Chapter 7: One sample tests

(Ott & Longnecker Sections: 5.4-5.7, 5.9)

Part 2

https://dzwang91.github.io/stat324/

WISCONSIN

UNIVERSITY OF WISCONSIN–MADISON

Overview

Population mean

$$H_0 : \mu = \mu_0 \text{ v.s. } H_A : \mu \neq \mu_0$$

$$H_0 : \mu = \mu_0 \text{ v.s. } H_A : \mu < \mu_0$$

$$H_0 : \mu = \mu_0 \text{ v.s. } H_A : \mu > \mu_0$$

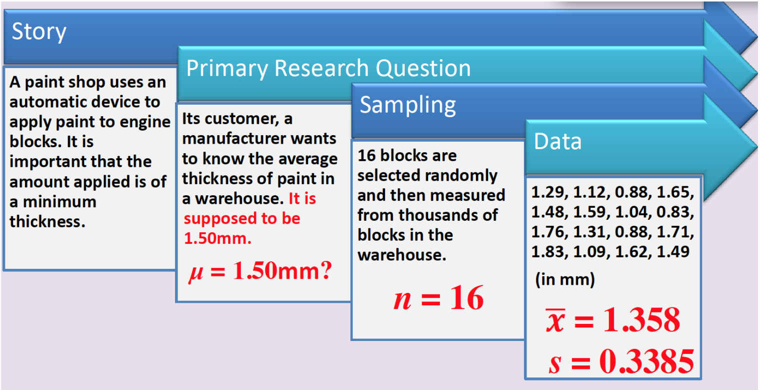
parametric tests

Sample mean test

Z test

T test

Example: two-tailed T test



Example: two-tailed T test

- Step 1:

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

- Step 2: Choose $\alpha = 0.05$
- Step 3: Rejection region is $T < -t_{15,0.025}$ or $T > t_{15,0.025}$. From t table, $t_{15,0.025} = 2.13$.
- Step 4: $t_{obs} = \frac{1.348 - 1.50}{\frac{0.3385}{\sqrt{16}}} = -1.796$. t_{obs} does not fall in the rejection region, so we do not reject the null.

Example: two-tailed T test

- Step 1:

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

- Step 2: Choose $\alpha = 0.05$
- Step 3: Rejection region is $T < -t_{15,0.025}$ or $T > t_{15,0.025}$. From t table, $t_{15,0.025} = 2.13$.
- Step 4: $t_{obs} = \frac{1.348-1.50}{\frac{0.3385}{\sqrt{16}}} = -1.796$. t_{obs} does not fall in the rejection region, so we do not reject the null.

Is this test statistically good?

Outline

1. Power calculation
2. Sample size calculation
3. Sign test

Power

Definition

The power of a test is the probability of rejecting H_0 given that the alternative hypothesis is true. That is

$$\text{Power} = 1 - \beta$$

where β is the Type II error.

- How can we calculate the power of a test?

Power

Definition

The power of a test is the probability of rejecting H_0 given that the alternative hypothesis is true. That is

$$\text{Power} = 1 - \beta$$

where β is the Type II error.

- How can we calculate the power of a test?

Definition

Simple hypothesis against a simple alternative:

$$H_0 : \mu = \mu_0 \text{ v.s. } H_A : \mu = \mu_1,$$

where μ_0 and μ_1 are two constants.

Vignette one: power for two-tailed T test

$$H_0 : \mu = 1.5 \text{ v.s. } H_A : \mu = 1.4$$

- Power:

$$\begin{aligned} P(\text{Reject } H_0 | H_A \text{ is true}) &= P\left(|T| > 2.13 \middle| \mu = 1.4\right) \\ &= P\left(\left|\frac{\bar{X} - 1.5}{S/\sqrt{n}}\right| > 2.13 \middle| \mu = 1.4\right) = ? \end{aligned}$$

Vignette one: power for two-tailed T test

$$H_0 : \mu = 1.5 \text{ v.s. } H_A : \mu = 1.4$$

- Power:

$$\begin{aligned} P(\text{Reject } H_0 | H_A \text{ is true}) &= P\left(|T| > 2.13 \middle| \mu = 1.4\right) \\ &= P\left(\left|\frac{\bar{X} - 1.5}{S/\sqrt{n}}\right| > 2.13 \middle| \mu = 1.4\right) = ? \end{aligned}$$

- What is the distribution of $\frac{\bar{X}-1.5}{S/\sqrt{n}}$ when $\mu = 1.4$?

Vignette one: power for two-tailed T test

$$H_0 : \mu = 1.5 \text{ v.s. } H_A : \mu = 1.4$$

- Power:

$$\begin{aligned} P(\text{Reject } H_0 | H_A \text{ is true}) &= P\left(|T| > 2.13 \middle| \mu = 1.4\right) \\ &= P\left(\left|\frac{\bar{X} - 1.5}{S/\sqrt{n}}\right| > 2.13 \middle| \mu = 1.4\right) =? \end{aligned}$$

- What is the distribution of $\frac{\bar{X}-1.5}{S/\sqrt{n}}$ when $\mu = 1.4$? NOT t distribution.

Power calculation continued

- Decomposition:

$$\frac{\bar{X} - 1.5}{S/\sqrt{n}} = \frac{\bar{X} - 1.4}{S/\sqrt{n}} + \frac{1.4 - 1.5}{S/\sqrt{n}}$$

Under $\mu = 1.4$, $\frac{\bar{X}-1.4}{S/\sqrt{n}}$ is t_{n-1} distribution.

Power calculation continued

- Decomposition:

$$\frac{\bar{X}-1.5}{S/\sqrt{n}} = \frac{\bar{X}-1.4}{S/\sqrt{n}} + \frac{1.4-1.5}{S/\sqrt{n}}$$

Under $\mu = 1.4$, $\frac{\bar{X}-1.4}{S/\sqrt{n}}$ is t_{n-1} distribution.

$$Power = P\left(\left|\frac{\bar{X}-1.4}{S/\sqrt{n}} + \frac{1.4-1.5}{S/\sqrt{n}}\right| > 2.13 \middle| \mu = 1.4\right)$$

$$= P\left(\left|t_{15} + \frac{1.4-1.5}{0.3385/\sqrt{16}}\right| > 2.13\right) = P(|t_{15} - 1.18| > 2.13)$$

$$= P\left(T_{15} < -0.95\right) + P\left(T_{15} > 3.31\right) = 0.179 + 0.002 = 0.181$$

Power calculation continued

- Decomposition:

$$\frac{\bar{X} - 1.5}{S/\sqrt{n}} = \frac{\bar{X} - 1.4}{S/\sqrt{n}} + \frac{1.4 - 1.5}{S/\sqrt{n}}$$

Under $\mu = 1.4$, $\frac{\bar{X}-1.4}{S/\sqrt{n}}$ is t_{n-1} distribution.

$$Power = P(|\frac{\bar{X} - 1.4}{S/\sqrt{n}} + \frac{1.4 - 1.5}{S/\sqrt{n}}| > 2.13|\mu = 1.4)$$
$$= P(|t_{15} + \frac{1.4 - 1.5}{0.3385/\sqrt{16}}| > 2.13) = P(|t_{15} - 1.18| > 2.13)$$
$$= P\left(T_{15} < -0.95\right) + P\left(T_{15} > 3.31\right) = 0.179 + 0.002 = 0.181$$

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

How can we increase power?

How can we increase power?

What is the required sample size to achieve specific power?

Outline

Vignette two: sample size calculation for two-tailed Z test

- Two-tailed Z test:
  - $H_0 : \mu = \mu_0$ v.s. $H_A : \mu \neq \mu_0$,
  - Test statistic: $Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$
  - Rejection region to control Type I error at significance level α:

$$Z > z_{\alpha/2} \text{ or } Z < -z_{\alpha/2}.$$

Vignette two: sample size calculation for two-tailed Z test

- Two-tailed Z test:
  - $H_0 : \mu = \mu_0$ v.s. $H_A : \mu \neq \mu_0$,
  - Test statistic: $Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$
  - Rejection region to control Type I error at significance level α:

$$Z > z_{\alpha/2} \text{ or } Z < -z_{\alpha/2}.$$

- Goal: find the required sample size to achieve power at least $1 - \beta$ for the specific simple alternative $H_A : \mu = \mu_1$

Vignette two: sample size calculation for two-tailed Z test

- Two-tailed Z test:
  - $H_0 : \mu = \mu_0$ v.s. $H_A : \mu \neq \mu_0$,
  - Test statistic: $Z = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$
  - Rejection region to control Type I error at significance level α:

$$Z > z_{\alpha/2} \text{ or } Z < -z_{\alpha/2}.$$

- Goal: find the required sample size to achieve power at least $1-\beta$ for the specific simple alternative $H_A : \mu = \mu_1$

$\Leftrightarrow$

$$\text{What is } n \text{ which satisfies } P\left(\left|\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right| < z_{\alpha/2} \middle| \mu = \mu_1\right) \leq \beta?$$

Vignette two: sample size calculation for two-tailed Z test

- Derivation:

$$P(|\frac{\bar{X}-\mu_1}{\sigma/\sqrt{n}} + \frac{\mu_1-\mu_0}{\sigma/\sqrt{n}}| < z_{\alpha/2}|\mu=\mu_1) \leq \beta$$

$$P(|Z + \frac{\mu_1-\mu_0}{\sigma/\sqrt{n}}| < z_{\alpha/2}) \leq \beta$$

$$P(-z_{\alpha/2} < Z + \frac{\mu_1-\mu_0}{\sigma/\sqrt{n}} < z_{\alpha/2}) \leq \beta$$

$$P(-z_{\alpha/2} - \frac{\mu_1-\mu_0}{\sigma/\sqrt{n}} < Z < z_{\alpha/2} - \frac{\mu_1-\mu_0}{\sigma/\sqrt{n}}) \leq \beta \qquad (\star)$$

Vignette two: sample size calculation for two-tailed Z test

- In order to make $(\star)$ hold, only need

$$P(Z < z_{\alpha/2} - \frac{\mu_1 - \mu_0}{\sigma/\sqrt{n}}) = \beta$$

Vignette two: sample size calculation for two-tailed Z test

- In order to make $(\star)$ hold, only need

$$P(Z < z_{\alpha/2} - \frac{\mu_1 - \mu_0}{\sigma/\sqrt{n}}) = \beta$$

$$z_{\alpha/2} - \frac{\mu_1 - \mu_0}{\sigma/\sqrt{n}} = -z_\beta$$

Vignette two: sample size calculation for two-tailed Z test

- In order to make $(\star)$ hold, only need

$$P(Z < z_{\alpha/2} - \frac{\mu_1 - \mu_0}{\sigma/\sqrt{n}}) = \beta$$

$$z_{\alpha/2} - \frac{\mu_1 - \mu_0}{\sigma/\sqrt{n}} = -z_\beta$$

$$n = \left(\frac{\sigma(z_{\alpha/2} + z_\beta)}{\mu_1 - \mu_0}\right)^2$$

Vignette two: sample size calculation for two-tailed Z test

- In order to make $(\star)$ hold, only need

$$P(Z < z_{\alpha/2} - \frac{\mu_1 - \mu_0}{\sigma/\sqrt{n}}) = \beta$$

$$z_{\alpha/2} - \frac{\mu_1 - \mu_0}{\sigma/\sqrt{n}} = -z_\beta$$

$$n = \left(\frac{\sigma\big(z_{\alpha/2} + z_\beta\big)}{\mu_1 - \mu_0}\right)^2$$

- How does n change when α and β become small?

Vignette two: sample size calculation for two-tailed Z test

- In order to make $(\star)$ hold, only need

$$P(Z < z_{\alpha/2} - \frac{\mu_1 - \mu_0}{\sigma/\sqrt{n}}) = \beta$$

$$z_{\alpha/2} - \frac{\mu_1 - \mu_0}{\sigma/\sqrt{n}} = -z_\beta$$

$$n = \left(\frac{\sigma\left(z_{\alpha/2} + z_\beta\right)}{\mu_1 - \mu_0}\right)^2$$

- How does n change when α and β become small?
- How does n change when the difference between μ_1 and μ_0 becomes small?

Back to example

- Determine the required sample size to have power 0.8 under $H_A : \mu = 1.4$

Back to example

- Determine the required sample size to have power 0.8 under $H_A : \mu = 1.4$
- We have $\beta = 1 - 0.8 = 0.2$, $z_{0.05/2} = 1.96$ and $z_{0.2} = 0.84$. Use $S = 0.3385$ as an estimator of σ. Then
$$n = \left(\frac{0.3385(1.96+0.84)}{1.5-1.4}\right)^2 = 89.8, \text{ round up to } 90.$$

Take-home exercise

Find the required sample size for right-tailed Z test which achieves power at least $1-\beta$ at significance level α for the specific simple alternative

$$H_A : \mu = \mu_1$$

Outline

Sign test

Definition

For a random variable X, population median is defined as a parameter m such that

$$P(X \leq m) = \frac{1}{2}$$

Sign test

Definition

For a random variable X, population median is defined as a parameter m such that

$$P(X \leq m) = \frac{1}{2}$$

- Collect an i.i.d. sample $X_1, \ldots, X_n$ from some population with population median m.
- **Step 1**:
  1. Two-tailed test:
  $$H_0 : m = m_0 \text{ v.s. } m \neq m_0$$
  2. Left-tailed test:
  $$H_0 : m = m_0 \text{ v.s. } m < m_0$$
  3. Right-tailed test:
  $$H_0 : m = m_0 \text{ v.s. } m > m_0$$

Sign test

- **Step 2**: delete data points which are equal to m_0. Count remaining sample size n^*.

Sign test

- **Step 2**: delete data points which are equal to m_0. Count remaining sample size n^*.
- **Step 3**: choose test statistic

  $B =$ number of data values greater than m_0 in the remaining sample.

Sign test

- **Step 2**: delete data points which are equal to m_0. Count remaining sample size n^*.
- **Step 3**: choose test statistic

  $B =$ number of data values greater than m_0 in the remaining sample.

- Under $H_0 : m = m_0$, $B \sim \text{Binomial}(n^*, 0.5)$.

Sign test

- **Step 2**: delete data points which are equal to m_0. Count remaining sample size n^*.
- **Step 3**: choose test statistic

  $B =$ number of data values greater than m_0 in the remaining sample.

- Under $H_0 : m = m_0$, $B \sim \text{Binomial}(n^*, 0.5)$.
- **Step 4**: calculate p-value. Let b be the observation of B.
  - $H_A : m > m_0$:
    $$\text{p-value} = P(B \geq b \mid B \sim \text{Binomial}(n^*, 0.5))$$
  - $H_A : m < m_0$:
    $$\text{p-value} = P(B \leq b \mid B \sim \text{Binomial}(n^*, 0.5))$$
  - $H_A : m \neq m_0$:
    $$\text{p-value} = 2 \times \min\{P(B \geq b), P(B \leq b)\}.$$

Example

The sanitation department in a large city is considering separating recyclable material out of the trash to save on landfill space and make money selling the recyclables. Based on data from other cities, it is determined that if at least half of the households in the city have more than 4.6 lbs of reclaimable recyclable material, then the separation will be profitable for the city. A random sample of 11 households yields the following data on pounds of recyclable material found in the trash:

14.2, 5.3, 2.9, 4.2, 1.2, 4.3, 1.1, 2.6, 6.7, 7.8, 25.9

Note that the median of the sample is 4.3 lbs.

Example continued

- $H_0 : m = 4.6$ vs. $H_A : m > 4.6$.
- test statistic: B=the number of observations in the sample that are greater than 4.6
- Under H_0, $B \sim Bin(11, 0.5)$.
- Observation of B:

  14.2, 5.3, 2.9, 4.2, 1.2, 4.3, 1.1, 2.6, 6.7, 7.8, 25.9

  so $B_{obs} = 5$.
- p-value$= P(B \geq 5) = 0.726$.
- Set $\alpha = 0.05$, then there is no sufficient evidence to reject the null.
- There is no strong evidence that separation of the recyclables would be profitable in this city.

Tests for symmetric distributions

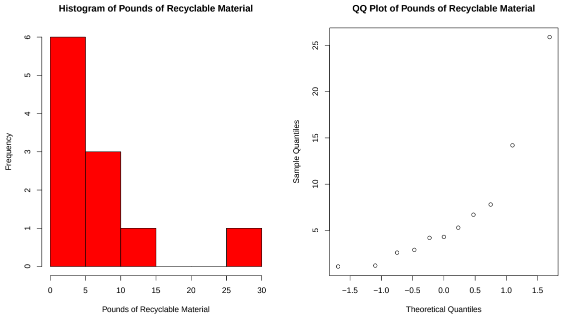
- For symmetric distributions, population median is equal to population mean, so T test and sign test can be used to test both of population mean and population median.

Tests for symmetric distributions

- For symmetric distributions, population median is equal to population mean, so T test and sign test can be used to test both of population mean and population median.
- Comparison between T test and sign test: T test requires the population distribution is normal, but sign test doesn't

Back to example

- Histogram and QQ plot:



The histogram looks neither normal nor symmetric, and the QQ plot also does not support normality. Therefore, we can not use T test.

What's the next?

We'll discuss how to test population proportions in next lecture.