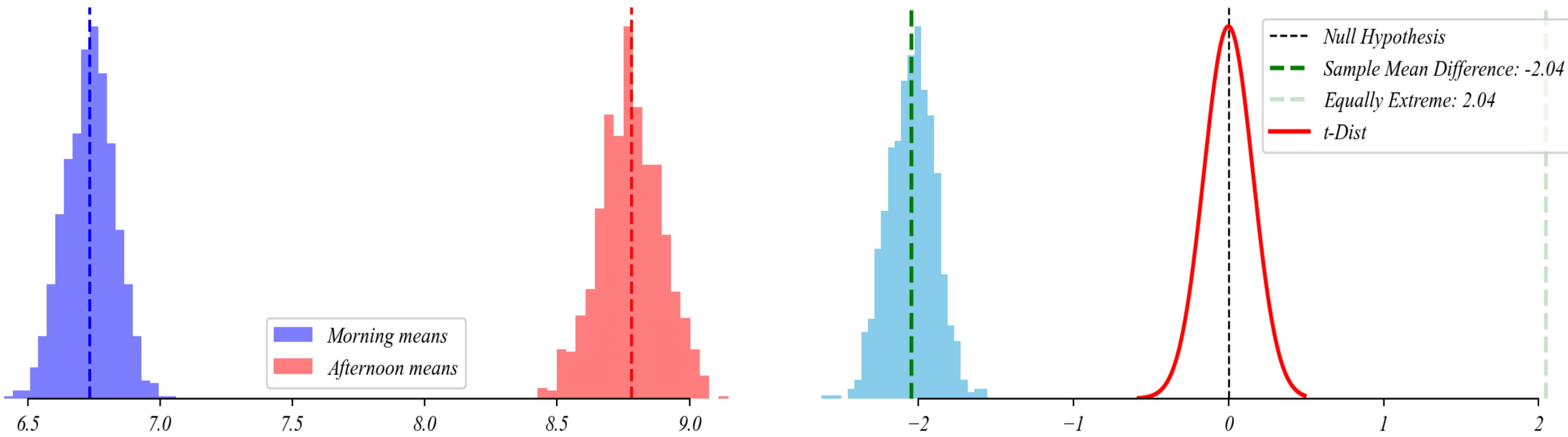
Two-Sample t-Test: Example

If the true means are not equal, the difference in sample means would not center on zero.



Two-Sample t-Test: Example

If the true means are not equal, the difference in sample means would not center on zero.



> the p-value tells us how often we'd see differences this extreme by chance

> here it looks like the area is basically zero

Two-Sample t-test: Step by Step

Example: Morning vs. Afternoon Wait Times

> load the data

```
morning = [
    7.2, 6.8, 5.9, 7.5, 6.5, 6.0, 7.1, 6.3, 6.7, 6.4,
    6.9, 7.2, 6.2, 6.8, 7.0, 6.5, 7.3, 6.6, 7.1, 6.7,
    6.3, 6.9, 6.4, 7.0, 6.6, 6.2, 6.8, 7.2, 6.5, 6.7
]

afternoon = [
    8.5, 9.2, 8.1, 8.8, 9.5, 8.3, 7.9, 9.0, 8.7, 9.3,
    8.6, 9.1, 8.4, 8.9, 9.4, 8.2, 9.0, 8.5, 8.8, 9.2,
    8.7, 9.1, 8.3, 8.9, 9.3, 8.0, 8.6, 9.2, 8.4, 8.8
]
```

> perform two-sample t-test using scipy.stats

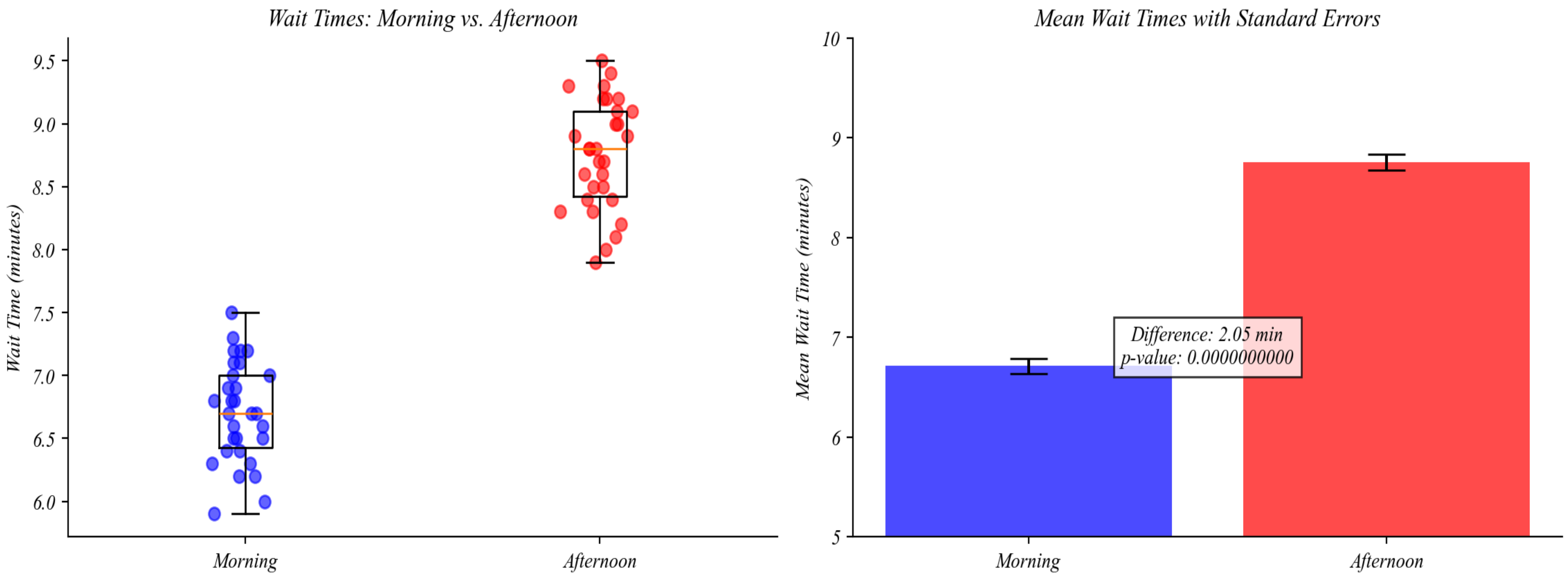
```
t_stat, p_value = stats.ttest_ind(morning, afternoon, equal_var=False)
```

> small p-value = surprising difference (reject null hypothesis of equal means)

> large p-value = difference could easily happen by chance (fail to reject null)

Visualizing the Difference

Morning vs. Afternoon Wait Times



> wait times are longer in the afternoon than the morning (p<0.001)

> reject the null that the means are equal

Key Insights

Connecting t-tests and regression

One-sample t-test:

- *Tests the sample mean against a specific null value*
- *Next time: regression with only an intercept*

Two-sample t-test:

- *Tests if the difference in sample means is zero*
- *Next time: regression with an intercept and one dummy variable*

Common Applications in Economics

Two-sample t-tests are one of the most common tests in economic research!

Labor Economics:

- *Wage gaps between different demographic groups*
- *Employment effects of policy changes*

Development Economics:

- *Impact of interventions on economic outcomes*
- *Differences between treatment and control groups*

Financial Economics:

- *Comparing returns across different time periods*
- *Testing market efficiency*

Looking Forward

Connecting to regression and extending to multiple groups

Next time:

- *Connecting to regression*
- *Comparing more than two groups (ANOVA)*

Coming soon:

- *Multiple regression*
- *Controlling for confounding variables*
- *Interaction effects*

> *all built on the same fundamental statistical framework*