

# Hypothesis Testing

Predictive Analytics

## The Tea Lady



**Figure:** The classic story of the lady who claimed she could tell if milk was added before or after the tea.

- ▶ This thought experiment was introduced by Ronald A. Fisher in his 1935 book, "The Design of Experiments".
- ▶ It illustrates the fundamental principles of hypothesis testing.

## Experimental Setup

- ▶ The lady is presented with 8 cups of tea.
- ▶ 4 cups have milk added first.
- ▶ 4 cups have tea added first.
- ▶ The cups are presented in a random order.
- ▶ The lady's task is to identify the 4 cups that had milk added first.

## Null and Alternative Hypotheses

- ▶ **Null Hypothesis ( $H_0$ ):** The lady has no ability to distinguish between the two methods of preparation. Any correct identification is due to random chance.
- ▶ **Alternative Hypothesis ( $H_1$ ):** The lady does have the ability to distinguish between the two methods.

## How many ways can she choose the cups?

The total number of ways to choose 4 cups from a set of 8 is given by the combination formula:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

In this case,  $n = 8$  and  $k = 4$ :

$$\binom{8}{4} = \frac{8!}{4!(8-4)!} = \frac{8 \times 7 \times 6 \times 5}{4 \times 3 \times 2 \times 1} = 70$$

There are **70** possible ways for the lady to choose the 4 cups.

## What is the probability of her guessing correctly?

- ▶ If the null hypothesis is true (she is guessing), there is only **one** way for her to choose all 4 cups correctly.
- ▶ The probability of this happening by chance is:

$$P(\text{all correct}) = \frac{\text{Number of ways to choose correctly}}{\text{Total number of ways to choose}} = \frac{1}{70} \approx 0.014$$

- ▶ This probability is the **p-value**. It's the probability of observing the data (or more extreme data) if the null hypothesis is true.

## What is the probability of her guessing exactly 3 correctly?

To get exactly 3 correct, she must choose 3 of the 4 correct cups  
AND 1 of the 4 incorrect cups.

- ▶ Ways to choose 3 correct cups from 4:

$$\binom{4}{3} = \frac{4!}{3!1!} = 4$$

- ▶ Ways to choose 1 incorrect cup from 4:

$$\binom{4}{1} = \frac{4!}{1!3!} = 4$$

The total number of ways to get exactly 3 correct is  $4 \times 4 = 16$ .  
The probability is:

$$P(3 \text{ correct}) = \frac{16}{70} \approx 0.2286$$

# What is Hypothesis Testing?

Hypothesis testing is a formal procedure for investigating our ideas about the world using statistics. It is used to test a claim about a population parameter using data from a sample.

The two competing hypotheses are:

- ▶ **Null Hypothesis ( $H_0$ ):** A statement of no effect or no difference.
- ▶ **Alternative Hypothesis ( $H_1$ ):** The claim we are trying to find evidence for.

# Summary of Errors

	Reality	
Decision	$H_0$ is True	$H_0$ is False
Reject $H_0$	Type I Error	Correct Decision
Fail to Reject $H_0$	Correct Decision	Type II Error