Chapter 7: One sample tests

Ott & Longnecker Sections: 5.4-5.7

Duzhe Wang

the Department of Statistics, UW-Madison

Part 1: one sample T-test
https://dzwang91.github.io/stat371/

UNIVERSITY OF WISCONSIN–MADISON

Example

Story

A paint shop uses an automatic device to apply paint to engine blocks. It is important that the amount applied is of a minimum thickness.

Primary Research Question

Its customer, a manufacturer wants to know the average thickness of paint in a warehouse. It is supposed to be 1.50mm.

$\mu = 1.50\text{mm}?$

Sampling

16 blocks are selected randomly and then measured from thousands of blocks in the warehouse.

$n = 16$

Data

1.29, 1.12, 0.88, 1.65, 1.48, 1.59, 1.04, 0.83, 1.76, 1.31, 0.88, 1.71, 1.83, 1.09, 1.62, 1.49

(in mm)

$\bar{x} = 1.358$

$s = 0.3385$

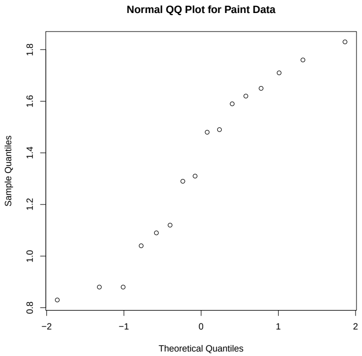
Statistical assumptions

Statistical assumptions

- Normality?

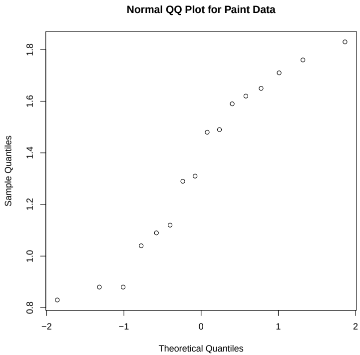
Statistical assumptions

- Normality?



Statistical assumptions

- Normality?



YES

Statistical assumptions

Statistical assumptions

- is σ known?

Statistical assumptions

- is σ known? NO

Statistical assumptions

- is σ known? NO
- is n large?

Statistical assumptions

- is σ known? NO
- is n large? NO

Statistical assumptions

- is σ known? NO
- is n large? NO
- what distribution will you use based on the above observations?

Hypotheses and T statistic

- Step 1: hypotheses

Hypotheses and T statistic

- Step 1: hypotheses

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

Hypotheses and T statistic

- Step 1: hypotheses

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

This type of H_A is called a **two-sided alternative**.

Hypotheses and T statistic

- Step 1: hypotheses

$$H_0\text{: } \mu = 1.5,\ H_A\text{: } \mu \neq 1.5$$

  This type of H_A is called a **two-sided alternative**.

- Step 2: decide a test statistic

Hypotheses and T statistic

- Step 1: hypotheses

$$H_0\text{: } \mu = 1.5,\ H_A\text{: } \mu \neq 1.5$$

  This type of H_A is called a **two-sided alternative**.

- Step 2: decide a test statistic

  " **what distribution will you use based on the above observations**?"

Hypotheses and T statistic

- Step 1: hypotheses

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

  This type of H_A is called a **two-sided alternative**.

- Step 2: decide a test statistic

  " **what distribution will you use based on the above observations**?"

$$\text{A fact: } T = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} \sim t_{n-1},$$

  where μ_0 is general notation for the value of μ under the null hypothesis; in this example, $\mu_0 = 1.5$.

Hypotheses and T statistic

- Step 1: hypotheses

$$H_0\text{: } \mu = 1.5, \; H_A\text{: } \mu \neq 1.5$$

  This type of H_A is called a **two-sided alternative**.

- Step 2: decide a test statistic

" **what distribution will you use based on the above observations**?"

$$\text{A fact: } T = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} \sim t_{n-1},$$

  where μ_0 is general notation for the value of μ under the null hypothesis; in this example, $\mu_0 = 1.5$.
  Therefore, we use T-statistic: $T = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}}$. In this example, $t_{obs} = \frac{1.348 - 1.50}{\frac{0.3385}{\sqrt{16}}} = -1.796$.

T-test using the rejection region

- Step 3: decide the rejection region

T-test using the rejection region

- Step 3: decide the rejection region

  " **Use the distribution of the test statistic to determine a rejection region that limits the type 1 error at significance level** α"

T-test using the rejection region

- Step 3: decide the rejection region

  “ **Use the distribution of the test statistic to determine a rejection region that limits the type 1 error at significance level** α”

  - α is usually given before sampling.
  - Its common values: $5\%, 1\%, 10\%$.
  - In this example, we set $\alpha = 0.05$.

T-test using the rejection region

- Step 3: decide the rejection region

  " **Use the distribution of the test statistic to determine a rejection region that limits the type 1 error at significance level** α"

  - α is usually given before sampling.
  - Its common values: $5\%, 1\%, 10\%$.
  - In this example, we set $\alpha = 0.05$.

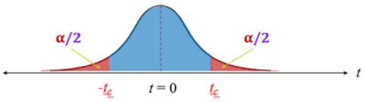
$$\alpha = P(\mathit{Reject}\ H_0 | H_0\ \mathit{true})$$

or

$$\alpha = P(\text{Test statistic falls in the rejection region } | H_0\ \mathit{true})$$

T-test using the rejection region

- Step 3: decide the rejection region

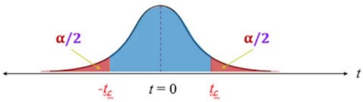


Since farther from 0 in either direction is more evidence against the null, we put half of α in the tails on either side of the distribution. Thus the **not-rejection region is defined by:**

$$P(-t_{(n-1,\alpha/2)} \leq T \leq t_{(n-1,\alpha/2)} \big| H_0 \; true) = 1 - \alpha.$$

T-test using the rejection region

- Step 3: decide the rejection region



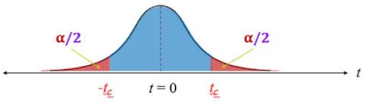
Since farther from 0 in either direction is more evidence against the null, we put half of α in the tails on either side of the distribution. Thus the **not-rejection region is defined by:**

$$P(-t_{(n-1,\alpha/2)} \leq T \leq t_{(n-1,\alpha/2)} \big| H_0 \; true) = 1 - \alpha.$$

Therefore, **the rejection region is** $T < -t_{n-1,\alpha/2}$, $T > t_{n-1,\alpha/2}$.

T-test using the rejection region

- Step 3: decide the rejection region



Since farther from 0 in either direction is more evidence against the null, we put half of α in the tails on either side of the distribution. Thus the **not-rejection region is defined by:**

$$P(-t_{(n-1,\alpha/2)} \leq T \leq t_{(n-1,\alpha/2)} \big| H_0 \textit{ true}) = 1 - \alpha.$$

Therefore, **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 $T > 2.13$.

Conclusion

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

Power computation

Suppose $\mu_A = 1.4$mm. (Notice the notation: μ_0 denotes the null value of μ, and μ_A denotes an alternative value of μ.)

$$Power = 1 - \beta = P(Reject\ H_0 | H_0\ false)$$

In this example,

$$Power = P\left(|T| > 2.13 \middle| \mu = \mu_A\right) = P\left(\left|\frac{\bar{X} - \mu_0}{S/\sqrt{n}}\right| > 2.13 \middle| \mu = \mu_A\right),$$

for $\mu_0 = 1.5$, $\mu_A = 1.4$, and $n = 16$.

Power computation

Suppose $\mu_A = 1.4$mm. (Notice the notation: μ_0 denotes the null value of μ, and μ_A denotes an alternative value of μ.)

$$Power = 1 - \beta = P(Reject\ H_0 | H_0\ false)$$

In this example,

$$Power = P\left(|T| > 2.13 \middle| \mu = \mu_A\right) = P\left(\left|\frac{\bar{X} - \mu_0}{S/\sqrt{n}}\right| > 2.13 \middle| \mu = \mu_A\right),$$

for $\mu_0 = 1.5$, $\mu_A = 1.4$, and $n = 16$.

Calculating this probability exactly is very challenging.

Power computation

We use an approximate method. The basic idea is to treat the sample standard deviation $s = 0.3385$ as fixed, and hence our standard error $s/\sqrt{16} = 0.3385/4 = 0.0846$ stays fixed. Next we re-write the rejection region in terms of $\bar{X}$.

One side of the rejection region is $T < -2.13$, so

$$\frac{\bar{X} - 1.5}{0.0846} < -2.13$$
$$\bar{X} < -2.13 * 0.0846 + 1.5 = 1.32$$

Similarly for the other side $T > 2.13$

$$\frac{\bar{X} - 1.5}{0.0846} > 2.13$$
$$\bar{X} > 0.0846 * 2.13 + 1.5 = 1.68$$

Power computation

Now the approximate power can be expressed as:

$$
\begin{aligned}
\text{Power}(\mu_A = 1.4) &= P(\bar{X} < 1.32 | \mu = 1.4) + P(\bar{X} > 1.68 | \mu = 1.4) \\
&= \mathrm{P}\left(\frac{\bar{X}-1.4}{0.0846} < \frac{1.32-1.4}{0.0846} \middle| \mu = 1.4\right) + P\left(\frac{\bar{X}-1.4}{0.0846} > \frac{1.68-1.4}{0.0846} \middle| \mu = 1.4\right) \\
&= \mathrm{P}(T_{15} < -0.95) + P\left(T_{15} > 3.31\right) \\
&= 0.179 + 0.002 \\
&= 0.181
\end{aligned}
$$

Power computation

Now the approximate power can be expressed as:

$$
\begin{aligned}
\text{Power}(\mu_A = 1.4) &= P(\bar{X} < 1.32 | \mu = 1.4) + P(\bar{X} > 1.68 | \mu = 1.4) \\
&= \text{P}\left(\frac{\bar{X} - 1.4}{0.0846} < \frac{1.32 - 1.4}{0.0846} \middle| \mu = 1.4 \right) + P\left(\frac{\bar{X} - 1.4}{0.0846} > \frac{1.68 - 1.4}{0.0846} \middle| \mu = 1.4 \right) \\
&= \text{P}(T_{15} < -0.95) + P\left(T_{15} > 3.31\right) \\
&= 0.179 + 0.002 \\
&= 0.181
\end{aligned}
$$

This power is low. This means that if the true population mean really was $\mu = 1.4$, we would be very unlikely to reject the null based on a sample of size 16.

Power computation

Now the approximate power can be expressed as:

$$\text{Power}(\mu_A = 1.4) = P(\bar{X} < 1.32|\mu = 1.4) + P(\bar{X} > 1.68|\mu = 1.4)$$
$$= \text{P}\left(\frac{\bar{X}-1.4}{0.0846} < \frac{1.32-1.4}{0.0846}\middle|\mu = 1.4\right) + P\left(\frac{\bar{X}-1.4}{0.0846} > \frac{1.68-1.4}{0.0846}\middle|\mu = 1.4\right)$$
$$= \text{P}(T_{15} < -0.95) + P\left(T_{15} > 3.31\right)$$
$$= 0.179 + 0.002$$
$$= 0.181$$

This power is low. This means that if the true population mean really was $\mu = 1.4$, we would be very unlikely to reject the null based on a sample of size 16.

Question: how can we increase the power?

Sample size computation

We can have the best of both worlds – controlling the probability of Type I and Type II error simultaneously (thereby controlling power) – provided we can collect the requisite number of samples. **What is the required sample size to achieve the power?**

Sample size computation

We can have the best of both worlds – controlling the probability of Type I and Type II error simultaneously (thereby controlling power) – provided we can collect the requisite number of samples. **What is the required sample size to achieve the power?**

It can be shown that if σ is known, the sample size n required to achieve power $1-\beta$ for a test of $H_0 : \mu = \mu_0$ vs. $H_A : \mu \neq \mu_0$ at a given alternative $\mu = \mu_A$ at level α is approximately:

$$n = \left(\frac{\sigma(z_{\alpha/2}+z_\beta)}{\mu_0-\mu_A}\right)^2,$$

where $z_{\alpha/2}$ and z_β are the $\alpha/2$ and β right-tailed critical values of the standard normal distribution.

Sample size computation

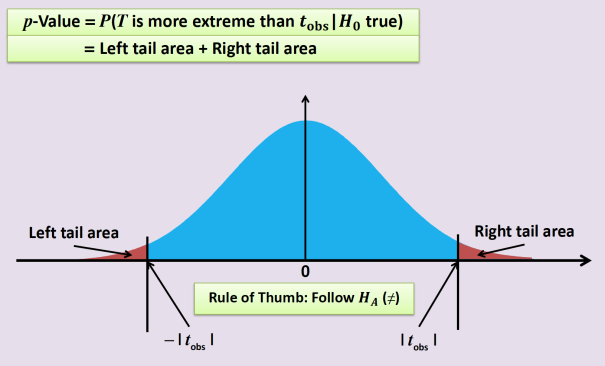
Back to the example. Now suppose we wanted to determine the sample size to have power 0.8 to reject when the true mean was $\mu_A = 1.4$. Thus our β value is 1-0.8 = 0.2. Use $\sigma = 0.3385$ as an estimator of σ. Using the standard normal table we have $z_{0.05/2} = 1.96$ and $z_{0.2} = 0.84$. By this formula we would need:

$$n = (\tfrac{0.3385(1.96+0.84)}{1.5-1.4})^2 = 89.8, \text{ round up to 90.}$$

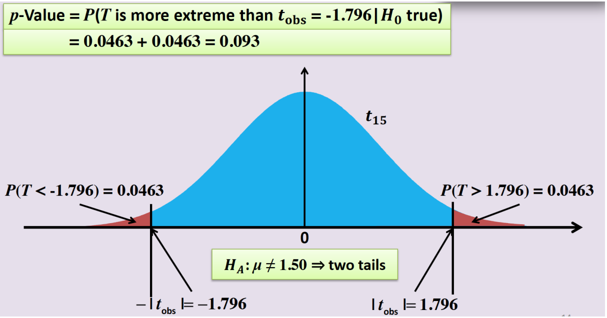
T-test using p-value

“ **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,



T-test using p-value



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-value is a measure of evidence in the sense that it is the smallest α value at which we'd just barely reject H_0!

Recap of the two-sided T-test

- We assume the data are realizations of n random variables $X_1, \ldots, X_n$ which are iid $\mathrm{N}(\mu, \sigma^2)$ where σ is unknown.
- We are testing $H_0 : \mu = \mu_0, H_A : \mu \neq \mu_0$.
- Let $T = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}}$. The two-sided T-test with significance level α has rejection region:

$$\text{Reject } H_0 \text{ when: } T < -t_{n-1,\alpha/2} \quad \text{or } T > t_{n-1,\alpha/2}.$$

- Let the observed (realized) T statistic be $t_{obs} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$. Then the p-value of the two-sided T-test is

$$\text{p-value} = P\left(T_{n-1} < -t_{obs}\right) + P\left(T_{n-1} > t_{obs}\right) = 2 * P\left(T_{n-1} > |t_{obs}|\right).$$

What's the next?

We'll discuss how to use bootstrap methods to make hypothesis testing in the next lecture.