

Chapter 7: One sample tests

(Ott & Longnecker Sections: 5.4-5.7)

Duzhe Wang

Part 1

<https://dzwang91.github.io/stat371/>



WISCONSIN
UNIVERSITY OF WISCONSIN-MADISON



Key concepts: One-tailed test, two-tailed test, sample mean test, Z test, T test, p-value

- 1 State the null hypothesis H_0 and the alternative hypothesis H_A . The goal is to test if H_A is “likely” true.
- 2 Choose a significance level α . Typically 0.05, 0.01.
- 3 Choose the test statistic $T_n(X_1, \dots, X_n)$ and establish the rejection region. (How do we choose the test statistic and rejection region?)
- 4 Data X_1, \dots, X_n are gathered, compute the realization of the test statistic. If the test statistic is in the rejection region, we reject H_0 and accept H_A (because of sufficient evidence in the sample in favor of H_A), otherwise we do not reject H_0 (because of insufficient evidence to support H_A).

- If

- $H_0 : \theta = \theta_0$
- $H_A : \theta > \theta_0$

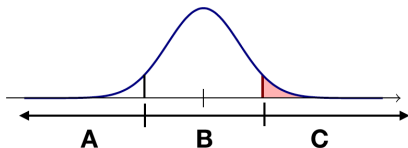
then this is a one-tailed test.

- If

- $H_0 : \theta = \theta_0$
- $H_A : \theta > \theta_0$

then this is a one-tailed test.

- If this is the distribution of the test statistic given the H_0 is true, then which part of A, B, C is the rejection region?

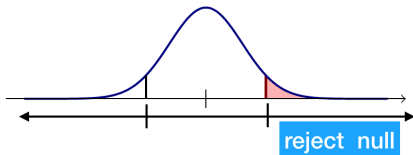


- If

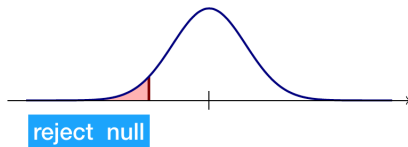
- $H_0 : \theta = \theta_0$
- $H_A : \theta > \theta_0$

then this is a one-tailed test.

- If this is the distribution of the test statistic given the H_0 is true, then which part of A, B, C is the rejection region?



- If
 - $H_0 : \theta = \theta_0$
 - $H_A : \theta < \theta_0$then this is also a one-tailed test.
- rejection region is in the left tail of the test statistic.

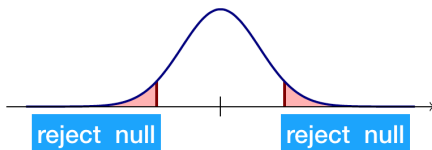


- If

- $H_0 : \theta = \theta_0$
- $H_A : \theta \neq \theta_0$

then this is a two-tailed test.

- rejection region is in both of the left and right tails of the test statistic.

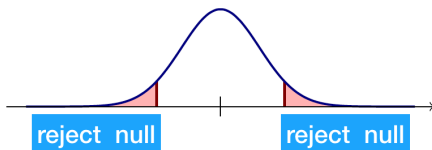


- If

- $H_0 : \theta = \theta_0$
- $H_A : \theta \neq \theta_0$

then this is a two-tailed test.

- rejection region is in both of the left and right tails of the test statistic.



- The rejection region corresponds to the alternative hypothesis.

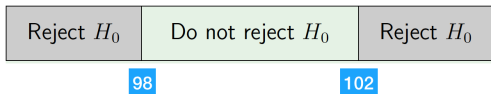


- Consider a production line of resistors that are supposed to be 100 Ohms. Assume $\sigma = 8$, so the hypotheses are
 - $H_0 : \mu = 100$
 - $H_A : \mu \neq 100$

An example of two-tailed test



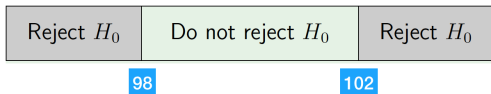
- Consider a production line of resistors that are supposed to be 100 Ohms. Assume $\sigma = 8$, so the hypotheses are
 - $H_0 : \mu = 100$
 - $H_A : \mu \neq 100$
- Let \bar{X} be the sample mean for a sample of size $n = 100$



An example of two-tailed test

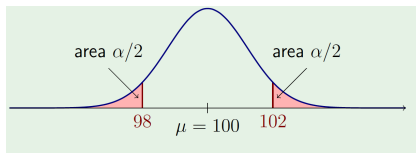


- Consider a production line of resistors that are supposed to be 100 Ohms. Assume $\sigma = 8$, so the hypotheses are
 - $H_0 : \mu = 100$
 - $H_A : \mu \neq 100$
- Let \bar{X} be the sample mean for a sample of size $n = 100$



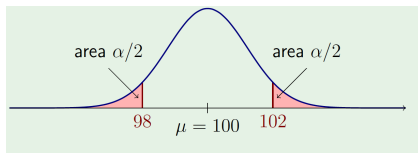
- In this case the test statistic is the sample mean. What is the sampling distribution of \bar{X} ?

An example of two-tailed test continued



- The sampling distribution of \bar{X} is a normal distribution with mean μ and standard deviation $\sigma/\sqrt{n} = 0.8$ due to [the central limit theorem](#).

An example of two-tailed test continued



- The sampling distribution of \bar{X} is a normal distribution with mean μ and standard deviation $\sigma/\sqrt{n} = 0.8$ due to [the central limit theorem](#).
- Then the probability of the Type I error is

$$\begin{aligned}\alpha &= \Pr(\bar{X} < 98 \text{ when } \mu = 100) + \Pr(\bar{X} > 102 \text{ when } \mu = 100) \\ &= \Pr\left(Z < \frac{98 - 100}{8/\sqrt{100}}\right) + \Pr\left(Z > \frac{102 - 100}{8/\sqrt{100}}\right) \\ &= \Pr(Z < -2.5) + \Pr(Z > 2.5) \\ &= 2 \times \Pr(Z < -2.5) = 2 \times 0.0062 = 0.0124.\end{aligned}$$



- In general we are often interested in testing

- $H_0 : \mu = \mu_0$
- $H_A : \mu \neq \mu_0$

based on the sample mean \bar{X} from samples X_1, \dots, X_n with **known population variance σ^2** .

- In general we are often interested in testing

- $H_0 : \mu = \mu_0$
- $H_A : \mu \neq \mu_0$

based on the sample mean \bar{X} from samples X_1, \dots, X_n with **known population variance σ^2** .

- Under $H_0 : \mu = \mu_0$, the probability of Type I error is computed using the sampling distribution of \bar{X} , which is normal distributed with mean μ_0 and standard deviation σ/\sqrt{n} .

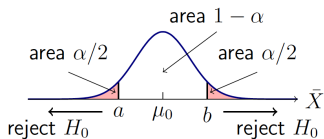
- In general we are often interested in testing

- $H_0 : \mu = \mu_0$
- $H_A : \mu \neq \mu_0$

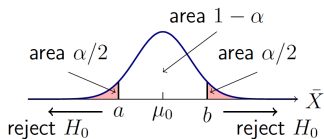
based on the sample mean \bar{X} from samples X_1, \dots, X_n with **known population variance σ^2** .

- Under $H_0 : \mu = \mu_0$, the probability of Type I error is computed using the sampling distribution of \bar{X} , which is normal distributed with mean μ_0 and standard deviation σ/\sqrt{n} .
- How do we decide the rejection region at the level of significance α ?
 - “Use the distribution of the test statistic to determine a rejection region that limits the type I error at significance level α ”

Two-tailed sample mean test continued

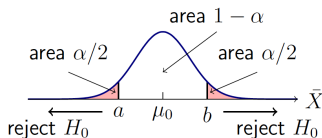


Two-tailed sample mean test continued



$$P(-z_{\alpha/2} < \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} < z_{\alpha/2}) = 1 - \alpha$$

Two-tailed sample mean test continued



$$P(-z_{\alpha/2} < \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} < z_{\alpha/2}) = 1 - \alpha$$

- Therefore, to design a test at the level of significance α we choose the critical values a and b as

$$a = \mu_0 - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$$b = \mu_0 + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

- we collect the sample, compute the sample mean \bar{X} and reject H_0 if $\bar{X} < a$ or $\bar{X} > b$.

$$H_0 : \mu = \mu_0, \text{ vs. } H_A : \mu \neq \mu_0$$

- Equivalently, we can choose the test statistic $Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$
- To design a test at the level of significance α , choose the rejection region $Z > z_{\alpha/2}$ and $Z < -z_{\alpha/2}$.

A batch of 100 resistors have an average of 102 Ohms. Assuming a population standard deviation of 8 Ohms, test whether the population mean is 100 Ohms at a significance level $\alpha = 0.05$.

A batch of 100 resistors have an average of 102 Ohms. Assuming a population standard deviation of 8 Ohms, test whether the population mean is 100 Ohms at a significance level $\alpha = 0.05$.

- Step 1: $H_0 : \mu = 100$, $H_A : \mu \neq 100$.

A batch of 100 resistors have an average of 102 Ohms. Assuming a population standard deviation of 8 Ohms, test whether the population mean is 100 Ohms at a significance level $\alpha = 0.05$.

- Step 1: $H_0 : \mu = 100$, $H_A : \mu \neq 100$.
- Step 2: Choose $\alpha = 0.05$

A batch of 100 resistors have an average of 102 Ohms. Assuming a population standard deviation of 8 Ohms, test whether the population mean is 100 Ohms at a significance level $\alpha = 0.05$.

- Step 1: $H_0 : \mu = 100$, $H_A : \mu \neq 100$.
- Step 2: Choose $\alpha = 0.05$
- Step 3: In this case, reject H_0 if $\bar{X} < a$ or $\bar{X} > b$ with

$$a = \mu_0 - z_{0.025} \frac{\sigma}{\sqrt{100}} = 100 - 1.96 \frac{8}{10} = 98.432$$

$$b = \mu_0 + z_{0.025} \frac{\sigma}{\sqrt{100}} = 100 + 1.96 \frac{8}{10} = 101.568$$

A batch of 100 resistors have an average of 102 Ohms. Assuming a population standard deviation of 8 Ohms, test whether the population mean is 100 Ohms at a significance level $\alpha = 0.05$.

- Step 1: $H_0 : \mu = 100$, $H_A : \mu \neq 100$.
- Step 2: Choose $\alpha = 0.05$
- Step 3: In this case, reject H_0 if $\bar{X} < a$ or $\bar{X} > b$ with

$$a = \mu_0 - z_{0.025} \frac{\sigma}{\sqrt{100}} = 100 - 1.96 \frac{8}{10} = 98.432$$

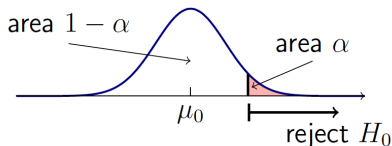
$$b = \mu_0 + z_{0.025} \frac{\sigma}{\sqrt{100}} = 100 + 1.96 \frac{8}{10} = 101.568$$

- Step 4: $\bar{X} = 102 > b$, therefore reject H_0 .

One-tailed sample mean test, right tail



We are interested in testing $H_0 : \mu = \mu_0$ vs. $H_A : \mu > \mu_0$ at the significance level α assuming the variance is known.



- Under $H_0 : \mu = \mu_0$, the probability of a Type I error is

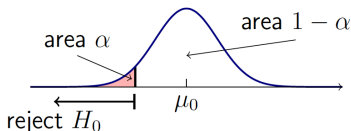
$$P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} > z_\alpha\right) = \alpha$$

- Thus the rejection region is $\bar{X} > \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}$. Note we use z_α instead of $z_{\alpha/2}$.

One-tailed sample mean test, left tail



We are interested in testing $H_0 : \mu = \mu_0$ vs. $H_A : \mu < \mu_0$ at the significance level α assuming the variance is known.



- Under $H_0 : \mu = \mu_0$, the probability of a Type I error is

$$P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} < -z_\alpha\right) = \alpha$$

- Thus the rejection region is $\bar{X} < \mu_0 - z_\alpha \frac{\sigma}{\sqrt{n}}$. Note we use z_α instead of $z_{\alpha/2}$.



A quality control engineer finds that a sample of 100 light bulbs had an average life-time of 470 hours. Assuming a population standard deviation of $\sigma = 25$ hours, test whether the population mean is 480 hours vs. the alternative hypothesis $\mu < 480$ at a significance level of $\alpha = 0.05$.

A quality control engineer finds that a sample of 100 light bulbs had an average life-time of 470 hours. Assuming a population standard deviation of $\sigma = 25$ hours, test whether the population mean is 480 hours vs. the alternative hypothesis $\mu < 480$ at a significance level of $\alpha = 0.05$.

- Step 1: $H_0 : \mu = 480$, $H_A : \mu < 480$.

A quality control engineer finds that a sample of 100 light bulbs had an average life-time of 470 hours. Assuming a population standard deviation of $\sigma = 25$ hours, test whether the population mean is 480 hours vs. the alternative hypothesis $\mu < 480$ at a significance level of $\alpha = 0.05$.

- Step 1: $H_0 : \mu = 480$, $H_A : \mu < 480$.
- Step 2: Choose $\alpha = 0.05$

A quality control engineer finds that a sample of 100 light bulbs had an average life-time of 470 hours. Assuming a population standard deviation of $\sigma = 25$ hours, test whether the population mean is 480 hours vs. the alternative hypothesis $\mu < 480$ at a significance level of $\alpha = 0.05$.

- Step 1: $H_0 : \mu = 480$, $H_A : \mu < 480$.
- Step 2: Choose $\alpha = 0.05$
- Step 3: In this case, reject H_0 if

$$\bar{X} < \mu_0 - z_{0.05} \frac{\sigma}{\sqrt{n}} = 480 - 1.645 \frac{25}{10} = 475.9$$

A quality control engineer finds that a sample of 100 light bulbs had an average life-time of 470 hours. Assuming a population standard deviation of $\sigma = 25$ hours, test whether the population mean is 480 hours vs. the alternative hypothesis $\mu < 480$ at a significance level of $\alpha = 0.05$.

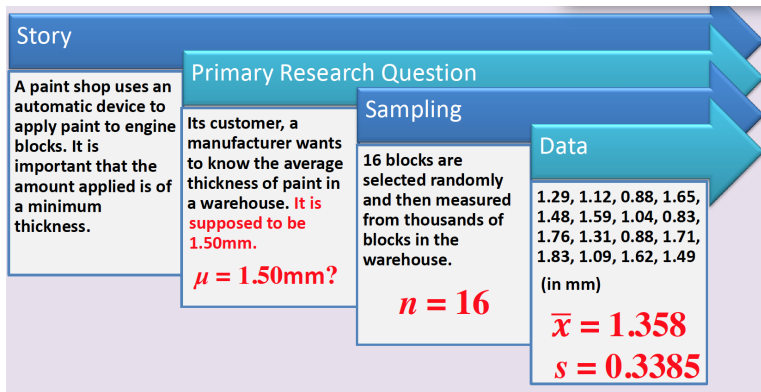
- Step 1: $H_0 : \mu = 480$, $H_A : \mu < 480$.
- Step 2: Choose $\alpha = 0.05$
- Step 3: In this case, reject H_0 if

$$\bar{X} < \mu_0 - z_{0.05} \frac{\sigma}{\sqrt{n}} = 480 - 1.645 \frac{25}{10} = 475.9$$

- Step 4: $\bar{X} = 470 < 475.9$, therefore reject H_0 .



What if the population standard deviation is **unknown**?

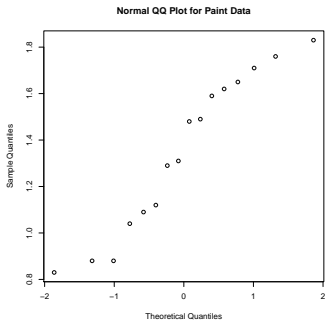




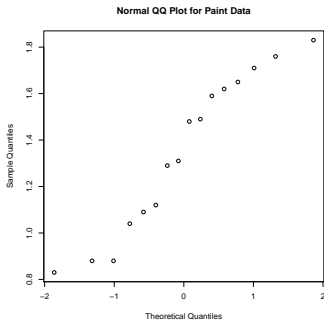


- Normality?

- Normality?

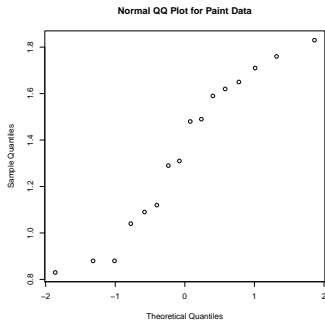


- Normality?



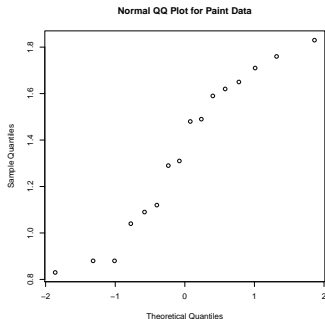
- Is population standard deviation σ known?

- Normality?



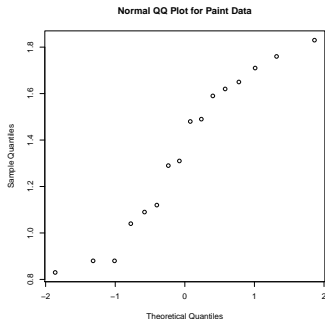
- Is population standard deviation σ known?
NO

- Normality?



- Is population standard deviation σ known?
NO
- Is the sample size n large?

- Normality?



- Is population standard deviation σ known?
NO
- Is the sample size n large?
NO

We are interested in testing $H_0 : \mu = \mu_0$ vs. $H_A : \mu \neq \mu_0$ at the significance level α based a sample X_1, \dots, X_n from a normal distribution, but with unknown variance σ^2 .

- Under $H_0 : \mu = \mu_0$, the sampling distribution of $\frac{\bar{X} - \mu_0}{s/\sqrt{n}}$ is a t-distribution with degrees of freedom $n - 1$.
- Choose test statistic $T = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$.
- Reject H_0 if $T < -t_{\alpha/2, n-1}$ or $T > t_{\alpha/2, n-1}$.
- For a one-sided test, $t_{\alpha/2, n-1}$ is replaced by $t_{\alpha, n-1}$ as usual.



- Step 1:

$$H_0: \mu = 1.5, H_A: \mu \neq 1.5$$

- Step 1:

$$H_0: \mu = 1.5, H_A: \mu \neq 1.5$$

- Step 2: Choose $\alpha = 0.05$

- Step 1:

$$H_0: \mu = 1.5, H_A: \mu \neq 1.5$$

- Step 2: Choose $\alpha = 0.05$
- Step 3: Use the T test statistic: $T = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}}$. In this example,

$$t_{obs} = \frac{1.348 - 1.50}{\frac{0.3385}{\sqrt{16}}} = -1.796. \text{ The rejection region is}$$

$T < -t_{n-1, \alpha/2}, T > t_{n-1, \alpha/2}$. In this example, $t_{15, 0.025} = 2.13$, so the rejection region is $T < -2.13$ or $T > 2.13$.

- Step 1:

$$H_0: \mu = 1.5, H_A: \mu \neq 1.5$$

- Step 2: Choose $\alpha = 0.05$

- Step 3: Use the T test statistic: $T = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}}$. In this example,

$$t_{obs} = \frac{1.348 - 1.50}{\frac{0.3385}{\sqrt{16}}} = -1.796. \text{ The rejection region is}$$

$T < -t_{n-1, \alpha/2}, T > t_{n-1, \alpha/2}$. In this example, $t_{15, 0.025} = 2.13$, so the rejection region is $T < -2.13$ or $T > 2.13$.

- Step 4: Make a conclusion: since $t_{obs} = -1.796$ does not fall in the rejection region, so we do not reject the null.

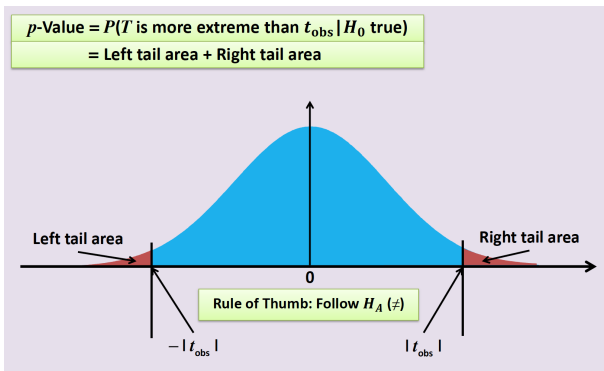


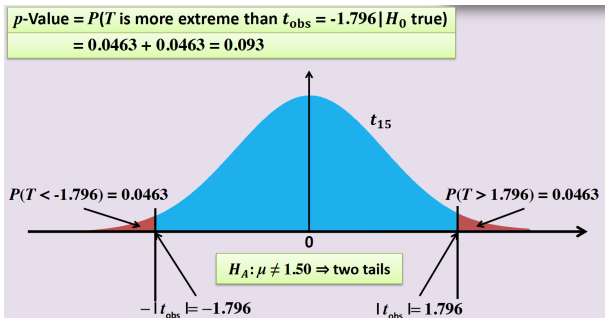
- In the approach we have taken so far, the significance level is pre-selected up front, either by choosing a given value of α or setting the rejection region explicitly.
- Suppose a hypothesis test is performed at a significance level of 0.05, but someone else wants to test with a stricter significance level of 0.01, this requires **recomputing** the rejection region.
- **Using p-value is better than using rejection region.**

Hypothesis testing using the p-value



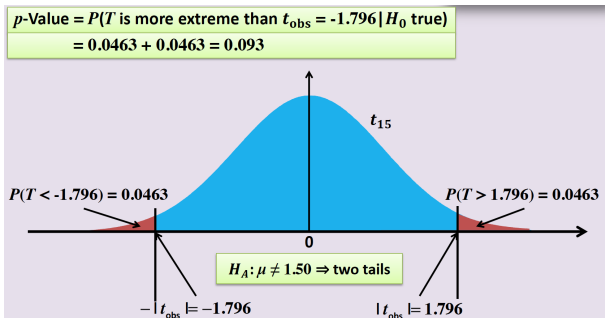
- “ If the p-value is smaller than the given significance level α , we would reject the null, otherwise we would not reject the null.”
- For a two-sided test,





- If we choose $\alpha = 0.05$, then since $0.094 > 0.05$, we would not reject the null. However, if we had chosen $\alpha = 0.1$, we would have rejected.

An example revisited



- If we choose $\alpha = 0.05$, then since $0.094 > 0.05$, we would not reject the null. However, if we had chosen $\alpha = 0.1$, we would have rejected.
- $p\text{-value}$ is a measure of evidence in the sense that it is the minimum probability of a Type I error with which H_0 can still be rejected.



We'll discuss how to calculate the sample size and power in hypothesis testing in the next lecture.