Member-only story

One-Sample Hypothesis Tests, with Python

The Complete Beginner's Guide to perform One-Sample Hypothesis Tests (with code!)

Chao De-Yu · Follow
Published in Level Up Coding
5 min read · May 20, 2021

Listen Share More

Photo by Luca Bravo on Unsplash

A Hypothesis Test is a statistical test that is used to test the assumption or hypothesis made and draw a conclusion about the entire population. In this article, I will share how to run one-sample hypothesis testing on a single population with different scenarios.

The FIVE STEPS begin here:

1. Define the Null Hypothesis (H_0)
   The null hypothesis is the starting line of a hypothesis test or initial claim what someone believes he has evidence for.
   Always uses **parameters** such as **μ** or **σ** or **π**, depending on types of test.
   A null hypothesis always contains an equal sign: =, ≤, or ≥.

2. Define the Alternative Hypothesis (H_1)
   The **opposing statement to H_0** and is a suspicious claim or finding what someone wants to prove.
   Always uses **parameters** such as **μ** or **σ** or **π**, depending on types of test.
   H_1 **never** has an equal sign in its statement, it takes on strict inequalities: ≠ or < or >.

3. Set the Level of Significance (α)
   Is the probability of committing Type I error, where
   α = P (Type I error) = P (rejecting H_0 when H_0 is true)
   Preset and usually selected as 0.01, 0.05, 0.1.

4. Collect data and calculate the Test Statistic

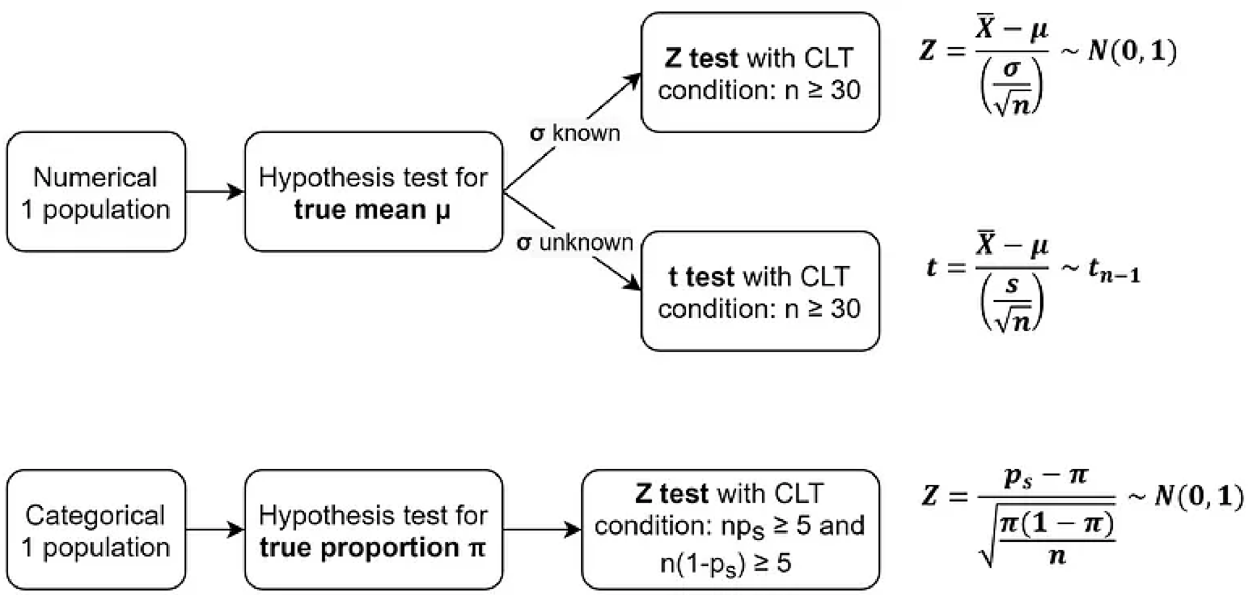

Figure 1: Test statistics depend on different situation

5. Construct the Rejection and Non-Rejection Regions

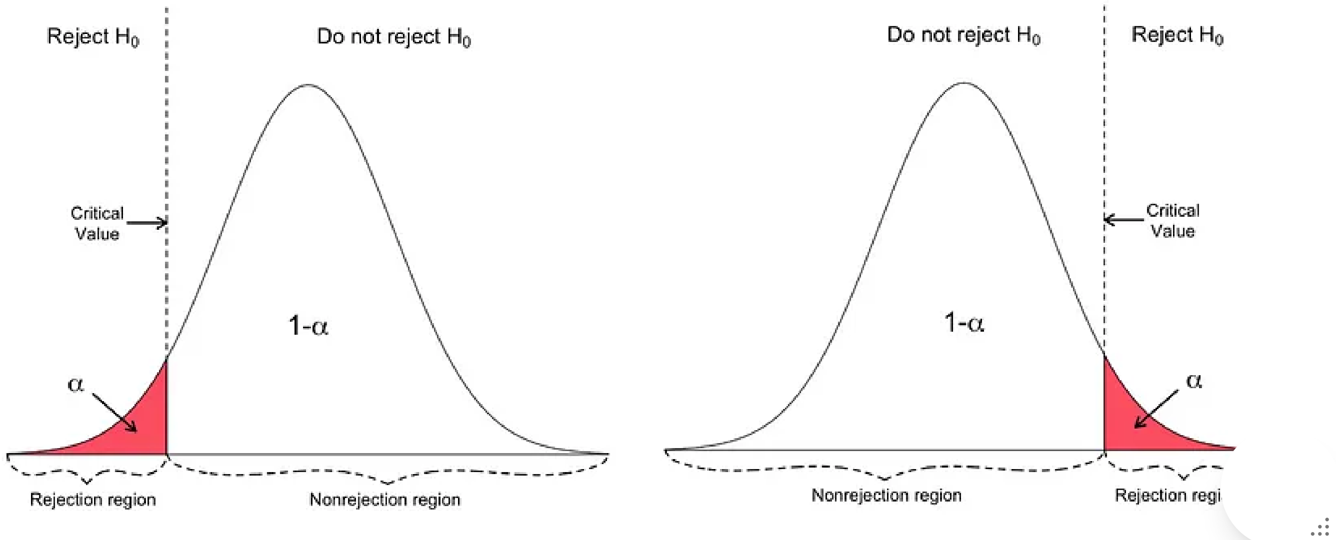

Figure 2: One-tailed test — left-tailed and right-tailed [1]

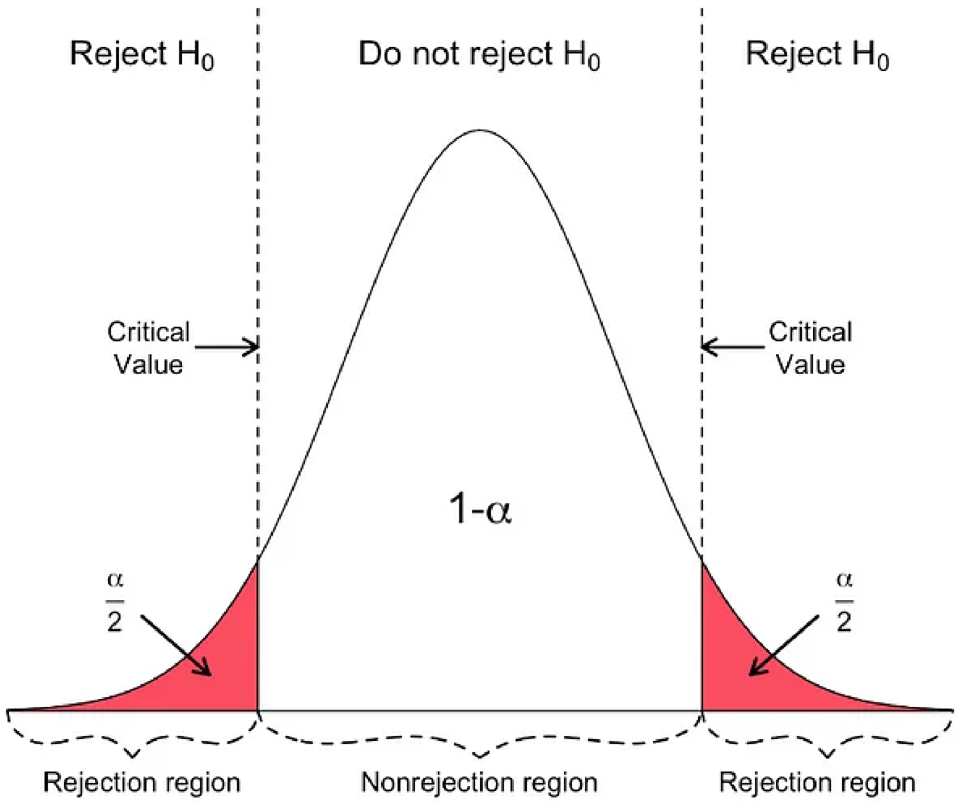

Figure 3: Two-tailed test [1]

The decision rule or the critical value of a test statistic divides the entire area under the curve (Z curve or t curve) into two main regions: the tiny **rejection region** and the monstrous **non-rejection region**.

Finally, based on test statistics and the decision rule, we make the decision to determine whether to reject or not to reject H_0 at α significance level.

Going through TEST STATISTICS with examples:

Example 1: Carl Reinhold August Wunderlich was a doctor who pioneered the measurement of average healthy human body temperature which he found to be 98.6°F. Almost 150 years later, researchers at the University of Maryland evaluate the famous 98.6°F standard. They measured the body temperatures of **148** healthy men and women, who yielded an average of **98.24923°F**. With the traditionally **known standard deviation** of human body temperatures at about **0.63°F**, is there evidence at **1% level of significance** that the true average human body temperature **differs from 98.6°F**? [2]

Let **μ** be the true mean human body temperature in °F
Given: n = 148, $\bar{x}$ = 98.24923°F, σ = 0.63°F
Now let's follow the five steps:
1. H_0: μ = 98.6°F
2. H_1: μ ≠ 98.6°F (2 tailed test)
3. α = 0.01
4. Following Figure 1, it is a one-sample two-tailed **Z test** for true mean **μ** with known variance **σ**.

```
import math
import scipy.stats as stats

x_bar = 98.24923
u = 98.6
sigma = 0.63
n = 148
alpha = 0.01

# possible types "right-tailed, left-tailed, two-tailed"
tail_hypothesis_type = "two-tailed"

if tail_hypothesis_type == "two-tailed":
    alpha /= 2

print("One-Sample", tail_hypothesis_type, "Z-test of true mean")
print("-------------------------------------------------------------------------------------")
# The p-value approach
print("Approach 1: The p-value approach to hypothesis testing in the decision rule")
if n >= 30:
    print("Since sample size >= 30, by CLT ")
    z_score = (x_bar - u)/(sigma/math.sqrt(n))
    if tail_hypothesis_type == "left-tailed":
        p_value = stats.norm.cdf(z_score)
    elif tail_hypothesis_type == "right-tailed":
        p_value = 1 - stats.norm.cdf(z_score) # or stats.norm.sf(z_score)
    else:
        z_score = abs(z_score)
        p_value = 1 - stats.norm.cdf(z_score)

    conclusion = "Failed to reject the null hypothesis."
    if p_value <= alpha:
        conclusion = "Null Hypothesis is rejected."

    print("z-score is:", z_score, " and p value is:", p_value)
    print(conclusion)
else:
    print("CLT is not satisfied")

# The critical value approach
print("\n---------------------------------------------------------------------------------------
print("Approach 2: The critical value approach to hypothesis testing in the decision rule")
if n >= 30:
    print("Since sample size >= 30, by CLT ")
    z_score = (x_bar - u)/(sigma/math.sqrt(n))
    critical_value = stats.norm.ppf(abs(alpha))

    conclusion = "Failed to reject the null hypothesis."
```

```
        conclusion = "Failed to reject the null hypothesis."
        if tail_hypothesis_type == "left-tailed":
            if z_score < critical_value:
                conclusion = "Null Hypothesis is rejected."
        elif tail_hypothesis_type == "right-tailed":
            critical_value = abs(critical_value)
            if z_score > critical_value:
                conclusion = "Null Hypothesis is rejected."
        else:
            z_score = abs(z_score)
            critical_value = abs(critical_value)
            if z_score > critical_value:
                conclusion = "Null Hypothesis is rejected."

        print("z-score is:", z_score, " and critical value is:", critical_value)
        print(conclusion)
    else:
        print("CLT is not satisfied")
```

```
One-Sample two-tailed Z-test of true mean
--------------------------------------------------------------------------------------
Approach 1: The p-value approach to hypothesis testing in the decision rule
Since sample size >= 30, by CLT
z-score is: 6.7734940087386954  and p value is: 6.285416631612861e-12
Null Hypothesis is rejected.

--------------------------------------------------------------------------------------
Approach 2: The critical value approach to hypothesis testing in the decision rule
Since sample size >= 30, by CLT
z-score is: 6.7734940087386954  and critical value is: 2.575829303548901
Null Hypothesis is rejected.
```

Conclusion: We have enough evidence that the true mean body temperature is NOT 98.6°F, at 1% significance level.

Example 2: In 2010, researchers use the Length Ratio to study individual facial attractiveness. They have done 4 different experiments and for simplicity, I will use Experiment 1 to do the test. [3]

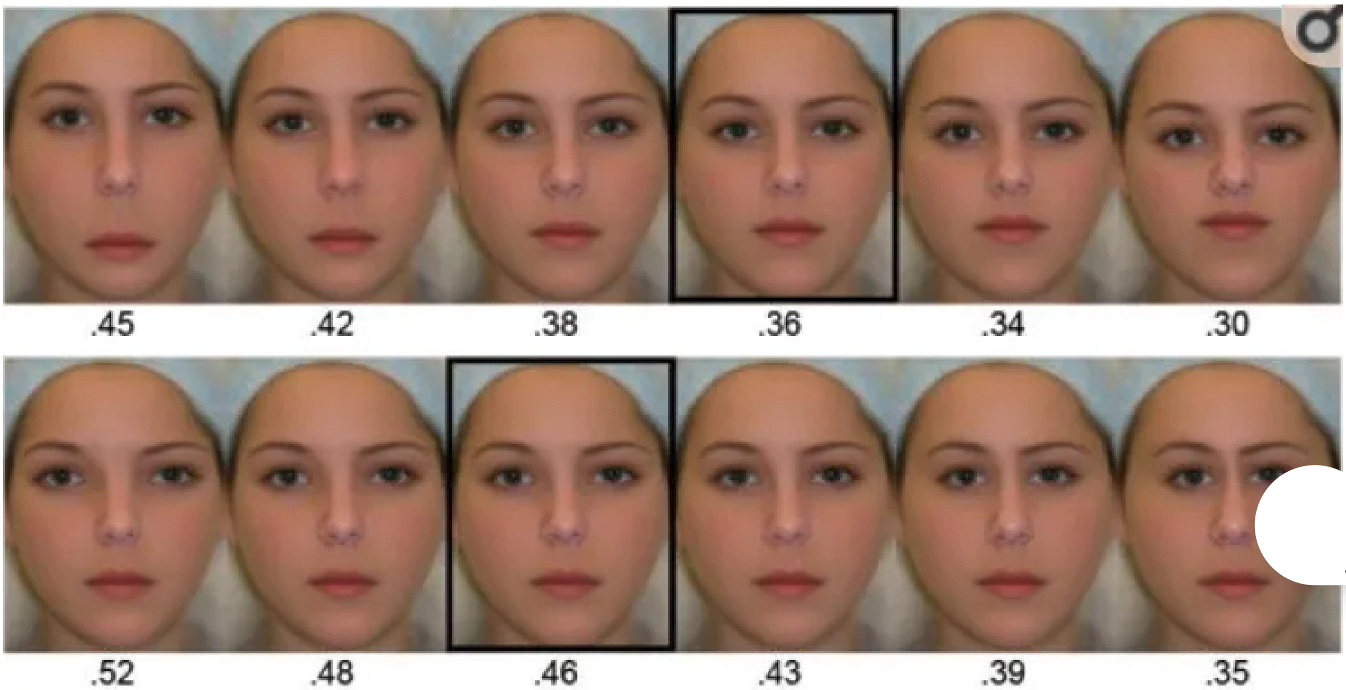

Figure 4: Experiment 1 [3]

Experiment 1: Researchers modified the length ratio of a single human face to obtain a series of faces as shown in Figure 4. The modified faces and the original face were then paired with each other to create face pairs with identical facial features but different length ratios. Participants were shown random face pairs to judge which face appeared more attractive. From **110** results, the **sample mean length ratio** for the most attractive face was **0.36**, with a **standard deviation** of **0.017.**

From the experiment, is there evidence that the true mean length ratio for face attractiveness is smaller than the classical golden ratio of 0.38 at 2.5% significance level?

Let **μ** be the true mean length ratio for face attractiveness
Given: $n = 110$, $\bar{x} = 0.36$, $s = 0.017$
Now let's follow the five steps:
1. H_0: $\mu \geq 0.38$
2. H_1: $\mu < 0.38$ (left-tailed test)
3. $\alpha = 0.025$
4. Following Figure 1, it is a one-sample left-tailed **t-test** for true mean **μ** with unknown variance **σ**.

```python
import math
import scipy.stats as stats

x_bar = 0.36
u = 0.38
std = 0.017
n = 110
alpha = 0.025

# possible types "right-tailed, left-tailed, two-tailed"
tail_hypothesis_type = "left-tailed"

if tail_hypothesis_type == "two-tailed":
    alpha /= 2

print("One-Sample", tail_hypothesis_type, "t-test of true mean")
print("--------------------------------------------------------------------------------------")
# The p-value approach
print("Approach 1: The p-value approach to hypothesis testing in the decision rule")
if n >= 30:
    print("Since sample size >= 30, by CLT ")
    t_score = (x_bar - u)/(std/math.sqrt(n))

    if tail_hypothesis_type == "left-tailed":
        p_value = stats.t.cdf(t_score, n-1)
    elif tail_hypothesis_type == "right-tailed":
        p_value = 1 - stats.t.cdf(t_score, n-1) # or stats.norm.sf(t_score, n-1)
    else:
        t_score = abs(t_score)
        p_value = 1 - stats.t.cdf(t_score, n-1)

    conclusion = "Failed to reject the null hypothesis."
    if p_value <= alpha:
        conclusion = "Null Hypothesis is rejected."

    print("t-score is:", t_score, " and p value is:", p_value)
    print(conclusion)
else:
    print("CLT is not satisfied")

# The critical value approach
print("\n------------------------------------------------------------------------------------
print("Approach 2: The critical value approach to hypothesis testing in the decision rule")
if n >= 30:
    print("Since sample size >= 30, by CLT ")
    t_score = (x_bar - u)/(std/math.sqrt(n))
    critical_value = stats.t.ppf(abs(alpha), n-1)
```

```
    conclusion = "Failed to reject the null hypothesis."
    if tail_hypothesis_type == "left-tailed":
        if t_score < critical_value:
            conclusion = "Null Hypothesis is rejected."
    elif tail_hypothesis_type == "right-tailed":
        critical_value = abs(critical_value)
        if t_score > critical_value:
            conclusion = "Null Hypothesis is rejected."
    else:
        t_score = abs(t_score)
        critical_value = abs(critical_value)
        if t_score > critical_value:
            conclusion = "Null Hypothesis is rejected."

    print("t-score is:", t_score, " and critical value is:", critical_value)
    print(conclusion)
else:
    print("CLT is not satisfied")
```

```
One-Sample left-tailed t-test of true mean
--------------------------------------------------------------------------------
Approach 1: The p-value approach to hypothesis testing in the decision rule
Since sample size >= 30, by CLT
t-score is: -12.338927625531204  and p value is: 1.0137870701492986e-22
Null Hypothesis is rejected.

--------------------------------------------------------------------------------
Approach 2: The critical value approach to hypothesis testing in the decision rule
Since sample size >= 30, by CLT
t-score is: -12.338927625531204  and critical value is: -1.9819674896884745
Null Hypothesis is rejected.
```

Conclusion: We have sufficient evidence that the true mean length ratio is smaller than 0.38, at 2.5% significance level.

Example 3: Suppose the CEO claims that *at least* 80 percent of the company's 1,000,000 customers are very satisfied. Again, **100** customers are surveyed using simple random sampling. The result stated **73 percent** are very satisfied. Based on these results, should we accept or reject the CEO's hypothesis? Assume a **significance level of 0.05.** [4]

Let **π** be the true proportion of very satisfied customers
Given: n = 100, p_s= 0.73
Now let's follow the five steps:

1. H_0: $\pi \geq 0.8$
2. H_1: $\pi < 0.8$ (left-tailed test)
3. $\alpha = 0.05$
4. Following Figure 1, it is a one-sample left-tailed Z test for true proportion π.

```
import math
import scipy.stats as stats

pi = 0.8
ps = 0.73
n = 100
alpha = 0.05

# possible types "right-tailed, left-tailed, two-tailed"
tail_hypothesis_type = "left-tailed"

if tail_hypothesis_type == "two-tailed":
    alpha /= 2

print("One-Sample", tail_hypothesis_type, "Z-test of true proportion")
print("--------------------------------------------------------------------------------")
# The p-value approach
print("Approach 1: The p-value approach to hypothesis testing in the decision rule")
if n*ps >= 5 and n*(1-ps) >= 5:
    print("Since n*ps >= 5 and n*(1-ps) >= 5, by CLT ")
    z_score = (ps - pi)/math.sqrt((pi*(1-pi))/n)
    if tail_hypothesis_type == "left-tailed":
        p_value = stats.norm.cdf(z_score)
    elif tail_hypothesis_type == "right-tailed":
        p_value = 1 - stats.norm.cdf(z_score) # or stats.norm.sf(z_score)
    else:
        z_score = abs(z_score)
        p_value = 1 - stats.norm.cdf(z_score)

    conclusion = "Failed to reject the null hypothesis."
    if p_value <= alpha:
        conclusion = "Null Hypothesis is rejected."

    print("z-score is:", z_score, " and p value is:", p_value)
    print(conclusion)
else:
    print("CLT is not satisfied")

# The critical value approach
print("\n-------------------------------------------------------------------------------
print("Approach 2: The critical value approach to hypothesis testing in the decision rule")
if n*ps >= 5 and n*(1-ps) >= 5:
    print("Since n*ps >= 5 and n*(1-ps) >= 5, by CLT ")
    z_score = (ps - pi)/math.sqrt((pi*(1-pi))/n)
    critical_value = stats.norm.ppf(abs(alpha))

    conclusion = "Failed to reject the null hypothesis."
    if tail_hypothesis_type == "left-tailed":
```

```
    if tail_hypothesis_type == "left-tailed":
        if z_score < critical_value:
            conclusion = "Null Hypothesis is rejected."
    elif tail_hypothesis_type == "right-tailed":
        critical_value = abs(critical_value)
        if z_score > critical_value:
            conclusion = "Null Hypothesis is rejected."
    else:
        z_score = abs(z_score)
        critical_value = abs(critical_value)
        if z_score > critical_value:
            conclusion = "Null Hypothesis is rejected."

    print("z-score is:", z_score, " and critical value is:", critical_value)
    print(conclusion)
else:
    print("CLT is not satisfied")
```

```
One-Sample left-tailed Z-test of true proportion
--------------------------------------------------------------------------------
Approach 1: The p-value approach to hypothesis testing in the decision rule
Since n*ps >= 5 and n*(1-ps) >= 5, by CLT
z-score is: -1.7500000000000016  and p value is: 0.04005915686381696
Null Hypothesis is rejected.

--------------------------------------------------------------------------------
Approach 2: The critical value approach to hypothesis testing in the decision rule
Since n*ps >= 5 and n*(1-ps) >= 5, by CLT
z-score is: -1.7500000000000016  and critical value is: -1.6448536269514729
Null Hypothesis is rejected.
```

Conclusion: We have sufficient evidence that the true proportion of very satisfied customers is smaller than 0.8 percent, at 5% significance level.

Recommended Reading

Two-Sample Hypothesis Tests, with Python

The Complete Beginner Guide to perform Two-Sample Hypothesis Tests (with code!)

levelup.gitconnected.com

Chi-Square Test, with Python

The Complete Beginner Guide to perform Chi-Square Test (with code!)

towardsdatascience.com

Open in app ↗

Two-Way ANOVA Test, with Python

The Complete Beginner's Guide to perform Two-Way ANOVA Test (with code!)

towardsdatascience.com

McNemar's Test, with Python

The Complete Beginner's Guide to perform McNemar's Test (with code!)

towardsdatascience.com

References

[1] "Critical Value and the p-Value • SOGA • Department of Earth Sciences." [Online]. Available: https://www.geo.fu-berlin.de/en/v/soga/Basics-of-statistics/Hypothesis-Tests/Introduction-to-Hypothesis-Testing/Critical-Value-and-the-p-Value-Approach/index.html

[2] P. A. Mackowiak, S. S. Wasserman, and M. M. Levine, "A Critical Appraisal of 98.6°F, the Upper Limit of the Normal Body Temperature, and Other Legacies of Carl Reinhold August Wunderlich," *JAMA J. Am. Med. Assoc.*, vol. 268, no. 12, pp. 1578–1580, Sep. 1992.

[3] P. M. Pallett, S. Link, and K. Lee, "New 'golden' ratios for facial beauty," *Vision Res.*, vol. 50, no. 2, pp. 149–154, Jan. 2010.

[4] "Hypothesis Test: Proportion." [Online]. Available: https://stattrek.com/hypothesis-test/proportion.aspx

Hypothesis Testing Statistics Data Science Data Analysis

Follow

Written by Chao De-Yu

319 Followers · Writer for Level Up Coding

Data Analyst | MSc. Artificial Intelligence | LinkedIn — https://www.linkedin.com/in/thet-thet-yee-deyu/

Recommended from Medium