Inference Worksheet

Step 1: State the Hypotheses

We begin by specifying the hypotheses:

- **Null Hypothesis (H_0):** The population mean is equal to 3 hours ($\mu = 3$).
  Example: Children watch TV for 3 hours per day.
- **Alternative Hypothesis (H_1):** The population mean differs from 3. Depending on research focus, H_1 can be:
  - **Two-tailed:** $\mu \neq 3$
  - **One-tailed greater:** $\mu > 3$
  - **One-tailed less:** $\mu < 3$

This is like a courtroom analogy: the defendant (null hypothesis) is assumed innocent until there is sufficient evidence against them.

Step 2: Set the Criteria for the Decision

We choose the **significance level (α)**, usually 5% (0.05).

- This represents the maximum probability of committing a **Type I error** (rejecting a true H_0).
- In the figure provided, α is split into **two tails** (α/2 each side), creating rejection regions.
- For medical research, a stricter α (e.g., 0.01) is often used.

Definition:

- **Level of significance (α):** The threshold for deciding whether sample evidence is strong enough to reject H_0.

Step 3: Compute the Test Statistic

We compare the observed sample mean to the hypothesised population mean.

Suppose we collected a sample and found the average TV viewing = **4 hours**.

We then calculate the **test statistic**:

$$t = \frac{\acute{x} - \mu}{SE}$$

Where:

- $\acute{x}$= sample mean
- μ = population mean under H_0
- SE = standard error of the mean

The test statistic tells us how far our sample mean is from the hypothesised mean in terms of standard errors.

Step 4: Make a Decision

Using the computed test statistic and sampling distribution:

- If the **p-value < α**, reject H_0 (evidence against the null).
- If the **p-value ≥ α**, fail to reject H_0 (evidence is insufficient).

Example:

- If $p = 0.03$ at $\alpha = 0.05$ → reject H_0.
- If $p = 0.10$ at $\alpha = 0.05$ → fail to reject H_0.

Thus, there are two decisions possible:

1. Reject the null hypothesis.
2. Accept (fail to reject) the null hypothesis.

Hypothesis Testing and Sampling Distributions

- The **sample mean** is an unbiased estimator of the population mean.
- The **sampling distribution of the mean** is normally distributed (Central Limit Theorem).
- To locate the probability of observing a sample mean, we need:
    1. The hypothesised population mean.
    2. The standard error of the mean.

Because we observe only samples (not the entire population), there is always a chance of error in conclusions.

Errors in Hypothesis Testing

Decision	**H_0 is True**	**H_0 is False**
Accept H_0	Correct ($1 - \alpha$)	Type II Error (β)
Reject H_0	Type I Error (α)	Correct (Power = $1-\beta$)

- **Type I Error (α):** Rejecting H_0 when it is actually true (false positive).
- **Type II Error (β):** Failing to reject H_0 when it is actually false (false negative).
- **Power ($1-\beta$):** Probability of correctly rejecting a false H_0.

Analogy:

- **Type I error:** Convicting an innocent person.
- **Type II error:** Letting a guilty person go free.

The Decision to Reject H_0

- When rejecting H_0, we might be correct or incorrect.
- The probability of making a **Type I error** is set by our choice of α (significance level).
- This is why medical research uses smaller α values, as the consequences of a false positive are severe.

Reporting Significance Levels

- Rejecting H_0 at $\alpha = 0.05$ does **not** mean there is a 95% chance that H_0 is false.
- Instead, it means the probability of observing the result (or more extreme) if H_0 were true is < 5%.

Importance of Setting Hypotheses in Advance

- Hypotheses (H_0 and H_1) must be defined **before** inspecting the data.

- Switching to a one-tailed test after seeing the data inflates the risk of false positives (equivalent to a two-tailed test with α doubled).
- Doing so is misleading and can invalidate results.

Hypothesis Tests and "Proof"

- Hypothesis testing cannot **prove** or **disprove** a hypothesis.
- Instead, it assesses whether sample evidence is consistent or inconsistent with H_0.
- A **p-value = 0** is impossible in practice, so absolute proof is unattainable.

Related vs. Unrelated Samples

- **Related samples tests (paired samples)** are more powerful since individual differences cancel out.
- **Unrelated samples tests (independent samples)** may require larger sample sizes.
- However, repeated measurements in related samples may risk participant dropout, which must be considered in study design.

Conclusion

Hypothesis testing is a structured process involving:

1. Stating H_0 and H_1.
2. Choosing a significance level (α).
3. Computing a test statistic.
4. Making a decision based on p-values.

Errors are possible, but careful choice of significance level and study design reduces risks. Hypothesis testing does not prove or disprove but instead provides evidence for or against the null hypothesis.