Two samples t-test

Jean Peyen

Queen's University of Belfast

August 30, 2024

Recap: One sample t-test

- **Sample:** $x_1, x_2, \cdots, x_n$, continuous valued, with unknown variance, either normally distributed or n is large.

Recap: One sample t-test

- **Sample:** $x_1, x_2, \cdots, x_n$, continuous valued, with unknown variance, either normally distributed or n is large.
- **Null hypothesis $\mathbf{H_0}$:** The true expectation (mean) μ of the random distribution from which the x_i values are drawn is equal to a specified value μ_0, i.e.,

$$H_0 : \mu = \mu_0.$$

Recap: One sample t-test

- **Sample:** $x_1, x_2, \cdots, x_n$, continuous valued, with unknown variance, either normally distributed or n is large.
- **Null hypothesis $\mathbf{H_0}$:** The true expectation (mean) μ of the random distribution from which the x_i values are drawn is equal to a specified value μ_0, i.e.,

$$H_0 : \mu = \mu_0.$$

- **Test statistic:**

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}},$$

Recap: One sample t-test

- **Sample:** $x_1, x_2, \cdots, x_n$, continuous valued, with unknown variance, either normally distributed or n is large.
- **Null hypothesis $\mathbf{H_0}$:** The true expectation (mean) μ of the random distribution from which the x_i values are drawn is equal to a specified value μ_0, i.e.,

$$H_0 : \mu = \mu_0.$$

- **Test statistic:**

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}},$$

- **Sampling distribution:** Under H_0, the test statistic follows a Student's t-distribution with $n-1$ degrees of freedom.

Activity. *We want to test whether the amount of fat in an egg yolk meets the standard, which is* $300g/kg$. *Measurements have been performed on a sample of 10 eggs, and the following results have been obtained: the sample mean* $\bar{x} = 301.5\ g/kg$ *and the sample variance* $s^2 = 10.00\ (g/kg)^2$. *Can we conclude that the fat content in the egg yolks meets the standard at a significance level of 0.05?*

Solution. In this case we have to perform a two-sided t-test with the null assumption

$$H_0 : \mu = 300.$$

Test statistic $t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{301.5 - 300}{\frac{\sqrt{10.00}}{\sqrt{10}}} = 1.5$, **Degrees of freedom** $df = n - 1 = 9$

Critical value $t_{0.025,9} \approx \pm 2.262$

Conclusion $|t| = 1.5 < 2.262$. The mean fat content does not differ significantly from the standard.

Two samples t-test (Welch's)

- **Samples:** Two independent samples, $x_{11}, x_{21}, \ldots, x_{n_a 1}$ and $x_{12}, x_{22}, \ldots, x_{n_b 2}$, continuous valued, with unknown and unequal variances, either normally distributed or n_a, n_b are large.

Two samples t-test (Welch's)

- **Samples:** Two independent samples, $x_{11}, x_{21}, \ldots, x_{n_a 1}$ and $x_{12}, x_{22}, \ldots, x_{n_b 2}$, continuous valued, with unknown and unequal variances, either normally distributed or n_a, n_b are large.
- **Null hypothesis $\mathbf{H_0}$:** The true expectations (means) μ_1 and μ_2 of the random distributions from which the x_{i1} and x_{i2} values are drawn are equal, i.e.,

$$H_0 : \mu_1 = \mu_2.$$

Two samples t-test (Welch's)

- **Samples:** Two independent samples, $x_{11}, x_{21}, \ldots, x_{n_a 1}$ and $x_{12}, x_{22}, \ldots, x_{n_b 2}$, continuous valued, with unknown and unequal variances, either normally distributed or n_a, n_b are large.
- **Null hypothesis $\mathbf{H_0}$:** The true expectations (means) μ_1 and μ_2 of the random distributions from which the x_{i1} and x_{i2} values are drawn are equal, i.e.,

$$H_0 : \mu_1 = \mu_2.$$

- **Test statistic:**

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_a} + \frac{s_2^2}{n_b}}},$$

Two samples t-test (Welch's)

- **Samples:** Two independent samples, $x_{11}, x_{21}, \ldots, x_{n_a 1}$ and $x_{12}, x_{22}, \ldots, x_{n_b 2}$, continuous valued, with unknown and unequal variances, either normally distributed or n_a, n_b are large.
- **Null hypothesis $\mathbf{H_0}$:** The true expectations (means) μ_1 and μ_2 of the random distributions from which the x_{i1} and x_{i2} values are drawn are equal, i.e.,

$$H_0 : \mu_1 = \mu_2.$$

- **Test statistic:**

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_a} + \frac{s_2^2}{n_b}}},$$

- **Sampling distribution:** Under H_0, the test statistic follows a Student's t-distribution with Welch-Satterthwaite degrees of freedom:

$$\text{df} \approx \frac{\left(\frac{s_1^2}{n_a} + \frac{s_2^2}{n_b}\right)^2}{\frac{\left(\frac{s_1^2}{n_a}\right)^2}{n_a - 1} + \frac{\left(\frac{s_2^2}{n_b}\right)^2}{n_b - 1}}.$$

Activity. *A competitor claims that their steel cheaper and stronger than the steel of our supplier. The goal is to test whether there is a significant difference in the tensile strength between the competitor's steel and our supplier's steel. Tensile testing has been performed on two sets of steel samples: 10 samples from Supplier A and 10 samples from Competitor B. The results obtained are as follows:*

$$\text{Supplier A's Steel: } \{123, 116, 110, 119, 112, 127, 118, 105, 120, 130\} \text{ kpsi}$$
$$\text{Competitor B's Steel: } \{97, 118, 105, 122, 89, 108, 113, 124, 101, 110\} \text{ kpsi}$$

Can it be concluded at a significance level of 0.05 that Supplier A's steel is stronger than Competitor B's?

Solution.

1. Hypotheses:

Null Hypothesis (H_0): The mean tensile strength of Supplier A's steel is not better than the mean tensile strength of Competitor B's steel. $H_0 : \mu_A = \mu_B$

Alternative Hypothesis (H_1): The mean tensile strength of Supplier A's steel is greater than the mean tensile strength of Competitor B's steel. $H_1 : \mu_A > \mu_B$

Solution.

1. Hypotheses:

Null Hypothesis (H_0): The mean tensile strength of Supplier A's steel is not better than the mean tensile strength of Competitor B's steel. $H_0 : \mu_A = \mu_B$

Alternative Hypothesis (H_1): The mean tensile strength of Supplier A's steel is greater than the mean tensile strength of Competitor B's steel. $H_1 : \mu_A > \mu_B$

2. Check Normality:

Use the following code to check if the tensile strength data follows a normal distribution:

```
stats.probplot(data, dist="norm", plot=plt)
```

Scipy.stats probplot Documentation

Solution.

1. Hypotheses:

Null Hypothesis (H_0): The mean tensile strength of Supplier A's steel is not better than the mean tensile strength of Competitor B's steel. $H_0 : \mu_A = \mu_B$

Alternative Hypothesis (H_1): The mean tensile strength of Supplier A's steel is greater than the mean tensile strength of Competitor B's steel. $H_1 : \mu_A > \mu_B$

2. Check Normality:

Use the following code to check if the tensile strength data follows a normal distribution:

```
stats.probplot(data, dist="norm", plot=plt)
```

Scipy.stats probplot Documentation

3. Perform the Welch's t-test:

Use Welch's t-test to compare the tensile strengths with the following code:

```
t, p = stats.ttest_ind(A, B, equal_var=False, alternative='greater')
```

Scipy.stats ttest_ind Documentation

Solution.

1. Hypotheses:

Null Hypothesis (H_0): The mean tensile strength of Supplier A's steel is not better than the mean tensile strength of Competitor B's steel. $H_0 : \mu_A = \mu_B$

Alternative Hypothesis (H_1): The mean tensile strength of Supplier A's steel is greater than the mean tensile strength of Competitor B's steel. $H_1 : \mu_A > \mu_B$

2. Check Normality:

Use the following code to check if the tensile strength data follows a normal distribution:

```
stats.probplot(data, dist="norm", plot=plt)
```

Scipy.stats probplot Documentation

3. Perform the Welch's t-test:

Use Welch's t-test to compare the tensile strengths with the following code:

```
t, p = stats.ttest_ind(A, B, equal_var=False, alternative='greater')
```

Scipy.stats ttest_ind Documentation

4. Conclusion:

Compare the p-value obtained from the t-test with the significance level ($\alpha = 0.05$). If $p < 0.05$, reject the null hypothesis and conclude that Supplier A's steel is significantly stronger than Competitor B's steel. If $p \geq 0.05$, do not reject the null hypothesis.

```
import scipy.stats as stats
import matplotlib.pyplot as plt

# Data
A = [123, 116, 110, 119, 112, 127, 118, 105, 120, 130]
B = [97, 118, 105, 122, 89, 108, 113, 124, 101, 110]

# QQ plot to check normality
stats.probplot(A, dist="norm", plot=plt)
plt.show()
stats.probplot(B, dist="norm", plot=plt)
plt.show()

# Perform Welch's t-test (two-sample t-test with unequal variances)
t, p = stats.ttest_ind(A, B, equal_var=False, alternative='greater')

# Output the results
print(f"T-statistic: {t:.4f}")
print(f"P-value: {p:.4f}")
```

Homework. *Guinness Brewery is testing whether a new brewing process results in a different alcohol concentration in their Guinness Draught compared to the traditional process. Two separate batches of beer are brewed, one using the traditional method and the other using the new method. Samples of beer from each batch are measured for their alcohol concentration (ABV).*
The data collected are as follows:

- *Traditional Method: 4.2%, 4.3%, 4.1%, 4.4%, 4.2%, 4.3%, 4.1%, 4.2%, 4.3%, 4.2%*
- *New Method: 4.5%, 4.6%, 4.5%, 4.7%, 4.6%, 4.5%, 4.6%, 4.7%, 4.5%, 4.6%*

These data are assumed to be normally distributed. Follow the steps below to perform a Welch's t-test to determine if there is a statistically significant difference in the mean alcohol concentration between the two brewing methods. Assume a significance level of 0.05.

1. State the null and alternative hypotheses. [1 mark]
2. Calculate the sample mean and standard deviation for each group. [2 marks]
3. Calculate the test statistic using Welch's t-test formula. [2 marks]
4. Determine the degrees of freedom using the Welch-Satterthwaite formula. [2 marks]
5. Find the critical value and make a conclusion. [3 marks]

Next session: What if we have more than two groups?

Question: We have k samples and want to check if any of these samples is statistically different from the others.

- **Option 1:** Perform multiple pairwise tests with Bonferroni correction.

Next session: What if we have more than two groups?

Question: We have k samples and want to check if any of these samples is statistically different from the others.

- **Option 1:** Perform multiple pairwise tests with Bonferroni correction.
- **Option 2:** Analysis of Variance (ANOVA)
  - **Between-group variability (MSB):**

$$\text{MSB} = \frac{\sum_{i=1}^{k} n_i (\bar{x}_i - \bar{x})^2}{k-1}$$

  - **Within-group variability (MSW):**

$$\text{MSW} = \frac{\sum_{i=1}^{k} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2}{N-k}$$

  - **Test statistic:**

$$F = \frac{\text{MSB}}{\text{MSW}}$$

  - **Sampling distribution:** Under H_0, F follows an F-distribution with $(k-1, N-k)$ degrees of freedom.

Sources

- Douglas C. Montgomery, George C. Runger, and Norma F. Hubele. *Engineering Statistics*, 5th ed. John Wiley & Sons, 2010, Chapter 5.
- James T. McClave, P. George Benson, and Terry Sincich. *Statistics for Business and Economics*, 13th ed. Pearson, 2018, Chapter 8.
- *Python For Beginners*. `https://www.python.org/about/gettingstarted/`
- "scipy.stats.ttest_ind". *SciPy Manual, two samples t-test and QQ plot.* `https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.ttest_ind.html` `https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.probplot.html`
- The lecture content is available at `https://etamunu.github.io/`