You want to know if **gender** (male/female) changes which product people like (Product A or Product B).

**What is a hypothesis?**

When doing a test like this, you start with two statements:

1. **Null hypothesis (H0):**  
   This means *“There is NO relationship.”*  
   So, gender **does NOT affect** product preference.
2. **Alternative hypothesis (H1):**  
   This means *“There IS a relationship.”*  
   So, gender **DOES affect** product preference.

**What is a p-value?**

* The **p-value** tells you **how likely** it is to get the results you saw **if the null hypothesis were true**.
* Think of it as the chance that your results happened **just by random luck**.

**What is alpha (α)?**

* Alpha is a **threshold** you set to decide when to say a result is "significant."
* Commonly, we pick **0.05** (which means 5%).
* If the **p-value is less than alpha (p < 0.05)**, it means the results are **unlikely to be just luck** → so we **reject the null hypothesis**.

**What does “reject the null hypothesis” mean?**

* You say: “The data shows a real relationship.”
* In this case: **Gender affects product preference**.

**What if p-value is greater than alpha?**

* You say: “There is not enough evidence to say there is a relationship.”
* So, **we do NOT find gender affects product preference** — but that doesn’t 100% prove it doesn’t.
* It just means the data doesn’t give strong enough evidence to say it does.

**Example:**

* Alpha = 0.05 (5%)
* p-value = 0.03 (3%) → **3% chance this happened by luck** → since 3% < 5%, we reject null → gender influences preference.
* p-value = 0.10 (10%) → 10% chance this happened by luck → since 10% > 5%, we fail to reject null → gender does NOT seem to influence preference.

**Why not just say “accept null” when p-value > alpha?**

Because failing to find evidence **is not the same as proving no relationship exists**.  
You just say: “We can’t say there’s an effect based on this data.”

**📊 1. Chi-Square Test**

* **Purpose**: Tests relationships between **categorical variables**.
* **Used For**:
  + **Chi-Square Test of Independence** – Are two categorical variables related?
  + **Chi-Square Goodness of Fit** – Does a sample match a population distribution?
* **Data Type**: Categorical (e.g., gender, yes/no, product type).
* **Assumption**: Expected frequency in each cell should be ≥ 5.
* **Example**: Is there a relationship between gender and product preference?

**📈 2. T-Test**

* **Purpose**: Compares **means of groups** when **sample size is small (< 30)** or **population standard deviation is unknown**.
* **Used For**:
  + **One-sample t-test** – Is the sample mean different from a known value?
  + **Independent t-test (two-sample)** – Are the means of two groups different?
  + **Paired t-test** – Compare means of the same group at two different times.
* **Data Type**: Continuous (e.g., height, weight, score).
* **Assumption**: Data should be normally distributed, especially for small samples.
* **Example**: Is the average test score of class A different from class B?

**📉 3. Z-Test**

* **Purpose**: Compares **means or proportions** when **sample size is large (≥ 30)** and **population standard deviation is known**.
* **Used For**:
  + **One-sample z-test** – Is the sample mean different from a known mean?
  + **Two-sample z-test** – Are the means of two groups different?
  + **Proportion z-test** – Compare sample proportion with a population proportion.
* **Data Type**: Continuous or proportion.
* **Assumption**: Normally distributed data or large sample size (Central Limit Theorem).
* **Example**: Is the average salary in one city different from another when population variances are known?

**🧪 4. "P-Test"**

* **Not a separate test**, but often used to refer to:
  + **The p-value** resulting from **any** statistical test (t, z, chi2, ANOVA).
  + A p-value helps determine whether the result is statistically significant (e.g., p < 0.05).

**🧭 Summary Table:**

| **Test** | **Data Type** | **Purpose** | **Assumptions** |
| --- | --- | --- | --- |
| Chi-Square | Categorical | Test independence or fit | Expected freq ≥ 5 |
| T-Test | Continuous | Compare means (small sample) | Normality, unknown σ |
| Z-Test | Continuous / Proportion | Compare means or proportions (large sample) | Known σ or large sample size |
| "P-Test" | N/A | Refers to use of **p-value** | Interprets significance, not a test |