0.1 Learning Objectives

By the end of this lecture, you will be able to:

- Distinguish one-sample from two-sample t-tests
- Apply the five-step hypothesis testing procedure to a real dataset
- Understand the structure and goals of the course project

1. Worked Example: Hypothesis Testing

1.1 Quick Review: Two-Sample t-Test

Last lecture we compared **two groups**

1.1 Quick Review: Two-Sample t-Test

bikers vs. drivers:

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

1.1 Quick Review: Two-Sample t-Test

1.1 Quick Review: Two-Sample t-Test

bikers vs. drivers:

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

But what if we only have **one sample** and want to test it against a **known or claimed value**?

1.1 Quick Review: Two-Sample t-Test

1.1 Quick Review: Two-Sample t-Test

bikers vs. drivers:

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

But what if we only have **one sample** and want to test it against a **known or claimed value**?

Two-sample t-test

"Is there a difference between groups A and B?"

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

One-sample t-test

"Does the population mean equal some claimed value μ_0?"

$$t = \frac{\overline{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

1.1 Quick Review: Two-Sample t-Test

1.1 Quick Review: Two-Sample t-Test

bikers vs. drivers:

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

But what if we only have **one sample** and want to test it against a **known or claimed value**?

Two-sample t-test

"Is there a difference between groups A and B?"

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

One-sample t-test

"Does the population mean equal some claimed value μ_0?"

$$t = \frac{\overline{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

1.1 Quick Review: Two-Sample t-Test

Today: a worked example of each.

1.2 The Scenario

A coffee shop owner claims that a new espresso machine produces shots in **under 25 seconds** on average.

1.2 The Scenario

A coffee shop owner claims that a new espresso machine produces shots in **under 25 seconds** on average.

A barista is skeptical. She times $n = 12$ consecutive shots during a morning rush:

Shot	1	2	3	4	5	6
Time (s)	24.2	26.1	25.8	23.9	27.0	24.5
Shot	7	8	9	10	11	12
Time (s)	25.3	26.7	24.8	25.1	26.4	25.0

1.2 The Scenario

A coffee shop owner claims that a new espresso machine produces shots in **under 25 seconds** on average.

A barista is skeptical. She times $n = 12$ consecutive shots during a morning rush:

Shot	1	2	3	4	5	6
Time (s)	24.2	26.1	25.8	23.9	27.0	24.5
Shot	7	8	9	10	11	12
Time (s)	25.3	26.7	24.8	25.1	26.4	25.0

Is there evidence that the true mean extraction time is greater than 25 seconds?

1.3 Step 1: State the Hypotheses

Recall: The **null hypothesis** (H_0) is the "no effect" claim. The **alternative** (H_1) is what we're trying to find evidence for.

1.3 Step 1: State the Hypotheses

Recall: The **null hypothesis** (H_0) is the "no effect" claim. The **alternative** (H_1) is what we're trying to find evidence for.

$$H_0 : \mu = 25 \text{ seconds}$$

$$H_1 : \mu > 25 \text{ seconds}$$

1.3 Step 1: State the Hypotheses

Recall: The **null hypothesis** (H_0) is the "no effect" claim. The **alternative** (H_1) is what we're trying to find evidence for.

$$H_0 : \mu = 25 \quad \text{seconds}$$

$$H_1 : \mu > 25 \quad \text{seconds}$$

This is a **one-tailed test**

1.3 Step 1: State the Hypotheses

we only care if the mean is *greater* than 25.

1.3 Step 1: State the Hypotheses

we only care if the mean is *greater* than 25.

We'll use significance level $\alpha = 0.05$.

1.4 Step 2: Compute Sample Statistics

From our 12 observations:

1.4 Step 2: Compute Sample Statistics

From our 12 observations:

$$\overline{x} = \frac{24.2 + 26.1 + 25.8 + \cdots + 25.0}{12} = \frac{304.8}{12} = 25.4$$

1.4 Step 2: Compute Sample Statistics

From our 12 observations:

$$\overline{x} = \frac{24.2 + 26.1 + 25.8 + \cdots + 25.0}{12} = \frac{304.8}{12} = 25.4$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})^2$$

1.4 Step 2: Compute Sample Statistics

From our 12 observations:

$$\overline{x} = \frac{24.2 + 26.1 + 25.8 + \cdots + 25.0}{12} = \frac{304.8}{12} = 25.4$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})^2$$

$$s^2 = \frac{1}{11}\left[(24.2 - 25.4)^2 + (26.1 - 25.4)^2 + \cdots + (25.0 - 25.4)^2\right]$$

1.4 Step 2: Compute Sample Statistics

From our 12 observations:

$$\overline{x} = \frac{24.2 + 26.1 + 25.8 + \cdots + 25.0}{12} = \frac{304.8}{12} = 25.4$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

$$s^2 = \frac{1}{11}\left[(24.2 - 25.4)^2 + (26.1 - 25.4)^2 + \cdots + (25.0 - 25.4)^2\right]$$

$$s^2 \approx 0.937 \quad \rightarrow \quad s \approx 0.968$$

1.5 Step 3: Calculate the Test Statistic

Recall: For a one-sample t-test against a hypothesized mean μ_0:

$$t = \frac{\overline{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

1.5 Step 3: Calculate the Test Statistic

Recall: For a one-sample t-test against a hypothesized mean μ_0:

$$t = \frac{\overline{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

$$t = \frac{25.4 - 25}{\frac{0.968}{\sqrt{12}}} = \frac{0.4}{0.279} \approx 1.43$$

1.5 Step 3: Calculate the Test Statistic

Recall: For a one-sample t-test against a hypothesized mean μ_0:

$$t = \frac{\overline{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

$$t = \frac{25.4 - 25}{\frac{0.968}{\sqrt{12}}} = \frac{0.4}{0.279} \approx 1.43$$

Degrees of freedom: $\mathrm{df} = n - 1 = 11$

1.6 Step 4: Find the p-Value

We need: $P(t \geq 1.43)$ with $\mathrm{df} = 11$

1.6 Step 4: Find the p-Value

We need: $P(t \geq 1.43)$ with $\mathrm{df} = 11$

Using a t-table or software:

```
1 - pt(1.43, df=11)
```

Vim

```
1 -
scipy.stats.t.cdf(1.43,
df=11)
```

Python

1.6 Step 4: Find the p-Value

We need: $P(t \geq 1.43)$ with $\mathrm{df} = 11$

Using a t-table or software:

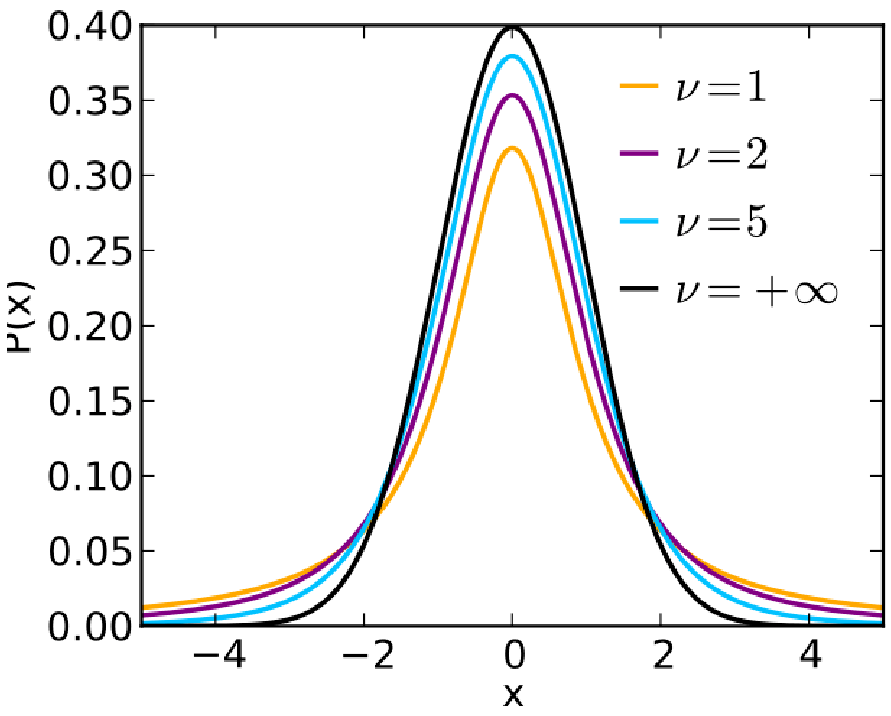
```r
1 - pt(1.43, df=11)
```

Vim

```python
1 -
scipy.stats.t.cdf(1.43,
df=11)
```

Python

$$p \approx 0.09$$



1.7 Step 5: Make a Decision

Our decision rule: reject H_0 if $p < \alpha$.

1.7 Step 5: Make a Decision

Our decision rule: reject H_0 if $p < \alpha$.

$$p = 0.09 > \alpha = 0.05$$

1.7 Step 5: Make a Decision

Our decision rule: reject H_0 if $p < \alpha$.

$$p = 0.09 > \alpha = 0.05$$

We fail to reject H_0.

1.7 Step 5: Make a Decision

Our decision rule: reject H_0 if $p < \alpha$.

$$p = 0.09 > \alpha = 0.05$$

We fail to reject H_0.

Note: At the 5% significance level, there is **not enough evidence** to conclude that the mean extraction time exceeds 25 seconds.

The sample mean of 25.4s is higher, but with $n = 12$ observations and this variability, the difference could plausibly be due to chance.

1.8 What If We Had More Data?

The barista collects $n = 50$ shots and finds $\overline{x} = 25.4$, $s = 0.97$.

1.8 What If We Had More Data?

The barista collects $n = 50$ shots and finds $\overline{x} = 25.4$, $s = 0.97$.

$$t = \frac{25.4 - 25}{\frac{0.97}{\sqrt{50}}} = \frac{0.4}{0.137} \approx 2.92$$

1.8 What If We Had More Data?

The barista collects $n = 50$ shots and finds $\overline{x} = 25.4$, $s = 0.97$.

$$t = \frac{25.4 - 25}{\frac{0.97}{\sqrt{50}}} = \frac{0.4}{0.137} \approx 2.92$$

$p \approx 0.003$ (with $\mathrm{df} = 49$)

1.8 What If We Had More Data?

The barista collects $n = 50$ shots and finds $\overline{x} = 25.4$, $s = 0.97$.

$$t = \frac{25.4 - 25}{\frac{0.97}{\sqrt{50}}} = \frac{0.4}{0.137} \approx 2.92$$

$p \approx 0.003$ (with $\mathrm{df} = 49$)

Now $p < 0.05$, so we **reject H_0**.

1.8 What If We Had More Data?

The barista collects $n = 50$ shots and finds $\overline{x} = 25.4$, $s = 0.97$.

$$t = \frac{25.4 - 25}{\frac{0.97}{\sqrt{50}}} = \frac{0.4}{0.137} \approx 2.92$$

$p \approx 0.003$ (with $\mathrm{df} = 49$)

Now $p < 0.05$, so we **reject H_0**.

Note: Same sample mean, same standard deviation — but **more data** gives us more confidence. Larger n means smaller standard error $\frac{s}{\sqrt{n}}$, which makes the t-statistic larger.

1.9 Summary: Five Steps of Hypothesis Testing

1. **State hypotheses:** H_0 and H_1, choose α

1.9 Summary: Five Steps of Hypothesis Testing

1. **State hypotheses:** H_0 and H_1, choose α
2. **Compute sample statistics:** $\overline{x}$, s

1.9 Summary: Five Steps of Hypothesis Testing

1. **State hypotheses:** H_0 and H_1, choose α
2. **Compute sample statistics:** $\overline{x}$, s
3. **Calculate the test statistic:** $t = \frac{\overline{x} - \mu_0}{\frac{s}{\sqrt{n}}}$

1.9 Summary: Five Steps of Hypothesis Testing

1. **State hypotheses:** H_0 and H_1, choose α
2. **Compute sample statistics:** $\overline{x}$, s
3. **Calculate the test statistic:** $t = \frac{\overline{x}-\mu_0}{\frac{s}{\sqrt{n}}}$
4. **Find the p-value:** from the t-distribution with $n-1$ df

1.9 Summary: Five Steps of Hypothesis Testing

1. **State hypotheses:** H_0 and H_1, choose α
2. **Compute sample statistics:** $\overline{x}$, s
3. **Calculate the test statistic:** $t = \frac{\overline{x}-\mu_0}{\frac{s}{\sqrt{n}}}$
4. **Find the p-value:** from the t-distribution with $n-1$ df
5. **Decide:** if $p < \alpha$, reject H_0; otherwise, fail to reject

1.10 Try It: Two-Sample Problem

Two sections of CHE 2A take the same final exam. Section A ($n_1 = 35$) averages 74 with $s_1 = 12$. Section B ($n_2 = 40$) averages 78 with $s_2 = 10$.

Is there a significant difference between the sections at $\alpha = 0.05$?

1.10 Try It: Two-Sample Problem

Two sections of CHE 2A take the same final exam. Section A ($n_1 = 35$) averages 74 with $s_1 = 12$. Section B ($n_2 = 40$) averages 78 with $s_2 = 10$.

Is there a significant difference between the sections at $\alpha = 0.05$?

Think (1 minute): Set up H_0 and H_1, then compute the t-statistic.

1.10 Try It: Two-Sample Problem

Two sections of CHE 2A take the same final exam. Section A ($n_1 = 35$) averages 74 with $s_1 = 12$. Section B ($n_2 = 40$) averages 78 with $s_2 = 10$.

Is there a significant difference between the sections at $\alpha = 0.05$?

Think (1 minute): Set up H_0 and H_1, then compute the t-statistic.

Discuss with your neighbor (1 minute): What did you get? One-tailed or two-tailed?

1.10 Try It: Two-Sample Problem

Two sections of CHE 2A take the same final exam. Section A ($n_1 = 35$) averages 74 with $s_1 = 12$. Section B ($n_2 = 40$) averages 78 with $s_2 = 10$.

Is there a significant difference between the sections at $\alpha = 0.05$?

Think (1 minute): Set up H_0 and H_1, then compute the t-statistic.

Discuss with your neighbor (1 minute): What did you get? One-tailed or two-tailed?

Try it yourself

Talk to your neighbor and try to solve this problem.

1.11 Solution

$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{74 - 78}{\sqrt{\frac{144}{35} + \frac{100}{40}}} = \frac{-4}{\sqrt{4.11 + 2.5}} = \frac{-4}{2.57} \approx -1.56$$

1.11 Solution

$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{74 - 78}{\sqrt{\frac{144}{35} + \frac{100}{40}}} = \frac{-4}{\sqrt{4.11 + 2.5}} = \frac{-4}{2.57} \approx -1.56$$

Two-tailed test ($H_1 : \mu_1 \neq \mu_2$), so $p = 2 \times P(t \leq -1.56)$

1.11 Solution

$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{74 - 78}{\sqrt{\frac{144}{35} + \frac{100}{40}}} = \frac{-4}{\sqrt{4.11 + 2.5}} = \frac{-4}{2.57} \approx -1.56$$

Two-tailed test ($H_1 : \mu_1 \neq \mu_2$), so $p = 2 \times P(t \leq -1.56)$

$p \approx 0.12 > 0.05$

1.11 Solution

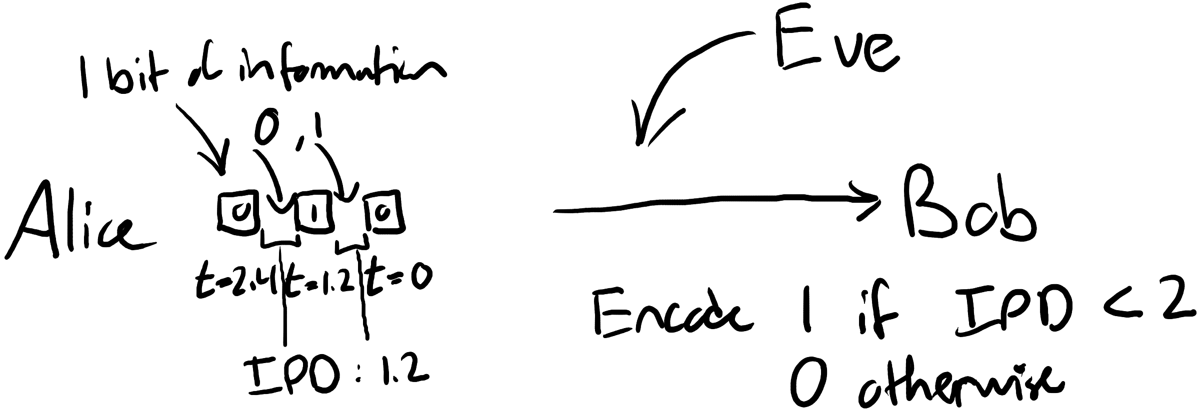
$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{74 - 78}{\sqrt{\frac{144}{35} + \frac{100}{40}}} = \frac{-4}{\sqrt{4.11 + 2.5}} = \frac{-4}{2.57} \approx -1.56$$

Two-tailed test ($H_1 : \mu_1 \neq \mu_2$), so $p = 2 \times P(t \leq -1.56)$

$p \approx 0.12 > 0.05$

Fail to reject H_0. Not enough evidence to say the sections differ.

2. Course Project: Covert Timing Channels



2.1 The Scenario

Alice wants to send a secret message to **Bob**. **Eve** monitors all network traffic.

2.1 The Scenario

Alice wants to send a secret message to **Bob**. **Eve** monitors all network traffic.

Alice and Bob set up a normal Skype call (the **overt** channel). Alice hides secret bits in the **timing between packets**.

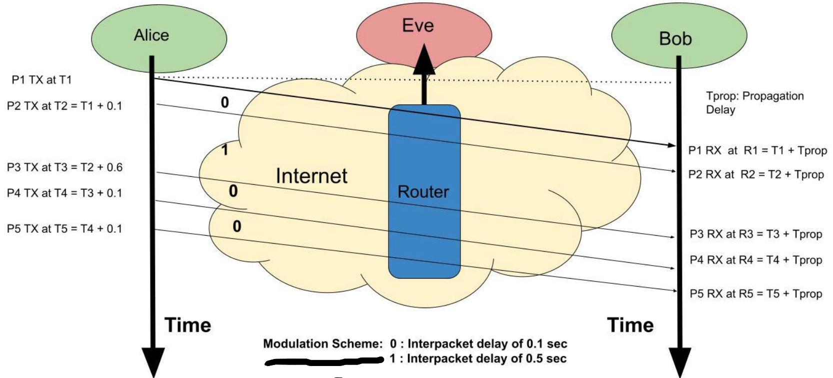
2.1 The Scenario

Alice wants to send a secret message to **Bob**. **Eve** monitors all network traffic.

Alice and Bob set up a normal Skype call (the **overt** channel). Alice hides secret bits in the **timing between packets**.

The **inter-packet delay** (IPD) is the time between consecutive packets. Alice encodes bits by choosing different delays:

Bit	IPD	Meaning
0	0.1 sec	Short delay
1	0.5 sec	Long delay



FM

AM

Bit

0 $IPD \sim N(0.1, 0.05)$

1 $IPD \sim N(0.5, 0.05)$

2.2 The Challenge

Your goal: design encoding schemes and evaluate whether Eve can **statistically distinguish** the covert traffic from normal traffic.

2.2 The Challenge

Your goal: design encoding schemes and evaluate whether Eve can **statistically distinguish** the covert traffic from normal traffic.

You'll use tools from this course to both **design** and **detect** covert channels:

2.2 The Challenge

Your goal: design encoding schemes and evaluate whether Eve can **statistically distinguish** the covert traffic from normal traffic.

You'll use tools from this course to both **design** and **detect** covert channels:

Histograms — visually compare IPD distributions

2.2 The Challenge

Your goal: design encoding schemes and evaluate whether Eve can **statistically distinguish** the covert traffic from normal traffic.

You'll use tools from this course to both **design** and **detect** covert channels:

Histograms — visually compare IPD distributions

QQ-plots — compare whether two datasets come from the same distribution

Simulation — model the source, buffer, and sender as a system

2.2 The Challenge

Your goal: design encoding schemes and evaluate whether Eve can **statistically distinguish** the covert traffic from normal traffic.

You'll use tools from this course to both **design** and **detect** covert channels:

Histograms — visually compare IPD distributions

QQ-plots — compare whether two datasets come from the same distribution

Simulation — model the source, buffer, and sender as a system

Distributions — exponential and uniform IPDs

2.3 Project Structure

Part 1: Design

(35 pts)

Implement encoding schemes and compare IPD distributions

Part 2: Detection

(40 pts)

Use QQ-plots to evaluate whether covert traffic is detectable

Part 3: Implementation

(25 pts)

Simulate the full system with a packet buffer

2.3 Project Structure

2.3 Project Structure

Part 1: Design

(35 pts)

Implement encoding schemes and compare IPD distributions

Part 2: Detection

(40 pts)

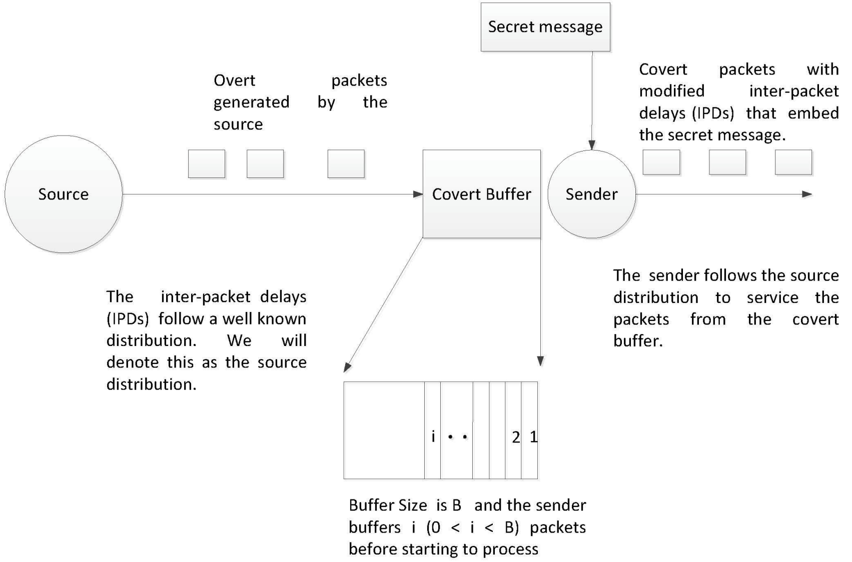
Use QQ-plots to evaluate whether covert traffic is detectable

Part 3: Implementation

(25 pts)

Simulate the full system with a packet buffer

- **Due date:** Check Canvas for the deadline
- **Submission:** Jupyter notebook (`.ipynb`) on Canvas
- **You will be given:** a CSV file with the overt packet stream, a character-to-binary table, and the secret message



2.3 Project Structure

2.3 Project Structure

Part 1: Design

(35 pts)

Implement encoding schemes and compare IPD distributions

Part 2: Detection

(40 pts)

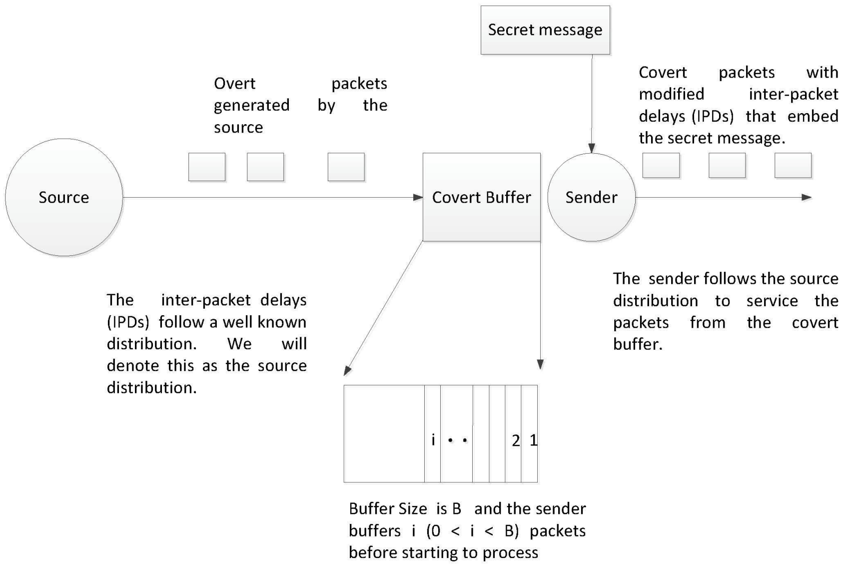
Use QQ-plots to evaluate whether covert traffic is detectable

Part 3: Implementation

(25 pts)

Simulate the full system with a packet buffer

- **Due date:** Check Canvas for the deadline
- **Submission:** Jupyter notebook (`.ipynb`) on Canvas
- **You will be given:** a CSV file with the overt packet stream, a character-to-binary table, and the secret message



2.3 Project Structure

Start early

2.3 Project Structure

Part 3 involves simulation that requires careful debugging.

2.4 Recap

Today we covered:

- **One-sample t-test:** compare a sample mean to a claimed value μ_0
- **Two-sample t-test:** compare two group means (from last lecture)
- **Five steps:** hypotheses, statistics, t-statistic, p-value, decision
- **More data** (larger n) gives more statistical power
- **Course project:** covert timing channels — design, detect, and simulate