Statistics

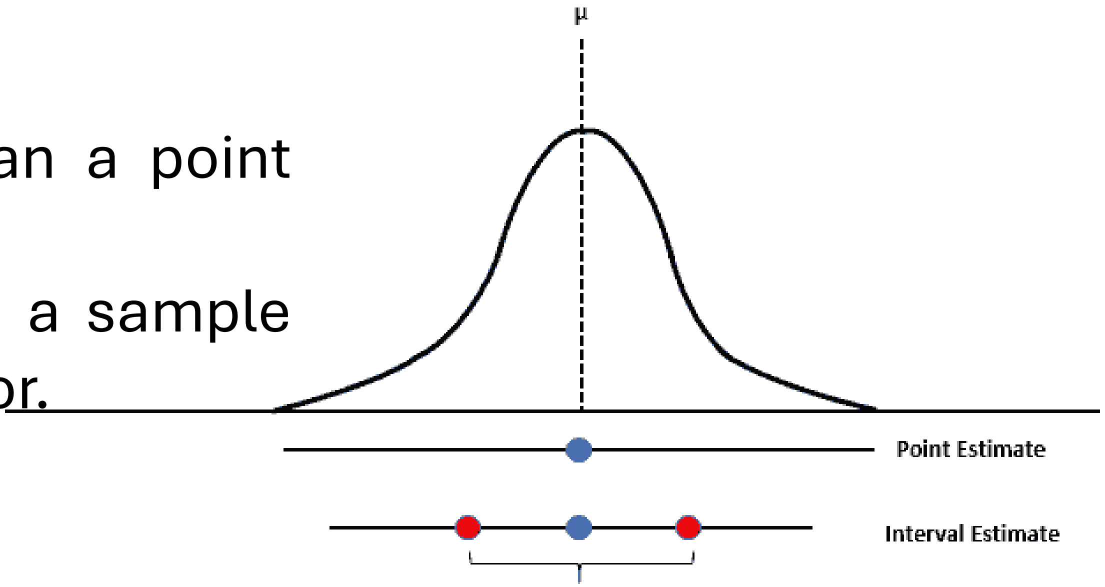
Point Estimate and Interval Estimate

- A **point estimate** is a single value estimate of a parameter. For instance, a sample mean is a point estimate of a population mean.
- A point estimate is a sample statistic calculated using the sample data to estimate the most likely value of the corresponding unknown population parameter. In other words, we derive the point estimate from a single value in the sample and use it to estimate the population value.
- Take a sample, find $\bar{x}$. It is a close approximation of μ. But, depending on your sample size, that may not be a good point estimate.
- In fact, the probability that a single sample statistic is equal to the population parameter is very unlikely.

Point Estimate and Interval Estimate

- An interval estimate gives you a range of values where the parameter is expected to lie.
- A confidence interval estimate is a range of values constructed from sample data so that the population parameter will likely occur within the range at a specified probability. Accordingly, the specified probability is the level of confidence.

- Broader and probably more accurate than a point estimate
- Any parameter estimate that is based on a sample statistic has some amount of sampling error.



Point Estimate and Interval Estimate

A Confidence interval is used to express the precision and ambiguity of a particular sampling method.

$$\mu = \overset{\text{Point Estimate}}{\bar{x}} \pm \overset{\text{Confidence Level}}{z_{\alpha/2}} * \overset{\text{Margin of Error}}{\frac{\sigma}{\sqrt{n}}}$$

- A confidence *interval* is a range of values that probably contain the population mean.
- A Confidence *level* is a percentage of certainty that, in any given sample, that confidence interval will contain the population means.
- The Point estimate is a statistic (value from a sample) used to estimate a parameter (value from the population).
- The margin of error is the maximum expected difference between the actual population parameter and a sample estimate of the parameter. In other words, it is the range of values above and below sample statistics.

Point Estimation

- Here, we assume that θ is an unknown parameter to be estimated.
- For example, θ might be the expected value of a random variable, θ=EX. Θ is a fixed (non-random) quantity.
- To estimate θ, we define a point estimator $\hat{\Theta}$ that is a function of the random sample, i.e.,

$$\hat{\Theta} = h(X_1, X_2, \cdots, X_n).$$

For example, if $\theta = EX$, we may choose $\hat{\Theta}$ to be the sample mean

$$\hat{\Theta} = \overline{X} = \frac{X_1 + X_2 + \ldots + X_n}{n}.$$

Point Estimation

Mean ($\bar{x}$) → Estimates Population Mean (μ)

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Variance (s^2) → Estimates Population Variance (σ^2)

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

Standard Deviation (s) → Estimates Population Standard Deviation (σ)

$$\hat{\sigma} = \sqrt{\hat{\sigma}^2}$$

Proportion (p-hat) → Estimates Population Proportion (p)

$$\hat{p} = \frac{x}{n}$$

Properties of Estimators

- Estimators should be **unbiased**.
  - expected value equals the true parameter value.
- The estimator should be **efficient.**
  - it has the **lowest variance** among all unbiased estimators of a parameter.
- An estimator should be **consistent.**
  - as the sample size increases, the estimated value gets closer to the true population parameter.
  - More data improves the accuracy of estimation.

Evaluating Estimators

- Three main desirable properties for point estimators

1. The **bias** of an estimator $\hat{\Theta}$ tells us on average how far $\hat{\Theta}$ is from the real value of θ.

i. Unbiasedness

- An estimator **θ^** is **unbiased** if its expected value is equal to the true population parameter (θ):

$$E(\hat{\theta}) = \theta$$

- Example: The sample mean $\bar{x}$ is an unbiased estimator of the population mean **μ**:

$$E(\bar{X}) = \mu$$

ii. *consistency*

- An estimator **θ̂** is **consistent** if it gets **closer to the true parameter (θ) as the sample size (n) increases**.

Let $\hat{\Theta}_1, \hat{\Theta}_2, \cdots, \hat{\Theta}_n, \cdots$, be a sequence of point estimators of θ. We say that $\hat{\Theta}_n$ is a **consistent** estimator of θ, if

$$\lim_{n\to\infty} P\big(|\hat{\Theta}_n - \theta| \geq \epsilon\big) = 0, \text{ for all } \epsilon > 0.$$

- Example: The **sample mean ($\bar{x}$)** is a consistent estimator of **μ** because as we take larger samples, it converges to **μ**.

iii. Efficiency

- Among multiple unbiased estimators, the **most efficient** estimator has the **smallest variance**.

$$Var(\hat{\theta}_1) < Var(\hat{\theta}_2) \Rightarrow \hat{\theta}_1 \text{ is more efficient}$$

Example: If we have two estimators of **μ**, the one with the smaller variance is preferred.

Estimator 1: $Var(\hat{\theta}_1) = 5$

Estimator 2: $Var(\hat{\theta}_2) = 2$

Hypothesis Testing

It refers to

- Making an assumption, called hypothesis, about a population parameter.
- Collecting sample data and calculating sample statistic.
- Using the sample statistic to evaluate the hypothesis (how likely is it that our hypothesized parameter is correct).
- To test the validity of our assumption we determine the difference between the hypothesized parameter value and the sample value.

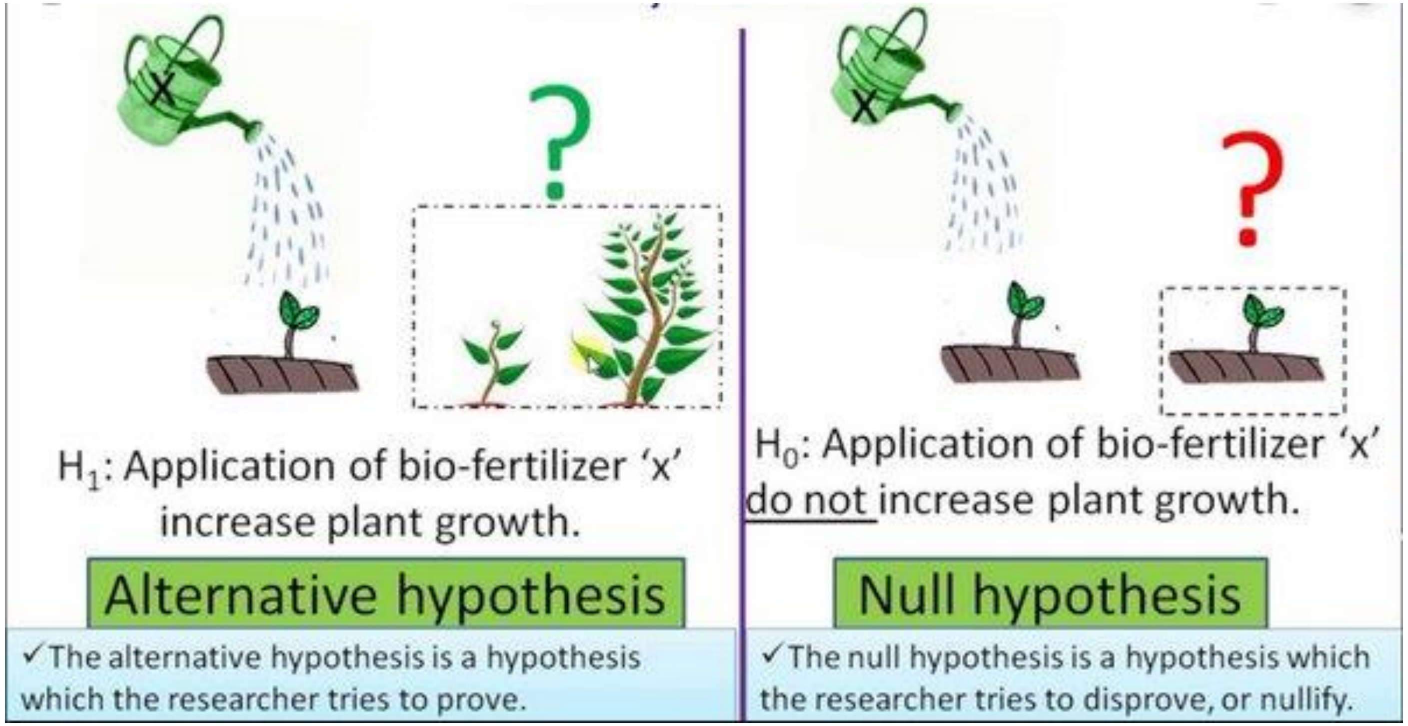
Hypothesis Testing

Example:

- A pharmaceutical company might be interested in knowing if a new drug is effective in treating a disease. Here, there are two hypotheses.
- The first one is that the drug is not effective (hypotheses H0), while the second hypothesis is that the drug is effective (hypotheses H1).
- The hypothesis H0 is called the **null hypothesis** and the hypothesis H1 is called the **alternative hypothesis**.

Hypothesis Testing

The null hypothesis represents the default assumption that no significant difference or relationship exists between the studied variables. In contrast, the alternative hypothesis represents the claim or hypothesis the researcher is testing.



Hypothesis Testing

You have a coin and you would like to check whether it is fair or not. More specifically, let θ be the probability of heads, $\theta = P(H)$. You have two hypotheses:

H_0 (the null hypothesis): The coin is fair, i.e. $\theta = \theta_0 = \frac{1}{2}$.

H_1 (the alternative hypothesis): The coin is not fair, i.e., $\theta \neq \frac{1}{2}$.

We need to design a test to either accept H_0 or H_1. To check whether the coin is fair or not, we perform the following experiment. We toss the coin 100 times and record the number of heads. Let X be the number of heads that we observe, so

$$X \sim Binomial(100, \theta).$$

Now, if H_0 is true, then $\theta = \theta_0 = \frac{1}{2}$, so we expect the number of heads to be close to 50. Thus, intuitively we can say that if we observe close to 50 heads we should accept H_0, otherwise we should reject it. More specifically, we suggest the following criteria: If $|X - 50|$ is less than or equal to some threshold, we accept H_0. On the other hand, if $|X - 50|$ is larger than the threshold we reject H_0 and accept H_1. Let's call that threshold t.

If $|X - 50| \leq t$, accept H_0.

If $|X - 50| > t$, accept H_1.

Hypothesis Testing

- **Level of significance**: It refers to the degree of significance in which we accept or reject the null hypothesis. 100% accuracy is not possible for accepting a hypothesis, so we select a level of significance. This is normally denoted with alpha(α) and generally, it is 0.05 or 5% which means your output should be 95% confident to give a similar kind of result in each sample.
- **Test Statistic:** Test statistic is the number that helps you decide whether your result is significant. It’s calculated from the sample data you collect it could be used to test if a machine learning model performs better than a random guess.
- **Critical value:** Critical value is a boundary or threshold that helps you decide if your test statistic is enough to reject the null hypothesis

Hypothesis Testing

- **P-value**: The p-value is the probability of observing a test statistic given that the null hypothesis is true.
  - A **small p-value** usually less than 0.05 means the results are unlikely to be due to random chance so we reject the null hypothesis.
  - A **large p-value** means the results could easily happen by chance so we don't reject the null hypothesis.

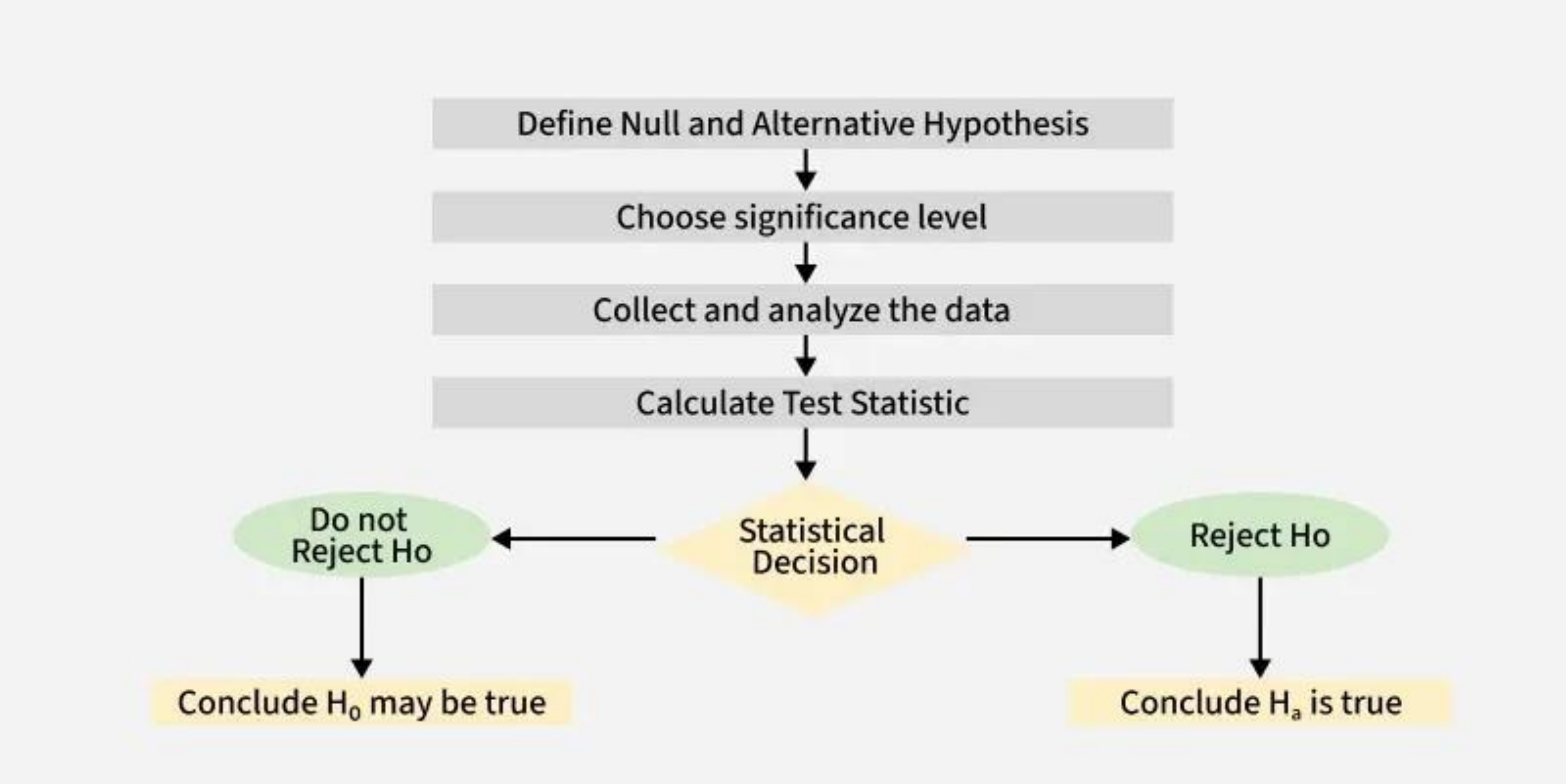
Hypothesis Testing

In hypothesis testing Type I and Type II errors are two possible errors that can happen when we are finding conclusions about a population based on a sample of data. These errors are associated with the decisions we made regarding the null hypothesis and the alternative hypothesis.

- **Type I error:** When we reject the null hypothesis although that hypothesis was true. Type I error is denoted by alpha(α).
- **Type II errors:** When we accept the null hypothesis, but it is false. Type II errors are denoted by beta(β).

	Null Hypothesis is True	**Null Hypothesis is False**
Null Hypothesis is True (Accept)	Correct Decision	Type II Error (False Negative)
Alternative Hypothesis is True (Reject)	Type I Error (False Positive)	Correct Decision

Hypothesis Testing



One-Sided (One-Tailed) Tests

- **Greater Than Test (Right-Tailed Test):** This type of one-sided test is used when you want to determine if a parameter or effect is greater than a specified value.
  - **Null Hypothesis (H0):** The parameter is less than or equal to the specified value.
  - **Alternative Hypothesis (H1 or Ha):** The parameter is greater than the specified value.
  - ***Example:*** Testing if a new drug improves patient recovery time **H0:** The drug does not improve recovery time. **Ha:** The drug improves recovery time.
- **Less Than Test (Left-Tailed Test):** This one-sided test is used when you want to determine if a parameter or effect is less than a specified value.
  - **Null Hypothesis (H0):** The parameter is greater than or equal to the specified value.
  - **Alternative Hypothesis (H1 or Ha):** The parameter is less than the specified value.
  - ***Example:*** Testing if a manufacturing process meets quality standards. **H0:** The process meets quality standards. **Ha:** The process does not meet quality standards.

Two-Sided (Two-Tailed) Tests

Two-Sided Test: This type of test is used when you want to determine if a parameter or effect is significantly different from a specified value, without specifying whether it's greater or less than that value.

- **Null Hypothesis (H0):** The parameter is equal to the specified value.
- **Alternative Hypothesis (H1 or Ha):** The parameter is not equal to the specified value.
- ***Example:*** Testing if a coin is fair (i.e., equally likely to land heads or tails). **H0:** The coin is fair. **Ha:** The coin is not fair.