Homework Stats 2
Week 2
6733172621 Patthadon Phengpinij

Collaborators. ChatGPT (for LaTeX styling and grammar checking)

---

1 Null Hypothesis Significance Test

Problem 1. Hamtaro and his entertainment:

Hamtaro operates an entertainment website called HamHub. Last Monday, the most famous website in this field of entertainment was blocked by the government. Hearing the news, Hamtaro wants to know whether the blockade has a significant effect on the number of HamHub's visitors?

- Before last Monday, the average number of visitors was $x_0 \sim \mathcal{N}(10000, \sigma^2)$.

- After last Monday, Hamtaro spent ten days collecting the number of users (stored in `X1`).

1.1 Can Hamtaro conclude that the blockade significantly increases the number of visitors with a significant level of 0.001?

Solution. To make the Hypothesis Testing, we first need to formulate the null hypothesis H_0 and alternative hypothesis H_A:

$$H_0 : \mu = 10000 \quad \text{(The blockade does not increase the number of visitors)}$$

$$H_A : \mu > 10000 \quad \text{(The blockade increases the number of visitors)}$$

Thus, this test is a one-sided test (right). Next, we want to calculate p-value to determine whether we can reject the null hypothesis or not.

Using the provided data, `X1`:

```
X1 = array([
    10190.25479236, 10082.65748517, 10161.37971691,
    10042.27783459, 10129.73858138,  9962.73586162,
    10187.78833611, 10013.48007958, 10372.98760763,
    10238.55408072
])
```

we can compute p-value from t score (9 degrees of freedom, since we have 10 samples) by using python code:

```
1  def calculate_sample_variance(X):
2    # HINT: Look up parameter ddof of numpy.var
3    # [YOUR CODE HERE]
4    return np.var(X, ddof=1)
5
6  def studentized_mean(X, mu0):
7    # [YOUR CODE HERE]
8    return (np.mean(X) - mu0) / ( np.sqrt(
       calculate_sample_variance(X) / X.shape[0]))
9
10 t_score = studentized_mean(X1, 10000)
11
12 # [YOUR CODE HERE]
13 # Calculate p-value.
14 # HINT: scipy.stats.t
15 pvalue = 1 - stats.t.cdf(abs(t_score), X1.shape[0] - 1)
16
17 # Check if it is equal to the library
18 result = stats.ttest_1samp(X1, 10000, alternative="
       greater")
19 print(tuple(float(p.round(6)) for p in [result[1], pvalue
       ]))
```

From the output of the code, we find that the p-value is approximately 0.002694 (for both methods). Because the p-value $p = 0.002694 > \alpha = 0.001$, we **do not reject** the null hypothesis H_0.

1.2 If the sample mean and variance are held the same, what is the minimum number of samples Hamtaro need to reject the null hypothesis? For the same observation effect, larger sample size will result in a significant result.

Solution. From the previous calculation, we have:

- Sample mean: $\bar{x} \approx 10138.19$

- Sample variance: $s^2 \approx 14400.28$

- Sample standard deviation: $s \approx 120.00$

- Sample size: $n = 10$

To find the minimum number of samples n needed to reject the null hypothesis at a significance level of $\alpha = 0.001$, we can use the formula for the t-score:

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

We want to find the smallest n such that the p-value is less than 0.001.

Using python code to find the minimum n:

```
1   sample_mean = np.mean(X1)
2   population_mean = 10000
3
4   sample_std_dev = np.std(X1, ddof=1)
5
6   alpha = 0.001
7
8   n = 10
9   while (True):
10      t = (sample_mean - population_mean) / (sample_std_dev
         / math.sqrt(n))
11      p_value = 1 - stats.t.cdf(abs(t), df=n-1)
12
13      if p_value < alpha:
14          break
15
16      n += 1
17
18  print(f"Minimum sample size required: {n}, p-value: {
        p_value.round(6)}")
```

From the result, we find that the minimum sample size required to reject the null hypothesis at a significance level of $\alpha = 0.001$ is:

$$\boxed{n = 13, \text{ with p-value } p = 0.000671}$$

Problem 2. T-Test:

Hamtaro performs a t-test for the null hypothesis $H_0 : \mu = 10$ at significance level $\alpha = 0.05$ from a dataset consisting of $n = 16$ elements with sample mean 11 and sample variance 4.

2.1 Should we reject the null hypothesis in favor of $H_A : \mu \neq 10$

2.2 What if we test against $H_{A'} : \mu > 10$?

Solution. These are hypothesis testing problems for a t-distribution with $n - 1 = 15$ degrees of freedom. Note that, the first alternative hypothesis H_A is two-sided, while the second alternative hypothesis $H_{A'}$ is one-sided (right).

Each subproblems can be solved using the following python code:

```
1   mu = 11; s2 = 4; n = 16
2
3   t_score = (mu - 10) / (math.sqrt(s2) / math.sqrt(n))
4   pvalue_1tail = 1 - stats.t.cdf(abs(t_score), n - 1)
5   pvalue_2tail = 2 * pvalue_1tail
```

The results are:

1. `pvalue_2tail = 0.0639`
   which is greater than $\alpha = 0.05$, so we $\boxed{\textbf{(2.1) do not reject}}$ the null hypothesis H_0.

2. `pvalue_1tail = 0.0320`
   which is lower than $\alpha = 0.05$, so we $\boxed{\textbf{(2.2) reject}}$ the null hypothesis H_0.

Problem 3. Hamtaro and his entertainment #2

The story in this problem is a parallel universe of problem 1.

Last Monday, Hamtaro added the new channel to the website, and he wanted to know its effects on the number of visitors. However, the most famous website in this field of entertainment was also blocked by the government on the same day. Since there was no sign of unblocking from the government, Hamtaro could not perform a hypothesis testing on only the factor of adding the new channel. How could Hamtaro know that the changes from adding the new channel are significant?

There are four scenarios in this problem:

1. Before the last Monday, the average number of visitors was

$$x_0 \sim \mathcal{N}(\mu_0, \sigma^2) \quad \text{(no block + no new channel)}.$$

2. After the last Monday, the average number of visitors are

$$x_1 \sim \mathcal{N}(\mu_1, \sigma^2) \quad \text{(block + new channel)}.$$

3. Days after removing the channel, the average number of visitors are

$$x_2 \sim \mathcal{N}(\mu_2, \sigma^2) \quad \text{(block + no new channel)}.$$

4. In an imaginative scenario that the new channel is added but the most famous website haven't been blocked, the average number of visitors is

$$x_3 \sim \mathcal{N}(\mu_3, \sigma_3^2) \quad \text{(no block + new channel)}.$$

Assuming that a user decides to visit the website because of the blockade, a new channel, or none of the two (independent).

3.1 Hamtaro found the p-value of 0.03 from doing a t-test on $H_A : x_1 > x_0$. Can he conclude that adding the new channel significantly increases the number of visitors? Justify your answer.

> **Solution.**
>
> **No**, Hamtaro cannot conclude that adding the new channel significantly increases the number of visitors based on the t-test result of $H_A : x_1 > x_0$ with a p-value of 0.03.
>
> This is because the observed increase in visitors could be attributed to the blockade rather than the new channel. To isolate the effect of the new channel, Hamtaro would need to compare the number of visitors in scenario 4 (x_3) with scenario 1 (x_0), which represents the situation without the blockade but with the new channel.

3.2 Hamtaro did another t-test and found the p-value of 0.1 from testing $H_A : x_1 > x_2$. Does he now have enough information to conclude anything about x_3 ?

> **Solution.**
>
> **No**, Hamtaro does not have enough information to conclude anything about x_3 based on the t-test result of $H_A : x_1 > x_2$ with a p-value of 0.1.
>
> The t-test comparing x_1 and x_2 only provides information about the effect of the new channel in the presence of the blockade. It does not provide any information about the scenario where the new channel is added without the blockade (x_3).

3.3 Does the current setups, 1. and 2., lead to the final question about the significance of adding the new channel?

- If yes, what should you do next to get the final answer?
- If no, Can we use the hypothesis testing answer to solve this problem?
  - If yes, design your testing, describe assumptions you made.
  - If no, explain why.

Solution.

From the current setups, 1. and 2., we **cannot** directly conclude the significance of adding the new channel.

To isolate the effect of the new channel, Hamtaro would need to conduct a hypothesis test comparing scenario 4 (x_3) with scenario 1 (x_0). This would involve collecting data on the number of visitors when the new channel is added without the blockade and comparing it to the baseline scenario without both the blockade and the new channel.

The null hypothesis H_0 and alternative hypothesis H_A for this test would be:

$$H_0 : \mu_3 = \mu_0 \quad \text{(The new channel does not increase the number of visitors)}$$

$$H_A : \mu_3 > \mu_0 \quad \text{(The new channel increases the number of visitors)}$$

The information from the previous hypothesis tests (comparing x_1 with x_0 and x_1 with x_2) cannot be directly used to solve this problem, as they do not isolate the effect of the new channel without the blockade.

TO SUBMIT

Problem 4. Hamtaro and his casino:

After opening HamHub for a short while, the website was also banned by the government since it contains some 'immoral' videos. Hamtaro then moves on and follows his other passionate dream of creating a gambling empire. Therefore, he hones his skills on public gambling websites which can be easily found even if they are illegal.

After playing for a while, he notices that the online gambling business has great business potential since the risk of gambling websites being banned is much lower than his previous entertainment business. Thus, he decides to open his own online casino.

At the opening date, he offers only a dice game. The rule is simple, the player selects a number and rolls a die. The player will receive a reward if the rolled number is the same as the one he chooses. Hamtaro wants to maximize his profit by cheating using a biased die. Since it is an online casino, he could easily change the biasness of the die after the player selects a number. However, the player is not a fool and would notice if it is too biased.

As a player,

4.1 Formulate the null hypothesis H_0 and alternative hypothesis H_A to investigate the biasness of the dice.

Solution. From the problem statement, we want to investigate whether the die is biased or not. Let p be the probability of rolling the selected number. We can formulate the hypotheses as follows:

$$H_0 : p = \frac{1}{6} \quad \text{(The die is fair)} \quad \text{and} \quad H_A : p < \frac{1}{6} \quad \text{(The die is biased)}$$

The null hypothesis H_0 states that the die is fair, meaning the probability of rolling the selected number is $\frac{1}{6}$. The alternative hypothesis H_A states that the die is biased, meaning the probability of rolling the selected number is less than $\frac{1}{6}$.

4.2 Should the H_A be one-sided or two-sided? What are the differences and benefits over another in this problem?

Solution. In this problem, $\boxed{\text{the alternative hypothesis } H_A \text{ should be one-sided}}$. A one-sided hypothesis test is appropriate here because we are only interested in detecting if the die is biased towards rolling the selected number more frequently than expected.

4.3 The player found the selected number is rolled out 3 out of 30 attempts. If he wants no more than 10% of type-I error, can he reject the H_0? Justify your answer.

Solution. This is a hypothesis testing problem for a binomial distribution.

$$H_0 : p = \frac{1}{6} \; ; X \sim \text{Binomial}(30, \frac{1}{6})$$

We need to determine whether the observed value $X = 3$ falls within the rejection region for a significance level of $\alpha = 0.10$. Because we are conducting a one-sided test, we need to consider only the lower tail of the distribution.

$$\text{The rejection region is } \alpha = 0.10$$

Calculating the cumulative probabilities, we find:

- The lower tail critical value k such that $P(X \leq k) \leq 0.1$

Consider the cumulative probabilities:

$$P(X \leq k) = \sum_{i=0}^{k} P(X = i) = \sum_{i=0}^{k} \binom{30}{i} \left(\frac{1}{6}\right)^i \left(\frac{5}{6}\right)^{30-i}$$

Calculating $P(X \leq k = 3)$:

$$P(X \leq 3) \approx 0.184$$

Since $P(X \leq 3) \approx 0.184 > 0.10$, we do not reject H_0.

4.4 If the player plays 200 games, what is the rejection region if he wants no more than 10% type-I error?

Solution. Like before, we need to determine the rejection region for a significance level of $\alpha = 0.10$. Consider the cumulative probabilities:

$$P(X \leq k) = \sum_{i=0}^{k} P(X = i) = \sum_{i=0}^{k} \binom{200}{i} \left(\frac{1}{6}\right)^i \left(\frac{5}{6}\right)^{200-i}$$

Using python to calculate the cumulative probabilities:

```
1   number_of_trials = 200
2   p_null = 1/6
3   alpha = 0.10
4   cumulative_prob = 0.0
5
6   for k in range(number_of_trials + 1):
7       cumulative_prob += math.comb(number_of_trials, k) * (
        p_null ** k) * ((1 - p_null) ** (number_of_trials - k)
        )
8       if cumulative_prob >= alpha:
9           lower_critical_value = k
10          break
11
12  print('Lower critical value:', lower_critical_value - 1)
```

from the result, we find that the rejection region is:

$$\boxed{\text{Rejection Region: } X \leq 26}$$

4.5 What would be the result in 4. if the true distribution is approximated by the Normal distribution?

Solution. Using the Normal approximation to the Binomial distribution, we have:

$$\mu = np = 200 \times \frac{1}{6} = \frac{100}{3}$$

$$\sigma = \sqrt{np(1-p)} = \sqrt{200 \times \frac{1}{6} \times \frac{5}{6}} = \sqrt{\frac{1000}{18}} = \frac{10\sqrt{5}}{3}$$

Using z-score table to find the critical value for $\alpha = 0.10$:

$$P(Z \leq z_\alpha) = 0.10 \implies z_\alpha \approx -1.28$$

Calculating the lower critical value using the z-score:

$$X = \mu + z_\alpha \sigma = \frac{100}{3} + (-1.28) \times \frac{10\sqrt{5}}{3} \approx 26.05$$

Thus, the rejection region using the Normal approximation is:

$$\boxed{\text{Rejection Region: } X \leq 26}$$

As a Hamtaro,
(**Hint:** Problem 6 and 7 are related to test power)

4.6 The mastermind Hamtaro observes that players will play no more than 200 games a day. He knows that some players studied **COMP ENG MATH 2** and might perform hypothesis testing to check whether Hamtaro cheats. Hamtaro assumes that the players will use a significant level of 0.01. He thinks that it is safe enough if the probability of being caught by a player is less than 0.05. What should be the lowest probability of rolling the selected number? (How much bias can he put in the dice) Answer in floating number with a precision of 3.

Solution. First, we need to determine the rejection region for a significance level of $\alpha = 0.01$ with $n = 200$. Using python to calculate the cumulative probabilities:

```
1  number_of_trials = 200
2  p_null = 1/6
3  alpha = 0.10
4  cumulative_prob = 0.0
5
6  for k in range(number_of_trials + 1):
7      cumulative_prob += math.comb(number_of_trials, k) * (
       p_null ** k) * ((1 - p_null) ** (number_of_trials - k)
       )
8      if cumulative_prob >= alpha:
9          lower_critical_value = k
10         break
11
12 print('Lower critical value:', lower_critical_value - 1)
```

From the result, we find that the rejection region is:

$$\text{Rejection Region: } X \leq 21$$

Next, we need to find the lowest probability p such that the power of the test is at least 0.95 (i.e., the probability of being caught is less than 0.05).

```
1  target_power = 0.95
2
3  for p in np.arange(0, 1, 0.001):
4      cumulative_prob = 0.0
5      for k in range(22):
6          cumulative_prob += math.comb(number_of_trials, k)
       * (p ** k) * ((1 - p) ** (number_of_trials - k))
7      power = 1 - cumulative_prob
8      if power >= target_power:
9          lowest_p = p
10         break
11
12 print('Lowest probability p:', round(lowest_p, 3))
```

From the result, the lowest probability p that Hamtaro can use is:

$$\boxed{p = 0.148}$$

4.7 What if Hamtaro accepts the probability of being caught equal to 0.01 instead? Answer in floating number with the precision of 5.

Solution. From the previous calculation, we need to find the lowest probability p such that the power of the test is at least 0.99 (i.e., the probability of being caught is less than 0.01).

```
1  target_power = 0.99
2
3  for p in np.arange(0, 1, 0.001):
4      cumulative_prob = 0.0
5      for k in range(22):
6          cumulative_prob += math.comb(number_of_trials, k)
       * (p ** k) * ((1 - p) ** (number_of_trials - k))
7      power = 1 - cumulative_prob
8      if power >= target_power:
9          lowest_p = p
10         break
11
12 print('Lowest probability p:', round(lowest_p, 3))
```

From the result, the lowest probability p that Hamtaro can use is:

$$\boxed{p = 0.167}$$

Problem 5. Hamtaro and the new AC company:

In a datacenter for HamHub, Hamtaro tried to control the temperature of the cloud storage room so that the failure rate of storage disks is as low as possible. Later, a new company came to him and offered a new AC system that, in theory, could provide a more stable room temperature. To strengthen their claim, the company sends the historical room temperature of the new AC, of which the target temperature is set 15, to Hamtaro.

Given that Hamtaro's existing AC provides the room temperature of $t \sim \mathcal{N}(15, 0.5^2)$:

```
1  temp_log = np.array([
2      14.66017243, 14.82134507, 14.75354867, 15.02847413,
3      15.96338554, 15.46598137, 15.35605532, 14.91048177,
4      15.13237189, 14.38789873, 15.76833691, 14.85383663,
5      15.28335022, 15.06718901, 15.44364169, 14.29511914,
6      15.13458572, 14.57428013, 15.14885716, 14.08580661,
7      15.60066540, 14.98109974, 14.95059512, 14.91460432,
8      14.68809902, 15.49886170, 14.99646465, 15.00654947,
9      14.65024467, 15.20684546, 15.54078700, 15.39207656,
10     14.53129171, 14.27527689, 14.37856735, 15.46854760,
11     14.94268835, 15.28311368, 14.88781520, 15.52350034,
12     14.35791689, 15.11741279, 15.41721681, 15.56690632,
13     15.30108101, 14.71389760, 15.39536719, 15.02994055,
14     14.74887633, 14.81419334, 15.36735467, 14.89706838,
15     14.89134826, 15.19781408, 15.32733540, 15.16729623,
16     14.82748547, 15.59488402, 15.49763473, 15.12876929,
17     14.11446324, 14.61298282, 14.57006854, 15.13227246,
18     14.68369474, 14.96443757, 15.73872741, 15.48498884,
19     15.35770021, 15.13471147, 14.94871779, 13.91322937,
20     14.84786617, 14.42086587, 15.26216287, 14.33225067,
21     14.94179209, 14.57095395, 15.12615130, 14.93201265,
22     14.82252959, 15.19061294, 15.33257912, 14.72448901,
23     15.54406202, 14.72704346, 14.99027730, 14.71477903,
24     14.90866689, 14.28862563, 15.04302902, 15.06973955,
25     14.51951387, 14.61413562, 14.58725869, 14.41407727,
26     15.05585075, 14.69229146, 14.30425173, 14.76913898,
27     14.27819269, 14.93917912, 14.22675051, 15.20964000,
28     14.96122782, 14.05371218, 15.10273752, 15.50886439,
```

```
29        15.43965366,  14.98863063,  15.34326459,  15.23694786,
30        13.90170147,  15.29660252,  15.26635161,  15.34710713,
31        14.34928594,  15.61509746,  15.80476574,  15.36769161,
32        14.52027993,  14.80624255,  14.58269606,  15.58830065,
33        14.25665696,  14.86914893,  15.40500584,  15.28855103,
34        15.43907472,  15.18196326,  15.47088551,  15.06327054,
35        15.01022434,  14.43508736,  15.37918870,  14.86202479,
36        15.16977660,  14.64346330,  15.72632770,  14.31813452,
37        15.30657752,  14.91471004,  15.14566170,  14.93856484,
38        15.14098396,  14.76996958,  15.38908210,  15.53549397,
39        15.28528007,  15.61416247,  14.45143470,  14.75105769,
40        14.22367585,  14.93898327,  14.61033024,  14.96348807,
41        15.24771829,  14.84653005,  15.36780845,  14.96846837,
42        14.66094081,  14.75905691,  14.96864336,  15.55687252,
43        14.62138304,  16.02201637,  14.95786084,  14.98549356,
44        15.18029872,  14.82305383,  15.09356200,  15.98065684,
45        15.27950419,  15.42169411,  15.66950953,  14.90725077,
46        13.69523862,  15.74709530,  14.93824139,  15.65590845,
47        14.69911713,  14.63306529,  15.09566097,  15.00531748,
48        15.06648240,  15.00496274,  15.15775270,  15.26365236,
49        14.98708579,  14.43256043,  15.58167070,  14.69227952,
50        15.22774367,  15.01510129,  15.03105086,  15.07222669,
51        15.22579141,  15.34835664,  15.14017702,  15.12604511
52  ])
```

5.1 Formulate the null and alternative hypotheses for determining whether the new AC is better than the existing one or not. List your assumptions that are required to make this experiment possible.

Solution. Firstly, let us set the significance level $\alpha = 0.05$.

From the problem statement, we want to investigate whether the new AC system provides a <u>more stable</u> room temperature compared to the existing one. Let σ^2_{new} be the variance of the room temperature provided by the new AC system.

We can formulate the hypotheses as follows:

$$\boxed{H_0 : \sigma^2_{\text{new}} = 0.5^2 \quad \text{(The new AC is as same as the existing one)}}$$

$$\boxed{H_A : \sigma^2_{\text{new}} < 0.5^2 \quad \text{(The new AC is better than the existing one)}}$$

Assumptions required to make this experiment possible:

- The room temperature data from both AC systems are normally distributed.

- The samples from both AC systems are independent of each other.

In order to test these hypotheses, we can use a Chi-square test for variance. The test statistic for the Chi-square test is given by:

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$$

Substitute the values into the formula to calculate the test statistic and compare it to the critical value from the Chi-square distribution table with $n-1$ degrees of freedom at a significance level of α.

Homework Stats 2

```
1  temp_log_std = np.std(temp_log, ddof=1)
2  n = temp_log.shape[0]
3
4  chi_square = (n - 1) * (temp_log_std ** 2) / (0.5 ** 2)
5  p_value = stats.chi2.cdf(chi_square, df=n - 1)
```

The result shows that the p-value $p = 0.0015$ is less than $\alpha = 0.05$, so we

$$\boxed{\textbf{reject the null hypothesis, } H_0}$$

and conclude that the new AC system is better than the existing one in terms of providing a more stable room temperature.

5.2 Can you decide which AC system is better? Justify your answer.

Solution. First, we need to compare the mean of the two AC systems in addition to the variance. Using the data stored in `temp_log`, we can calculate the mean and standard deviation of the new AC system.

$$\bar{t}_{new} = 15.005 \quad \text{(very close to the target temperature of 15)}$$

Combine this with the previous result that the variance of the new AC system is significantly lower than that of the existing one, we can conclude that the **new AC system is better** than the existing one in terms of stability (lower variance) and also kept the mean close to the target temperature.

<div align="center">

TO SUBMIT

</div>

Problem 6. Hamtaro and his casino:

Hamtaro also have a factory. He tried to boost the factory productivity by replacing the old machines with a new type-II variant. However, there is a concern from the local factory managers that Hamtaro might get bamboozled, since they do not observe an increase in productivity compared to the previous one. Therefore, to ease their concern, he decided to conduct a z-testing.

Given that the number of goods produced each day by the old machines was $x \sim \mathcal{N}(5000, 20^2)$:

6.1 Formulate the null and alternative hypothesis for determining whether the new machine is better than the previous one at a significant level $= 0.05$.

Solution. From the problem statement, we want to investigate whether the new machine increases productivity compared to the old machine. Let μ be the mean productivity of the new machine. We can formulate the hypotheses as follows:

$$\boxed{H_0 : \mu = 5000 \quad \text{(The new machine does not increase productivity)}}$$

$$\boxed{H_A : \mu > 5000 \quad \text{(The new machine increases productivity)}}$$

The null hypothesis H_0 states that the new machine does not increase productivity, meaning the mean productivity is equal to 5000. The alternative hypothesis H_A states that the new machine increases productivity, meaning the mean productivity is greater than 5000.

For the following problems, we will use the data from 4 factories provided below. The following information is given for the testing:

```python
from scipy.stats import norm
import numpy as np

# 30 days of product quantity in 4 factories

fac_0 = np.array([
    4993.89323126, 5021.67118211, 5023.54710937,
    4999.11746331, 5001.53450095, 4986.27990953,
    4987.12362188, 5004.91535087, 4999.97591193,
    5038.09176163, 4993.94184053, 5026.52644680,
    5040.62862593, 4979.91124088, 5008.59143715,
    5016.45331659, 5013.63203948, 5010.84253735,
    5014.99772195, 5002.39462129, 5047.80507624,
    5007.23005532, 5019.87205007, 5005.76363012,
    4997.09106036, 4982.80291132, 5037.18158407,
    4996.54197735, 5007.57964251, 4971.18880247
])

fac_1 = np.array([
    5036.80041897, 4989.33779117, 4971.68709581,
    5041.92493487, 5041.64823146, 5026.33602398,
    5009.58334612, 4989.05382998, 5031.17423169,
    4992.20198911, 4970.63425555, 5007.17615704,
    4993.84416738, 5028.59671588, 5009.91388156,
    5049.64187466, 5015.12711371, 5032.29005130,
    5013.66869347, 4988.21257317, 5020.44276181,
    4985.62886808, 5022.46800468, 5042.35501669,
    5001.61539080, 5012.14209858, 5006.14666402,
    4999.93219541, 5002.77927647, 5002.20750425
])

fac_2 = np.array([
    5029.95293241, 5019.47959949, 4976.84278360,
    4985.22792264, 4994.97618684, 5026.75059569,
    5015.71350753, 5008.46632673, 5037.96915682,
    4990.38948551, 4988.70822060, 5032.42440206,
    5036.41040953, 5003.75236158, 5002.62361815,
    4998.89320570, 5000.51153033, 5002.19196574,
    5023.74534474, 5032.03601587, 5000.10614764,
    4989.74566985, 4985.97436664, 4973.63380449,
    5028.58100504, 4997.84267810, 5011.42021980,
    5018.71432385, 4969.03296199, 5009.23456565
])
```

```
1  fac_3 = np.array([
2      4962.36508403,  5015.91734917,  5030.86885403,
3      5012.74787091,  5036.94455211,  4995.21037570,
4      5029.84241184,  5015.68062582,  4996.43546786,
5      4999.57614716,  5006.88735305,  5035.10432486,
6      5017.33437936,  5006.70625696,  5007.97827037,
7      4981.80482708,  5020.78603239,  4993.12742287,
8      4996.10718141,  4988.00315629,  5003.00004152,
9      4949.54117305,  5008.62500480,  5004.09075453,
10     5026.56246304,  5011.02296759,  5010.67413795,
11     4990.58062539,  5009.64435703,  5001.94134280
12  ])
```

For the problem 6.2., we will create a `NDArray` that contains all the data from the 4 factories, `fac_whole`. And, we also create `factories`, which is a dictonary that maps the factory name to its corresponding data.

```
1  fac_whole = np.concatenate((fac_0, fac_1, fac_2, fac_3))
2
3  factories = {
4      "Factory 0": fac_0,
5      "Factory 1": fac_1,
6      "Factory 2": fac_2,
7      "Factory 3": fac_3,
8      "All Factories": fac_whole
9  }
```

6.2 From the testing, can Hamtaro conclude that factory productivity increased as a whole?

Solution. To determine whether Hamtaro can conclude that factory productivity increased as a whole, we need to perform a z-test using the provided data. Assuming we have the sample mean $\bar{x}$ and sample size n, we can calculate the z-score as follows:

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

where $\mu_0 = 5000$ is the mean under the null hypothesis, and $\sigma = 20$ is the standard deviation of the old machine.

Next, we compare the calculated z-score to the critical value for a one-tailed test at a significance level of $\alpha = 0.05$. Considering the z-table, the critical value is approximately:

$$P(Z \leq z) = 0.05 \implies z_\alpha \approx 1.645$$

.
Calculating the z-score with the provided data will give us the final conclusion.

$$z = \frac{\bar{x} - 5000}{20/\sqrt{n}}$$

Next, using python to calculate the z-score and compare it with the critical value:

```python
mu_0 = 5000
sigma_0 = 20
z_alpha = norm.ppf(0.95)

for name, data in factories.items():
    n = len(data)
    sample_mean = np.mean(data)

    z_score = (sample_mean - mu_0) / (sigma_0 / np.sqrt(n
    ))

    decision = "Reject" if z_score > z_alpha else "Accept
    "

    print(f"{name}: z-score = {z_score:.4f}, Decision: {
    decision}")
```

The result from the code above are as follows:

Factory	z-score	Decision
Factory 0	2.1647	Reject
Factory 1	3.0542	Reject
Factory 2	1.7468	Reject
Factory 3	1.5072	Accept
All Factories	4.2365	Reject

The results show that for "All Factories", the z-score is greater than the critical value, leading to the rejection of the null hypothesis.

Thus, Hamtaro can conclude that factory productivity increased as a whole.

6.3 Can Hamtaro say the same for each individual factory?

Solution. Using the results from problem 6.2., we can see that for each individual factory:

Factory	z-score	Decision
Factory 0	2.1647	Reject
Factory 1	3.0542	Reject
Factory 2	1.7468	Reject
Factory 3	1.5072	Accept
All Factories	4.2365	Reject

Thus, Hamtaro cannot say the same for each individual factory. (Factory 3 does not show a significant increase in productivity.)

6.4 Repeat 6.1 - 6.3 again but with a t-test. Is there any difference from the z-test? What, and why does it happen?

Solution. To perform a t-test, we will use the sample standard deviation instead of the population standard deviation. The t-score is calculated as follows:

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

where s is the sample standard deviation.

The critical value for a one-tailed t-test at a significance level of $\alpha = 0.05$ with $n-1$ degrees of freedom can be found using the t-distribution table.

```python
from scipy.stats import t

for name, data in factories.items():
    n = len(data)
    sample_mean = np.mean(data)
    sample_std = np.std(data, ddof=1)

    t_score = (sample_mean - mu_0) / (sample_std / np.sqrt(n))
    t_alpha = t.ppf(0.95, df=n-1)

    decision = "Reject" if t_score > t_alpha else "Accept"

    print(f"{name}: t-score = {t_score:.4f}, Decision: {decision}")
```

The result from the code above are as follows:

Factory	t-score	Decision
Factory 0	2.3262	Reject
Factory 1	2.8827	Reject
Factory 2	1.8045	Reject
Factory 3	1.5589	Accept
All Factories	4.3410	Reject

The results of using the t-test are similar to those of the z-test.

The decisions are the same for all factories.

The reason for this similarity is that the sample size is **relatively large** (n=30 for each factory), which makes the t-distribution approach the normal distribution. Thus, the results of the t-test and z-test converge, leading to the same conclusions.