• Null Hypothesis (H0): The default assumption that no real difference or effect exists.

• Alternative Hypothesis (H1): The hypothesis that a difference or effect does exist.

• Degrees of Freedom (df): A parameter that affects the shape of distributions like the t-distribution; often (n1 + n2 − 2) for a two-sample t-test with equal variances.

• p-value: Probability of observing a test statistic at least as extreme as the one we got, if the null hypothesis is true.

• Critical Value: The boundary value of a test statistic beyond which we reject H0. Determined by α and df.

• Rejection Zone: Range of test statistic values so unlikely under H0 that we reject H0.

• Type I Error (α): Rejecting H0 when H0 is actually true (false positive).

• Type II Error (β): Failing to reject H0 when H0 is actually false (false negative).

• Power (1−β): Probability of rejecting a false H0.

**. Null Hypothesis (H₀)**

The **null hypothesis** represents the default assumption that there is **no effect, no difference, or no relationship** between variables.

**Example:**

A pharmaceutical company develops a new drug for lowering blood pressure. The null hypothesis would state:

* **H₀:** "The new drug has no effect on blood pressure compared to the existing drug."

**2. Alternative Hypothesis (H₁)**

The **alternative hypothesis** is the statement that there **is an effect, difference, or relationship** between variables.

**Example:**

Continuing from the drug example:

* **H₁:** "The new drug lowers blood pressure more effectively than the existing drug."

If sufficient evidence supports this claim, we reject H₀ in favor of H₁.

**3. Degrees of Freedom (df)**

Degrees of freedom **(df)** determine the shape of the probability distribution used in hypothesis testing. It is related to the sample size and the number of parameters being estimated.

**Example (Two-Sample t-test):**

If we compare blood pressure levels between two groups (patients taking the new drug vs. those taking the existing drug) with sample sizes n1=30 and n2=30, the degrees of freedom for a **two-sample t-test** assuming equal variances is:

df=(n1+n2−2)=(30+30−2)=58

The **df** determines the exact shape of the **t-distribution**, which affects critical values and p-values.

**4. p-value**

The **p-value** is the probability of obtaining a test statistic at least as extreme as the observed value, assuming H₀ is true.

* A **small p-value** (typically ≤ 0.05) suggests strong evidence against H₀, leading us to **reject** it.
* A **large p-value** (> 0.05) suggests weak evidence against H₀, meaning we **fail to reject** it.

**Example:**

* In the drug study, after running a statistical test, we obtain **p = 0.02**. Since p<0.05, we reject H₀ and conclude that the new drug **significantly reduces blood pressure**.

**5. Critical Value**

The **critical value** is a threshold that defines the boundary for rejecting H₀. It depends on:

* The **significance level (α)**, typically 0.05.
* The **degrees of freedom (df)**.

**Example (t-test, two-tailed, α = 0.05, df = 58):**

Using a t-table or calculator, the critical value for a **two-tailed** t-test with **α = 0.05** and **df = 58** is approximately **±2.00**.

* If our test statistic (t-value) falls beyond ±2.00, we **reject H₀**.
* Otherwise, we **fail to reject H₀**.

**6. Rejection Zone**

The **rejection zone** consists of values that are so extreme under H₀ that we consider them unlikely and reject H₀.

**Example:**

For a **two-tailed test at α = 0.05**, the rejection zones are:

* Left tail: t<−2.00
* Right tail: t>2.00

If our t-value falls in these regions, we reject H₀.

**7. Type I Error (False Positive, α)**

A **Type I Error** occurs when we reject H₀ when it is actually **true**.

**Example:**

Suppose the new drug **actually has no effect** on blood pressure, but due to random sampling variation, our test produces **p = 0.03**, leading us to reject H₀ falsely. This is a **Type I Error**.

* The probability of making a Type I Error is **α** (commonly set at 0.05 or 5%).
* This means we **incorrectly conclude the drug is effective when it isn’t**.

**8. Type II Error (False Negative, β)**

A **Type II Error** occurs when we **fail to reject H₀ when H₀ is false**.

**Example:**

Suppose the new drug **actually does lower blood pressure**, but due to a small sample size or high variability, we obtain **p = 0.08**, leading us to fail to reject H₀. This is a **Type II Error**.

* The probability of making a Type II Error is **β**.
* This means we **fail to detect a real effect**.

**9. Power of the Test (1 − β)**

The **power of a test** is the probability of **correctly rejecting H₀ when it is false**.

Power=1−β

A **high power (close to 1)** means the test is good at detecting true effects.

* Power increases with **larger sample size, lower variability, and stronger true effects**.

**Example:**

If a test has **80% power**, it means there is an **80% probability of correctly detecting a real effect** and only **20% chance of missing it (β = 0.20).**

**Final Summary Table**

| **Concept** | **Definition** | **Example** |
| --- | --- | --- |
| **H₀ (Null Hypothesis)** | No real effect exists | New drug has no effect on blood pressure |
| **H₁ (Alternative Hypothesis)** | A real effect exists | New drug lowers blood pressure |
| **Degrees of Freedom (df)** | Affects shape of test distribution | df=n1+n2−2 for t-test |
| **p-value** | Probability of getting observed result if H₀ is true | p=0.02 → reject H₀ |
| **Critical Value** | Boundary where H₀ is rejected | tcrit=±2.00 for α = 0.05 |
| **Rejection Zone** | Extreme values leading to H₀ rejection | t<−2.00 or t>2.00 |
| **Type I Error (α)** | Rejecting H₀ when it is true (False Positive) | Saying drug works when it doesn’t |
| **Type II Error (β)** | Failing to reject H₀ when it is false (False Negative) | Saying drug doesn’t work when it does |
| **Power (1−β)** | Probability of detecting a true effect | High power means lower Type II Error |

**Understanding Effect Size & Cohen’s d in Your Vole Population Analysis**

**Effect Size Calculation**

Effect size quantifies the **magnitude** of the difference between two groups. One common measure is **Cohen’s d**, which is calculated as:

d=Mean difference/Pooled standard deviation

From your R code:

r

CopyEdit

effect\_size <- (mean(island\_voles) - mean(mainland\_voles)) / common\_sd

print(paste("Cohen’s d =", round(effect\_size, 3)))

Here:

* MeanIsland=3.75
* MeanMainland=3.57
* Common SD=0.5

d=(3.75−3.57)/0.5=0.18/0.5=0.36

Thus, **Cohen’s d = 0.36**.

**Interpreting Cohen’s d**

Cohen’s d is interpreted as follows:

* **0.2 - 0.3** → Small effect
* **0.5** → Medium effect
* **0.8+** → Large effect

Since **d = 0.36**, this suggests a **small-to-moderate effect size**—the difference between **island and mainland vole populations’ dorsal sizes is present but not very strong**.

In a two-sided test with α = 0.05 and df = 20, how would you find the critical t value from standard t-tables or R’s qt() function?

**1. Using Standard t-Tables**

In a **two-tailed test**, the critical t-value corresponds to **α/2 in each tail**.  
For **α = 0.05**, we split it into two tails:

α/2=0.05/2=0.025

From **t-tables**, look up **df = 20** and **p = 0.025 (upper tail)**.

* The **critical t-value is approximately**: t0.025,20=2.086
* The rejection regions are **t < -2.086** or **t > 2.086**.

**2. Using R’s qt() Function**

R provides the **qt()** function to find the critical t-value:

# Find the critical t-value

qt(1 - 0.05/2, df = 20)

**Output:**

[1] 2.085963

So, the **critical t-value is ≈ ±2.086**.

**Conclusion**

* If **t-statistic > 2.086** or **t-statistic < -2.086**, **reject H₀**.
* Otherwise, **fail to reject H₀**.

Short Answer: Why do heavier tails in at- distribution with lower df require a bigger |t| to reject H0?

Heavier tails in a **t-distribution with lower degrees of freedom (df)** mean that extreme values are more likely compared to a **normal distribution**. This increased variability requires a **larger |t|** to reach statistical significance because:

1. **Greater Uncertainty**: With fewer samples (lower df), the estimate of the standard error is less precise, making it harder to distinguish real effects from random variation.
2. **Wider Critical Regions**: The critical t-value is larger for smaller df, meaning a more extreme t-statistic is needed to fall into the rejection zone.
3. **More Conservative Test**: The test compensates for increased variability by demanding stronger evidence (a larger |t|) before rejecting **H₀**.

Thus, as **df increases**, the t-distribution approaches the normal distribution, requiring a **smaller |t|** to reject **H₀**.

Questions:

•How does increasing n (samplesize) affect Type II error? Why?

* Increasing **sample size (n)** **reduces Type II error (β)**.
* A larger sample provides **more precise estimates** with **smaller standard errors**, making it easier to detect true differences.
* This increases the **test’s power (1−β1 )**, reducing the chance of **failing to reject H₀ when H₀ is actually false**.

•How does increasing trueEffect (the real difference in means) change the power?

* A larger true effect size makes it easier to detect a difference, increasing power.
* A greater separation between means leads to a larger t-statistic, making rejection of H₀ more likely.
* This reduces Type II error (β), as detecting a real difference becomes easier.

•What happens to TypeI error if your raise or lower α?

Raising α (e.g., from 0.05 to 0.10):

* Increases the probability of rejecting H₀ when it is true.
* Higher Type I error (false positive rate).
* More findings may be statistically significant but possibly false alarms.

Lowering α (e.g., from 0.05 to 0.01):

* Decreases the probability of rejecting H₀ when it is true.
* Lower Type I error but higher Type II error (risk of missing true effects).
* Makes the test more conservative, requiring stronger evidence to reject H₀.