Hypothesis Testing

Null

Alternative

P-value

Example: [Concrete beams]

10 concrete beams were each measured for flexural strength (MPa). The data is as follows.

$n=10$ → [1] 8.2 8.7 7.8 9.7 7.4 7.8 7.7 11.6 11.3 11.8

The sample mean was $\overset{\bar{x}}{9.2}$ MPa and the sample variance s^2 was 3.0933 MPa. Conduct a hypothesis test to find out if the flexural strength is different from 9.0 MPa. at $\alpha = 0.01$ level.

1. $H_0: \mu = \underset{\mu_0}{9}$ vs. $H_a: \mu \neq 9$

2. $\alpha = 0.01$

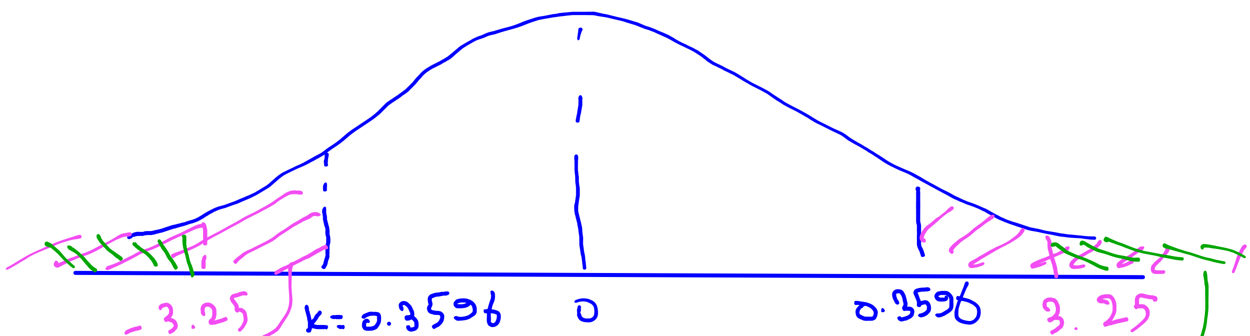
3. $n = 10\ (<25)$ + σ is unknown. So,

$$K = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} \quad \left(\text{we know that } K \sim t_{n-1}\right)$$

assumption for t-student also: $X_1, \ldots, X_{10} \overset{iid}{\sim} N(\mu, \sigma^2)$

$$4.\ K = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{9.2 - 9}{\sqrt{\frac{3.0933}{10}}} = 0.3596$$

$$\text{p-value} = P(|T| > K) = P(|T| > 0.3596)$$



$$t_{9,\,1-\alpha/2} = t_{9,\,0.995} = 3.25$$

by the t table

green shaded area

p-value

pink shaded area

by t-table $= \underbrace{P(|T| > 0.3596)}_{\text{p-value}} > \underbrace{P(|T| > t_{9,\,0.995})}_{\alpha = 0.01}$

5. Since p-value is $> \alpha = 0.01$, we fail to reject to null Hypothesis (H_0)

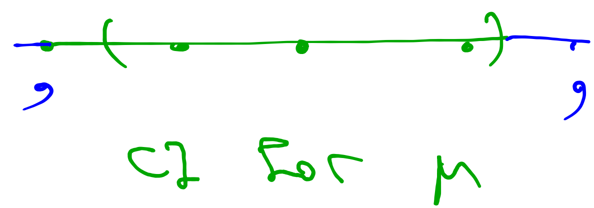
6. There is NOT enough evidence to conclude that the true mean flexural strength of the beams is different from 9 MPa.

Hypothesis Testing Using Confidence Interval

Hypothesis testing using the CI

We can also use the $1 - \alpha$ confidence interval to perform hypothesis tests (instead of p-values). The confidence interval will contain μ_0 when there is little to no evidence against H_0 and will not contain μ_0 when there is strong evidence against H_0.

$$\begin{cases} H_0: \mu = 9 \\ H_a: \mu \neq 9 \end{cases}$$



Hypothesis Testing

Null

Alternative

P-value

CI method

Hypothesis testing using the CI

Steps to perform a hypothesis test using a confidence interval:

$H_0: \mu = \mu_0$

$H_a: \begin{cases} \mu \neq \mu_0 \\ \mu > \mu_0 \\ \mu < \mu_0 \end{cases}$

1. State H_0 and H_1
2. State α, significance level
3. State the form of $100\,(1-\alpha)$ % CI along with all assumptions necessary. (use one-sided CI for one-sided tests and two-sided CI for two sided tests)
4. Calculate the CI
5. Based on $100\,(1-\alpha)$ % CI, either reject H_0 (if μ_0 is not in the interval) or fail to reject (if μ_0 is in the interval)
6. Interpret the conclusion in the content of the problem

confidence level.

Hypothesis Testing

Null

Alternative

P-value

CI method

Example:[Breaking strength of wire, cont'd]

Suppose you are a manufacturer of construction equipment. You make 0.0125 inch wire rope and need to determine how much weight it can hold before breaking so that you can label it clearly. You have breaking strengths, in kg, for 41 sample wires with sample mean breaking strength 91.85 kg and sample standard deviation 17.6 kg. Using the appropriate 95% confidence interval, conduct a hypothesis test to find out if the true mean breaking strength is above 85 kg.

n=41 >25 CLT

Steps:

1- $H_0: \quad \mu = 85$ *vs.* $H_1: \quad \mu > 85$

2- $\alpha = 0.05$

$100(1-\alpha)\% = 95\%$ => $\alpha = 0.05$ Significant level,

Hypothesis Testing

Null

Alternative

P-value

CI method

based on H_a

Example:[Breaking strength of wire, cont'd]

3- One-sided test and we care about the lower bound. So, we use $(\overline{X} - z_{1-\alpha}\frac{s}{\sqrt{n}}, +\infty)$.

4- From the example in previous set of slides, the CI is $(87.3422, +\infty)$.

$H_0: 85$ 87.3422

5- Since $\mu_0 = 85$ is not in the CI, we **reject** H_0.

6- There is significant evidence to conclude that the true mean breaking strength of wire is greater than the 85kg. Hence the requirement is met.

$H_a: \mu > 85$

Hypothesis Testing

Null

Alternative

P-value

CI method

Example: [Concrete beams, cont'd]

10 concrete beams were each measured for flexural strength (MPa). The data is as follows.

$n = 10$

[1] 8.2 8.7 7.8 9.7 7.4 7.8 7.7 11.6 11.3 11.8 $\rightarrow \bar{x} = \frac{1}{10}\sum_{i=1}^{10} x_i = 9.2$

The sample mean was $\bar{X}$ 9.2 MPa and the sample variance $s^2 =$ was 3.0933 $(MPa)^2$. At $\alpha = 0.01$, test the hypothesis that the true mean flexural strength is 10 MPa using a confidence interval. Steps:

write down the hypothesis $\rightarrow$

1- $H_0: \quad \mu = 10 \quad vs. \quad H_1: \quad \mu \neq 10$

2- $\alpha = 0.01$

$\rightarrow$ 3- This is two-sided test with $n = 10$ and 100 $(1-\alpha)$ % CI is

$n = 10$
σ unknown $\rightarrow$

$$\left(\bar{X} - t_{(n-1,1-\alpha/2)}\frac{s}{\sqrt{n}}, \bar{X} + t_{(n-1,1-\alpha/2)}\frac{s}{\sqrt{n}}\right)$$

two-sided test $\equiv$ two-sided CI

one-sided test $\equiv$ one-sided CI

$t_{9, 0.995} = 3.25$

Hypothesis Testing

Null

Alternative

P-value

CI method

Example:[Breaking strength of wire, cont'd]

4- Check that the CI is $(7.393, 11.007)$.

5- Since $\mu_0 = 10$ is within the CI, we **fail** to reject H_0.

6- There is not enough evidence to conclude that the true mean flexural strength is different from 10 Mpa.

Example:[Paint thickness, cont'd]

Consider the following sample of observations on coating thickness for low-viscosity paint.

[1] 0.83 0.88 0.88 1.04 1.09 1.12 1.29 1.31 1.48 1.49 1.59 1.62 1.65 1.71 ~~[15]~~ 1.76 1.83

$n=16$

Using $\boxed{\alpha = 0.1}$, test the hypothesis that the true mean paint thickness is 1.00 mm. Note, the 90\% confidence interval for the true mean paint thickness was calculated from before as $(1.201, 1.499)$.

1- $H_0: \quad \mu = 1 \; vs. \quad H_1: \quad \mu \neq 1$

2- $\alpha = 0.1$

3- This is two-sided test with $n = 16$, σ unknown, so $100\,(1-\alpha)$ % CI is

$$(\overline{X} - t_{(n-1,1-\alpha/2)}\frac{s}{\sqrt{n}}, \overline{X} + t_{(n-1,1-\alpha/2)}\frac{s}{\sqrt{n}})$$

$t_{15,\,1-\frac{0.1}{2}} = t_{15,\,.95}$

$n = 16 \,(< 25)$

+

σ unknown

=> use t dist.

Hypothesis Testing

Null

Alternative

P-value

CI method

Example:[Breaking strength of wire, cont'd]

> 4- The CI is $(1.201, 1.499)$.
>
> 5- Since $\mu_0 = 1$ is not in the the CI, we **reject** H_0.
>
> H_0
>
> 6- There is enough evidence to conclude that the true mean paint thickness is not 1mm.

$H_a: \mu \neq 1$

Section 6.4

Inference for matched pairs and two-sample data

Inference for matched pairs and two-sample data

An important type of application of confidence interval estimation and significance testing is when we either have *paired data* or *two-sample* data.

Recall: Matched pairs

Paired data is bivariate responses that consists of several determinations of basically the same characteristics

> **Example:**
>
> - Practice SAT scores *before* and *after* a preperation course
> - Severity of a disease *before* and *after* a treatment
> - Fuel economy of cars *before* and *after* testing new formulations of gasoline

Inference for matched pairs and two-sample data

One simple method of investigating the possibility of a consistent difference between paired data is to

> 1. Reduce the measurements on each object to a single difference between them
>
> 2. Methods of confidence interval estimation and significance testing applied to differences (using Normal or t distributions when appropriate)

Hypothesis Testing

Null

Alternative

P-value

CI method

Matched Pairs

Two-sample

paired data

Example:[Fuel economy]

Twelve cars were equipped with radial tires and driven over a test course. Then **the same twelve cars** (with the same drivers) were equipped with regular belted tires and driven over the same course.

① After each run, the cars gas economy (in km/l) was measured. Using significance level $\alpha = 0.05$ and the method of critical values, test for a difference in fuel economy between the radial tires and belted tires.

② Construct a 95% confidence interval for true mean difference due to tire type. (i.e μ_d)

car	1.0	2.0	3.0	4.0	5.0	6.0	7.0	8.0	9.0	10.0	11.0	12.0
① radial	4.2	4.7	6.6	7.0	6.7	4.5	5.7	6.0	7.4	4.9	6.1	5.2
② belted	4.1	4.9	6.2	6.9	6.8	4.4	5.7	5.8	6.9	4.7	6.0	4.9

$\mu_d \to$ difference.

$\therefore d = \text{radial}_1 - \text{belted}_1$

Example:[Fuel economy]

car	1.0	2.0	3.0	4.0	5.0	6.0	7.0	8.0	9.0	10.0	11.0	12.0
d_1: radial	4.2	4.7	6.6	7.0	6.7	4.5	5.7	6.0	7.4	4.9	6.1	5.2
d_2: belted	4.1	4.9	6.2	6.9	6.8	4.4	5.7	5.8	6.9	4.7	6.0	4.9
d	0.1	-0.2	0.4	0.1	-0.1	0.1	0.0	0.2	0.5	0.2	0.1	0.3

Since we have paired data, the first thing to do is to find the differences of the paired data. ($d = d_1 - d_2$, where d_1 is associated with radial and d_2 is associated with belted tires.)

Then writing down the information available:

$$n = 12, \quad \bar{d} = 0.142, \quad s_d = 0.198$$

$$\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i, \quad s_d^2 = \frac{1}{n-1}\sum_{i=1}^{n}(d_i - \bar{d})^2$$

Then we just need to apply steps of hypothesis testing. Note that the null hypothesis here is that **there is no difference** between the gas economy recorded of the two different tires.(i.e $\mu_d = 0$)

$H_0: \mu_d = 0$

Hypothesis Testing

Null

Alternative

P-value

CI method

Matched Pairs

Two-sample

Example:[Fuel economy]

1- $H_0: \quad \mu_d = 0 \; vs. \quad H_1: \quad \mu_d \neq 0$

2- $\alpha = 0.05$

μ_d

3- I will use the test statistics $K = \frac{\bar{d}-0}{s_d/\sqrt{n}}$ which has a $\boxed{t_{n-1}}$ distribution assuming that

- H_0 is true and
- $d_1, d_2, \cdots, d_{12}$ are iid $N(\mu_d, \sigma_d^2)$

n = 12
+
σ unknown
⇓
t-dist.

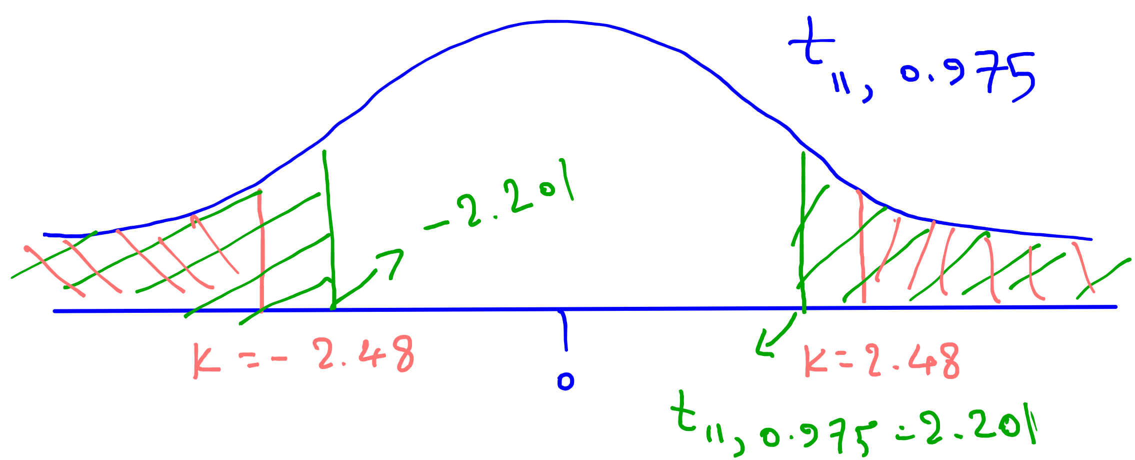
Example:[Breaking strength of wire, cont'd]

4- $K = \dfrac{0.421}{0.198/\sqrt{12}} = \boxed{2.48} \sim t_{(11,0.975)}$.

$$\begin{aligned}
p-value &= \boxed{P(|T|>K)} = P(|T|>2.48) \\
&= P(T>2.48) + P(T<-2.48) \\
\text{Software} \quad &= 1 - P(T<2.48) + P(T<-2.48) \\
(\text{by } \cancel{\text{the t table}}) \quad &= 1 - 0.9847 + 0.9694 = \boxed{0.03}
\end{aligned}$$

5- Since p-value < 0.05, we **reject** H_0.

6- There is enough evidence to conclude that fuel economy differs between radial and belted tires.



$$\underbrace{P(|T| > 2.48)}_{p\text{-value}} < \underbrace{P(|T| > t_{11, 0.975})}_{\alpha}$$