Seed Germination

Purva Parmar
20181081

Contents

1 Aim

To study germination of seeds under the influence of a designed experiment and measure the effects on root and shoot lengths.

2 Theory

2.1 Germination of a Seed

Seed Germination is the process by which a plant species grows from a seed to a plant.

In the initial stages, the seeds take up water rapidly, resulting in swelling and softening of the seed coat at

an optimum temperature.

The seed coat then ruptures, and a primary root is formed. The seed then starts to respire and produce proteins and metabolize the stored food.

In the final stage, the cells of the seed are elongated and divided. The cotyledons expand to form the leaves.

Seeds need water, oxygen, an optimum temperature and light to germinate properly. Nutrients are later on needed in the soil as well.

2.1.1 Factors affecting Seed Germination

- External Factors

  - Water

  - Temperature

  - Oxygen

- Internal Factors

  - Seed Dormancy

    Under certain conditions, the seeds can be prevented from germinating even under favourable conditions.

    * The seed coat is resistant to water and oxygen exchange.
    * Seeds with undeveloped or immature embryo do not germinate.
    * Some seeds contain plant growth regulators which inhibit seed germination.
    * Some seeds just simply require a large amount of time for germination.

Cotton Wool can be used for seed germination. The soft and tender fibres, combined with water acts as a good substitute for soil.

2.2 Statistics

Variations are inherent in biology. We need mathematical methods to study and analyze experimental data, and to sift out the variation due to experimental manipulation from inherent variation.

There are some measures for it.

2.2.1 Measures of central tendency

- **Mean**

  Mean of a data is the sum of all datapoints divided by the number of datapoints.

  So, if the values are $x_1, x_2, \ldots, x_n$, then the mean $\overline{x}$ is:

  $$\overline{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{\sum\limits_{i=1}^{n} x_i}{n}$$

  We represent mean of the sample by $\overline{x}$ and the mean of the population by μ.

  Mean is a good measure of central tendency, however, it can be affected by outliers. A few very large or very small values can significantly affect the mean.

- **Median**

  The median of a dataset is the value that divides the dataset into two.

  Half of the values in the data set lie above the median, and the other half lie below it.

- **Mode**

  The mode of a dataset is the most frequent value.

2.2.2 Variance

There are some measures of the spread of a data.

- **Variance**

  The variance of a sample is denoted by S^2, and that of a population by σ^2. It is defined by:

  $$S^2 = \frac{\sum\limits_{i=1}^{n}(x_i - \overline{x})^2}{n}$$

  $$\sigma^2 = \frac{\sum\limits_{i=1}^{n}(x_i - \mu)^2}{n}$$

- **Standard Deviation**

  The standard deviation is the most commonly used measure of spread or measure of variation. It measures spread around the mean.

  It is also a good indicator of outliers. It becomes useful in comparing two datasets with similar means.

  The standard deviation of a sample is denoted by S and that of a population by σ. It is defined by:

  $$S = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n}}$$

  $$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n}}$$

  – About 68% of the data lies in: $\overline{x} \pm S$
  – About 95% of the data lies in: $\overline{x} \pm 2S$
  – About 99% of the data lies in: $\overline{x} \pm 3S$

The Notation:

	Sample	Population
Mean	$\overline{x}$	μ
Variance	S^2	σ^2
Std Dev	S	σ

Standard Error of Mean

The Standard Error of Mean (SEM) is the expected value of standard deviation of means of several samples.

It can be estimated from a single sample as:

$$\text{SEM} = \frac{S}{\sqrt{n}}$$

where S is the Standard Deviation of the sample, and n is the sample size.

The Population Mean then can be represented as:

$$\mu = \overline{x} \pm \text{SEM}$$

2.3 Normal Distribution

A Normal Distribution is also called a Gaussian Distribution. It is a probability distribution that is symmetric about the mean, showing that the data near the mean are more frequent in occurence than the data far from the mean.

It appears as a bell curve in graph form.

The Normal Distribution is motivated by the Central Limit Theorem, which states that averages calculated from independent, identically distrubuted random variables have approximately normal distributions, regardless of the type of distribution from which the variables are sampled, provided it has a finite variance.

$$f(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

where
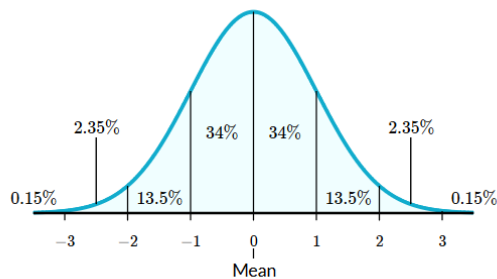$\mu = $ Mean
$\sigma^2 = $ Variance

Figure 1: Normal Distribution [Source: Khan Academy]

2.4 Hypothesis Testing

Hypothesis Testing is the use of statistics to determine the probability that a given hypothesis is true.

The best way to determine whether a statistical hypothesis is true would be to examine the entire population.

However, this is often impractical. We study random samples from the population and if the sample data is not consistent with the statistical hypothesis, the hypothesis is rejected.

There are two types of statistical hypothesese:

- **Null Hypothesis (H$_\circ$)**: Sample observations result purely from chance. In other words, the population mean of test (μ_T) is the same as population mean of control (μ_C).

  H$_\circ$: $\mu_\mathsf{C} = \mu_\mathsf{t}$

- **Alternative Hypothesis (H$_\mathsf{A}$)**: Observations are influenced by some non-random cause. In other words, the population mean from test is significantly different compared to the population mean from control.

  H$_\mathsf{A}$: $\mu_\mathsf{C} \neq \mu_\mathsf{T}$

2.4.1 Testing Procedure

1. Formulate the **Null Hypothesis** $H_\circ$ and **Alternative Hypothesis** H_A.

2. Identify a **test statistic** that can be used to assess the truth of the null hypothesis.

3. Compute the **P-value**, which is the Probability that a test statistic at least as significant as the one observed would be obtained assuming the Null Hypothesis were true.

   The smaller the P-value, the stronger the evidence against the Null Hypothesis.

4. Compare the P-value to an acceptable significance value α (called the **alpha value**).

   If $p \leqslant \alpha$, the observed effect is statistically significant. The Null Hypothesis is ruled out and the Alternative Hypothesis is valid.

   Usually, $\alpha = 0.05$.

2.4.2 Decision Errors

- **Type I Error**: Rejecting a Null Hypothesis that is actually true

  The probability of committing a Type I error is called the *significance level or* alpha, denoted by α.

- **Type II Error**: Failure to reject a Null Hypothesis that is actually false.

  The probability of committing a Type II Error is called *Beta, denoted by* β. *The probability of* not committing a Type II Error is called the *Power of the test.*

In a nutshell,

	Accept H$_\circ$	**Reject H$_\circ$**
H$_\circ$ Correct	✓	Type I Error
H$_\circ$ Wrong	Type II Error	✓

4

2.4.3 Paired t-test

The Paired t-test is used to compare means of two samples.

$$t = \frac{\overline{x_C} - \overline{x_T}}{\sqrt{\frac{S_C^2}{N_C} + \frac{S_T^2}{N_T}}}$$

The numerator takes the difference between means of samples from control and test. The larger this value is, the more is the effect of the test.

The denominator is the sum of the standard errors.

Overall, the t-value is a normalized value which shows how significantly different is the mean of control from the mean of test.

Larger the t-value, the more significant is the result.

Smaller t-values indicate overlap in samples. That is, the samples may be from the same population.

2.4.4 ANOVA

ANOVA stands for **Analysis of Variance**. It is a collection of statistical models and related estimation procedures to analyze the differences among group means in a sample.

If C is the control, and $T_1, T_2, \ldots, T_m$ are the tests, then for ANOVA:

$$H_0 = \mu_C = \mu_{T_1} = \mu_{T_2} = \cdots = \mu_{T_m}$$
$$H_A = \text{At least one mean is different}$$

To check if the means are equal, we do the **F-test**. Roughly speaking, it is:

$$F = \frac{\text{Variance between-groups}}{\text{Variance within-groups}}$$

3 Experiment

We took 11 Petri dishes, spread cotton on them, and added water.

We added 20 Moong seeds to each of them.

The 11 Plates were as follows:

1. `Control` - Just plain cotton and water

2. `GA` - Gibberellic Acid

3. `ABA` - Abscisic Acid

4. `GA + ABA` - Gibberellic Acid + Abscisic Acid

5. `GA (high conc.)` - Gibberellic Acid at a higher concentration

6. `M Control` - Control, but with a small Neodymium magnet below

7. `M1` - A large magnet below

8. `M2` - Another large magnet (but smaller than M1) below

9. `GA + M` - Gibberellic Acid and a Neodymium magnet below

10. `ABA + M` - Abscisic Acid with a Neodymium magnet below

11. `GA + ABA + M` - Gibberellic Acid, Abscisic Acid and a Neodymium magnet, all of them.

Additional Data:

- GA Concentration $= 1 \ \mu g/mL$
- Strength of M ≈ 0.3 T

- Strength of M1 ≈ 0.08 T

- Strength of M2 ≈ 0.04 T

These plates were left for a week in a room. Light was naturally as well as artificially available. Water was added periodically.

After a week, most of the seeds had germinated.

4 References

- Seed Germination - Byjus

- Normal Distribution - Investopedia

- Hypothesis Testing - Stat Trek

- Hypothesis Testing - Wolfram MathWorld

- Understanding t-tests - The Minitab Blog