Hang on for a Second

Let's review slide 58 again

Hypothesis Testing

Null

Alternative

P-value

CI method

Matched Pairs

Two-sample

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

$$\begin{aligned} p - value &= P(|T| > K) = P(|T| > 2.48) \\ &= P(T > 2.48) + P(T < -2.48) \\ &= 1 - P(T < 2.48) + P(T < -2.48) \\ (\text{by software}) &= 1 - 0.9847 + 0.9694 = 0.03 \end{aligned}$$

We have seen t-student table

How do we get that p-value usin software!!!

What is happening?

Hypothesis Testing

Null

Alternative

P-value

CI method

Matched Pairs

Two-sample

Unlike *standard Normal distribution* table which gives us probability under the standard Normal curve, t tables are quantile tables.

i.e We use the t table (Table B.4 in Vardeman and Jobe) to calculate quantiles.

To have exact probabilities, we need software.

Table B.4
t Distribution Quantiles

ν	$Q(.9)$	$Q(.95)$	$Q(.975)$	$Q(.99)$	$Q(.995)$	$Q(.999)$	$Q(.9995)$
1	3.078	6.314	12.706	31.821	63.657	318.317	636.607
2	1.886	2.920	4.303	6.965	9.925	22.327	31.598
3	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869

e.g. we use these quantiles to make confidence intervals.

$$\bar{x} \pm \boxed{t_{(n-1,\, 1-\alpha/2)}} \cdot s/\sqrt{n}$$

a quantile, not probability.

The approach in calculating p-value when t distribution is involved

Hypothesis Testing

Null

Alternative

P-value

CI method

Matched Pairs

Two-sample

Two important points:

> P-value and α are both probabilities. (so $\in [0, 1]$).
>
> They are areas under the curve in tails under *null hypothesis*.

H_0

Hypothesis Testing

Null

Alternative

P-value

CI method

Matched Pairs

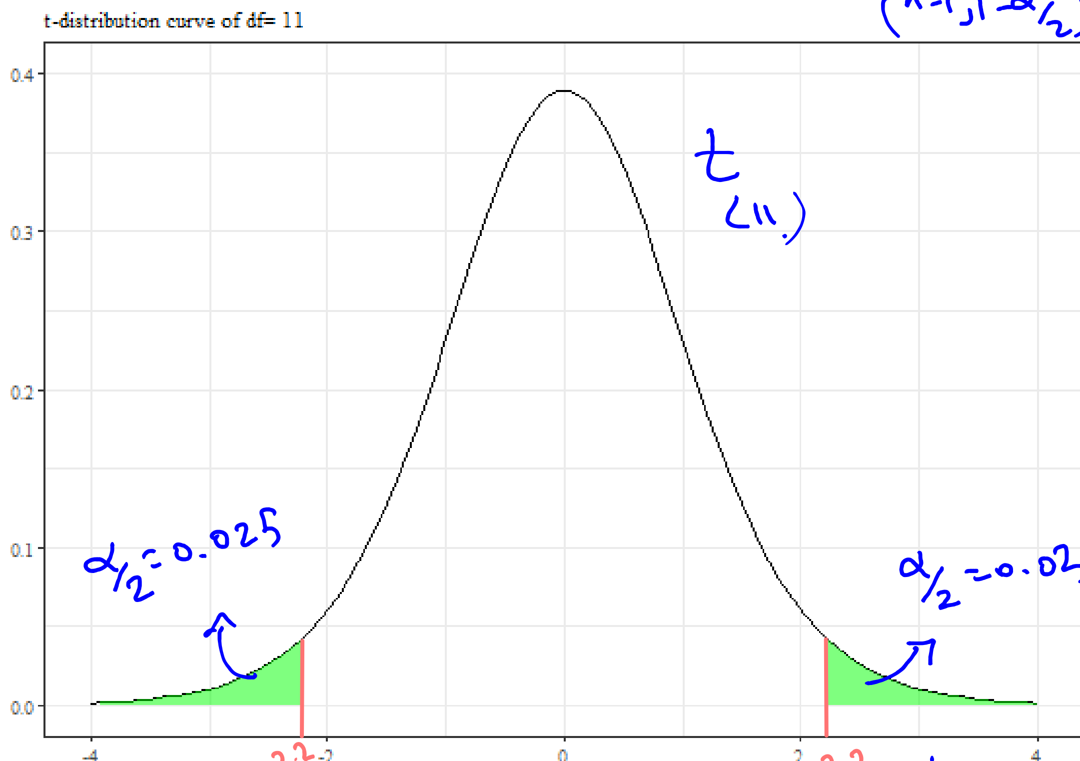
Two-sample

from example [breaking strength]

For a random variable with $\sim t_{(11,0.975)}$:

By the t table, the t quantile of $t_{(11,0.975)}$ is 2.2.

$t_{(n-1,1-\alpha/2)}$



$-2.2 \rightarrow -t_{(11,0.975)}$ $\quad$ $2.2 = t_{(11,0.975)}$

Total shaded area $= \alpha/2 + \alpha/2 = \alpha = 0.05$ (The significance level)

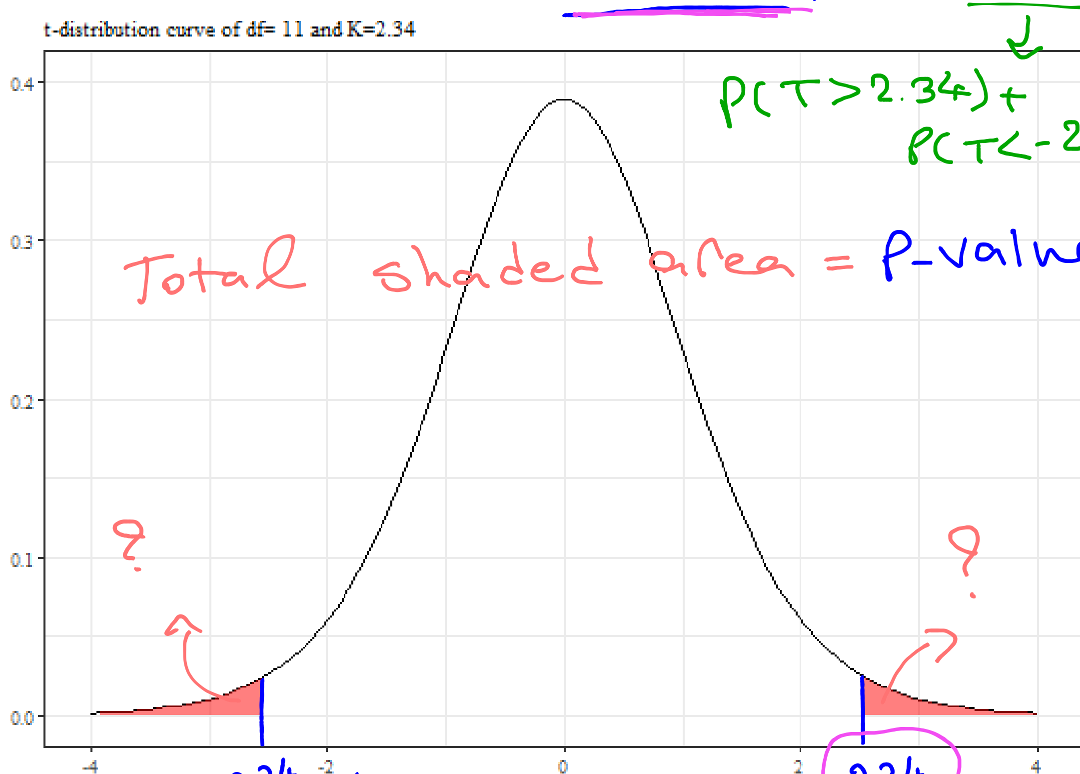
For the critical value we calculated under the null hypothesis:

$T \sim t_{(n-1)}$

The critical value calculated is $K = 2.34$

Recall: P-value $= P(|T| > K) = P(|T| > 2.34)$

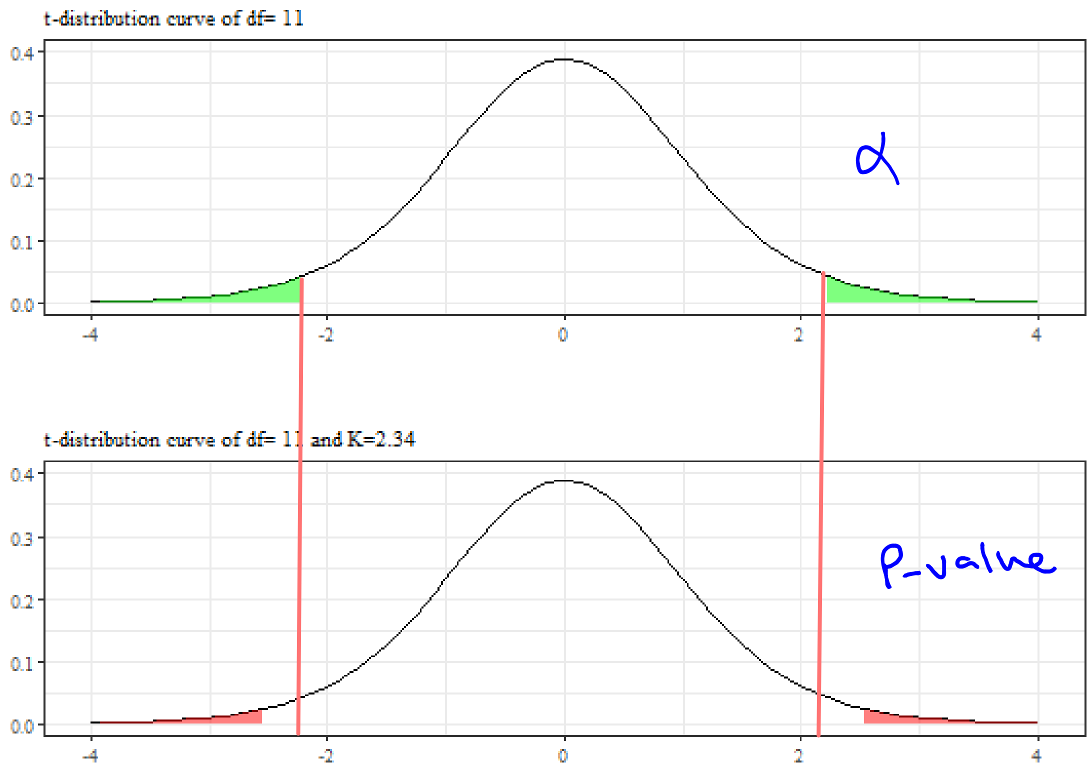
$P(T > 2.34) + P(T < -2.34)$



Here, we need software to find the total shaded area!

Both together

p-value ← $K \gtrless t_{(n-1,\, 1-\alpha/2)}$ → α



We reject the null if p-value < α.

> Remember p-value and α are areas under the curve

p-value < α

* Total green area : $\alpha = 0.05$

* Total red area : $P(|T| > k)$ = p-value = ?

(we don't need to find ? just need

→ to know if $? > \alpha$ or $? < \alpha$)

① If the red area (p-value) $< \alpha$ (= 0.05 in this problem)

⇒ Reject H_0.

② If the red area (p-value) $> \alpha$ (= 0.05 in this problem)

⇒ Fail to Reject H_0

The steps for p-value :

calculate p-value using table on slide 39.

1 - If you use Z => use Normal table to find p-value

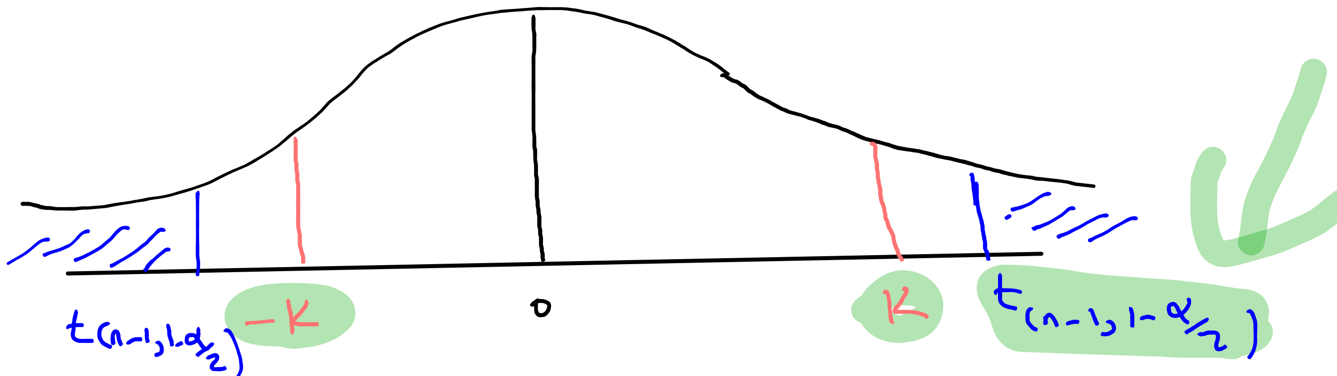
=> Then compare p-value & α (given in the problem) to reject or fail to reject H_0

2 - If you use T statistic, find the quantiles

$\rightarrow t_{(n-1,\,1-\alpha/2)}$ (or $t_{(n-1,\,1-\alpha)}$ depending on the problem).

$\nearrow$

(I suggest quickly plot $t_{(n-1,\,1-\alpha/2)}$ to understand better)



- write values of $\underline{\underline{K}}$ (critical value) and the corresponding $t_{(n-1,\,1-\alpha/2)}$

$H_0: \mu = \#$
$H_a: \mu \neq \#$

$$K = \frac{\bar{x} - \#}{s/\sqrt{n}}$$

- Note: area under the curve corresponding to $t_{(n-1,\,1-\alpha/2)}$ is equal to α. (e.g 0.05)

· Note: area under the curve corresponding to k (critical value) is p-value (which we don't need the exact value)

- Now compare the areas under the curve.

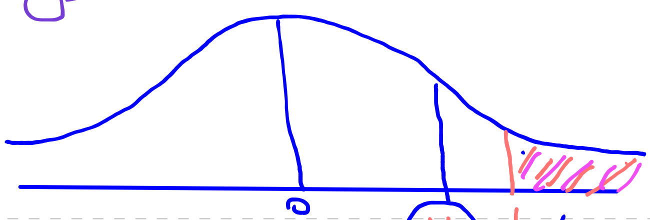
if p-value $< \alpha$ => reject H_0

if p-value $> \alpha$ => fail to reject H_0

Note Note Note:

* α is NOT always 0.05. (be careful)

* The test is NOT always two sided!



P-value $\rightarrow$ (K) $t_{(n-1,\ 1-\alpha)}$

Hypothesis Testing

Null

Alternative

P-value

CI method

Matched Pairs

Two-sample

Example:[End-cut router]

Consider the operation of an end-cut router in the manufacture of a company's wood product. Both a leading-edge and a trailing-edge measurement were made on each wooden piece to come off the router.

Paired data

Is the leading-edge measurement different from the trailing-edge measurement for a typical wood piece?

(Let's see if there's any difference between the measurement

Do a hypothesis test at $\alpha = 0.05$ to find out. Make a two-sided 95% confidence interval for the true mean of the difference between the measurements.

piece	1.000	2.000	3.000	4.000	5.000
leading_edge	0.168	0.170	0.165	0.165	0.170
trailing_edge	0.169	0.168	0.168	0.168	0.169
Difference: d_i	-.001	.002	-.003	-.003	.001

(–)

$n = 5$ $\quad \bar{d} = \frac{1}{5}(-.001 + .002 - .003 - 0.003 + .001)$, $S_d = \frac{1}{n-1}\sum_{i=1}^{5}(d_i - \bar{d})^2$

$= -8 \times 10^{-4}$ $\qquad = 0.0023$

Steps:

① $H_0: \mu_d = 0$ (There's no difference between measurements)
$H_a: \mu_d \neq 0$ (There's difference between measurements)

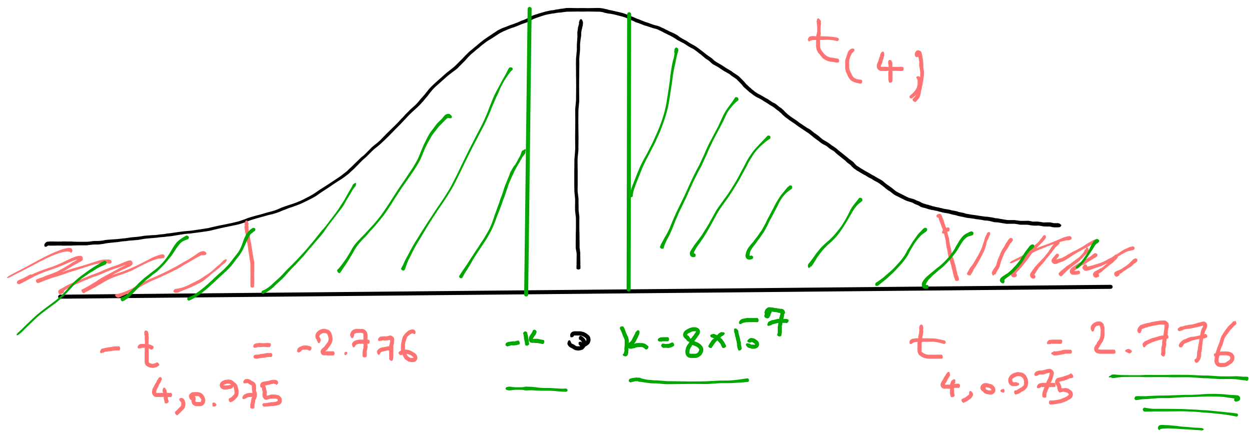
② $\alpha = 0.05$

③ since σ_d is unknown & $n = 5 < 25$, I will use

④ $$K = \frac{\mu_d - 0}{\frac{s_d}{\sqrt{n}}} = \frac{-8\times10^{-4} - 0}{\frac{0.0023}{\sqrt{5}}} = \boxed{-8\times10^{-7}}$$

$$\text{p-value} = P(|T| > |K|) = P(|T| > |-8\times10^{-7}|)$$
$$= P(|T| > 8\times10^{-7})$$

by the table: $t_{(n-1,\, 1-\alpha/2)} = t_{(5-1,\, 0.975)} = 2.776$



(obviously area under the curve corresponding to p-value > area under the curve corresponding to $t_{4,0.975}$ $(=\alpha)$)

⑤ Since p-value > α, we fail to reject H_0.

⑥ There is not enough evidence to conclude that

There's significant difference between the leading edge and trailing edge on average

---

Using 95% CI method:

$$\longrightarrow \bar{d} \pm \boxed{t_{(n-1,\, 1-\alpha/2)}}\, \frac{S_d}{\sqrt{n}}$$

$$\begin{cases} H_0: \mu_d = 0 \\ H_a: \mu_d \neq 0 \end{cases}$$

$$= -8\times10^{-4} \pm t_{(4,\, 0.975)} \cdot \frac{0.0023}{\sqrt{5}}$$

$$= -8\times10^{-4} \pm 2.776\,(0.001)$$

$$\Rightarrow = (-0.00358,\ 0.00198)$$

Since the 95% CI contains zero, we fail to reject H_0. There's not enough evidence to conclude that the leading edge measurement is significantly different from the trailing edge measurement.

Two-Sample Data

Hypothesis Testing

Null

Alternative

P-value

CI method

Matched Pairs

Two-sample

Two-sample data

Paired differences provide inference methods of a special kind for comparison. Methods that can be used to compare two means where two different *unrelated* samples will be discussed next.

> SAT score of high school A vs. high school B
>
> Severity of a disease in men vs. women
>
> Height of Liverpool soccerr players vs. Man United soccer players
>
> Fuel economy of gas formula type A vs. formula type B

Hypothesis Testing

Null

Alternative

P-value

CI method

Matched Pairs

Two-sample

Two-sample data

Notations:

	Sample 1	Sample 2
Sample size	n_1	n_2
* true means	$\boxed{\mu_1}$	$\boxed{\mu_2}$
Sample means	$\bar{x}_1$	$\bar{x}_2$
* true variance	σ_1^2	σ_2^2
Sample variance	s_1^2	s_2^2

$$H_0: \mu_1 = \mu_2 \equiv \mu_1 - \mu_2 = 0$$

$H_a: \mu_1 - \mu_2 \neq 0$ ✓

Large Samples

- Hypothesis Testing
- Null
- Alternative
- P-value
- CI method
- Matched Pairs
- **Two-sample**

(want to compare two true means μ_1, μ_2)

Large samples $(n_1 \geq 25, n_2 \geq 25)$

$H_0: \mu_1 = \mu_2 \Rightarrow \boxed{\mu_1 - \mu_2 = 0}$ vs. $H_a: \boxed{\mu_1 - \mu_2 \neq 0}$

The difference in sample means $\overline{x}_1 - \overline{x}_2$ is a natural statistic to use in comparing μ_1 and μ_2.

$\mu_1 - \mu_2 < 10$

or $\mu_1 - \mu_2 > 15$

i.e

$$\mathrm{E}(\overline{X}_1) = \mu_1 \quad \mathrm{E}(\overline{X}_2) = \mu_2 \quad \mathrm{Var}(\overline{X}_1) = \frac{\boxed{\sigma_1^2}}{n_1} \quad \mathrm{Var}(\overline{X}_2) = \frac{\sigma_2^2}{n_2}$$

If σ_1 and σ_2 are **known**, then we have

$$\mathrm{E}(\overline{X}_1 - \overline{X}_2) = \mathrm{E}(\overline{X}_1) - \mathrm{E}(\overline{X}_2) = \mu_1 - \mu_2$$

$$\mathrm{Var}(\overline{X}_1 - \overline{X}_2) = \mathrm{Var}(\overline{X}_1) + \mathrm{Var}(\overline{X}_2) = \boxed{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$= \mathrm{Var}(\overline{X}_1) + (-1)^2 \mathrm{Var}(\overline{X}_2)$

Hypothesis Testing

Null

Alternative

P-value

CI method

Matched Pairs

Two-sample

Large samples $(n_1 \geq 25, n_2 \geq 25)$

If, in addition, n_1 and n_2 are large,

$\overline{X}_1 \sim N(\mu_1, \frac{\sigma_1^2}{n_1})$ is independent of $\overline{X}_2 \sim N(\mu_2, \frac{\sigma_2^2}{n_2})$ (by CLT).

So that $\overline{X}_1 - \overline{X}_2$ is approximately Normal (trust me)

test statistic

$$Z = \frac{\overline{X}_1 - \overline{X}_2 - \overbrace{(\mu_1 - \mu_2)}^{\#}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0, 1)$$

Previously: $Z = \dfrac{\bar{X} - \mu^{\#}}{\sqrt{\sigma^2/n}} \sim N(0,1)$

(one sample)

Hypothesis Testing

Null

Alternative

P-value

CI method

Matched Pairs

Two-sample

Large samples $(n_1 \geq 25, n_2 \geq 25)$

So, if we want to test $H_0 : \mu_1 - \mu_2 = \#$ with some alternative hypothesis, σ_1 and σ_2 are known, and $n_1 \geq 25, n_2 \geq 25$, then we use the statistic

$$K = \frac{\overline{X}_1 - \overline{X}_2 - (\#)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

which has a $N(0, 1)$ distribution if

> 1. H_0 is true
> 2. The sample 1 points are iid with mean μ_1 and variance σ_1^2, and the sample 2 points are iid with mean μ_2 and variance σ_2^2.
>
> 3.Sample I is independent of sample II

Large samples $(n_1 \geq 25, n_2 \geq 25)$

The confidence intervals (2-sided, 1-sided upper, and 1-sided lower, respectively) for $\mu_1 - \mu_2$ are:

- *Two-sided* $100(1-\alpha)\%$ confidence interval for $\mu_1 - \mu_2$

$$(\overline{x}_1 - \overline{x}_2) \pm z_{1-\alpha/2} * \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

- *One-sided* $100(1-\alpha)\%$ confidence interval for $\mu_1 - \mu_2$ with a upper confidence bound

$$(-\infty\ ,\ (\overline{x}_1 - \overline{x}_2) \pm z_{1-\alpha} * \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}})$$

- *One-sided* $100(1-\alpha)\%$ confidence interval for μ with a lower confidence bound

$$((\overline{x}_1 - \overline{x}_2) \pm z_{1-\alpha} * \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}})\ ,\ +\infty)$$

Large samples $(n_1 \geq 25, n_2 \geq 25)$

If σ_1 and σ_2 are **unknown**, and $n_1 \geq 25, n_2 \geq 25$, then we use the statistic

replace σ_1^2 by s_1^2 and σ_2^2 by s_2^2

$$K = \frac{\overline{X}_1 - \overline{X}_2 - (\#)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

and confidence intervals (2-sided, 1-sided upper, and 1-sided lower, respectively) for $\mu_1 - \mu_2$:

- *Two-sided* $100(1-\alpha)\%$ confidence interval for $\mu_1 - \mu_2$

$$(\overline{x}_1 - \overline{x}_2) \pm z_{1-\alpha/2} * \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Hypothesis Testing

Null

Alternative

P-value

CI method

Matched Pairs

Two-sample

Large samples $(n_1 \geq 25, n_2 \geq 25)$

- *One-sided* $100(1-\alpha)\%$ confidence interval for $\mu_1 - \mu_2$ with a upper confidence bound

$$\left(-\infty\ ,\ (\bar{x}_1 - \bar{x}_2) \pm z_{1-\alpha} * \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}\right)$$

- *One-sided* $100(1-\alpha)\%$ confidence interval for μ with a lower confidence bound

$$\left((\bar{x}_1 - \bar{x}_2) \pm z_{1-\alpha} * \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}\ ,\ +\infty\right)$$

Hypothesis Testing

Null

Alternative

P-value

CI method

Matched Pairs

Two-sample

Example:[Anchor bolts]

An experiment carried out to study various characteristics of anchor bolts resulted in 78 observations on shear strength (kip) of 3/8-in. diameter bolts and 88 observations on strength of 1/2-in. diameter bolts.

Let Sample 1 be the 1/2 in diameter bolts and Sample 2 be the 3/8 indiameter bolts.

Using a significance level of $\alpha = 0.01$, find out if the 1/2 in bolts are more than 2 kip stronger (in shear strength) than the 3/8 in bolts. Calculate and interpret the appropriate 99% confidence interval to support the analysis.

given info.

- $n_1 = 88, n_2 = 78$
- $\overline{x}_1 = 7.14, \overline{x}_2 = 4.25$
- $s_1 = 1.68, s_2 = 1.3$

① $H_0: \mu_1 - \mu_2 = 2$ vs. $H_a: \mu_1 - \mu_2 > 2$

② $\alpha = 0.01$

Hypothesis Testing

Null

Alternative

P-value

CI method

Matched Pairs

Two-sample

Example:[Anchor bolts]

- $n_1 = 88, n_2 = 78$
- $\overline{x}_1 = 7.14, \overline{x}_2 = 4.25$
- $s_1 = 1.68, s_2 = 1.3$

③ since $n_1, n_2 > 25$, I'll use

$$\boxed{K = \frac{(\bar{x}_1 - \bar{x}_2) - 2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}}$$

If we assume H_0 is true, sample I is iid with mean μ_1, variance σ_1^2 independent of sample II iid with mean μ_2 and variance σ_2^2, $K \sim N(0,1)$

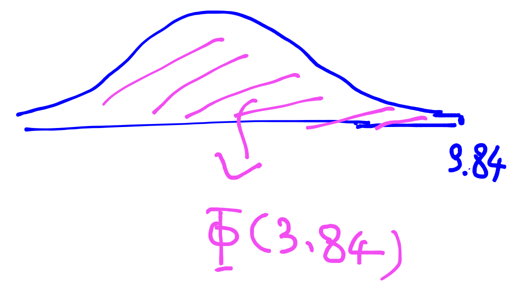
(4) $K = \dfrac{(\bar{x}_1 - \bar{x}_2) - 2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} = \dfrac{(7.14 - 4.25) - 2}{\sqrt{\dfrac{1.68^2}{88} + \dfrac{1.3^2}{78}}} = \boxed{3.84}$

p-value $= P(Z > K) = P(Z > 3.84)$

$H_a: \mu_1 - \mu_2 > 2$

$= 1 - \Phi(3.84)$ ≈ 1

$\approx 1 - 1 \approx 0$



(5) with a p-value $\approx 0 < \alpha = 0.01$, we reject H_0 in favor of H_a.

(6) There's enough evidence that the $\frac{1}{2}$ in bolts are more than 2 kip stronger than $\frac{3}{8}$ in bolts on average

99% lower bound CI (since $H_a: \mu_1 - \mu_2 > 2$)

$$\left(\bar{x}_1 - \bar{x}_2 - z_{1-\alpha} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}, +\infty\right)$$

$$= \left((7.14 - 4.25) - z_{1-0.01} \sqrt{\frac{1.68^2}{88} + \frac{1.3^2}{78}}, +\infty\right)$$

$$z_{.99}$$

$$= \left(2.89 - \boxed{2.33}(0.232), +\infty\right)$$

$$= (2.35, +\infty)$$

We're 99% confident that the true mean strength of the $\frac{1}{2}$ in bolts is at least 2.35 kip stronger than the true mean strength of the $\frac{3}{8}$ in bolts.

Small Samples

Hypothesis Testing

Null

Alternative

P-value

CI method

Matched Pairs

Two-sample

Small samples

If $n_1 < 25$ or $n_2 < 25$, then we need some **other assumptions** to hold in order to complete inference on two-sample data.

> We need two independent samples to be iid Normally distributed and $\sigma_1^2 \approx \sigma_2^2$

A test statistic to test $\mathrm{H}_0 : \mu_1 - \mu_2 = \#$ against some alternative is

$$* \quad K = \frac{\boxed{\overline{X}_1 - \overline{X}_2 - (\#)}}{S_p\sqrt{(\frac{1}{n_1} + \frac{1}{n_2})}}$$

where S_p^2 is called **pooled sample variance** and is defined as

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

Hypothesis Testing

Null

Alternative

P-value

CI method

Matched Pairs

Two-sample

Small samples

Also assuming

- 1 H_0 is true,
- 2 The sample 1 points are iid $N(\mu_1, \sigma_1^2)$, 3 the sample 2 points are iid $N(\mu_2, \sigma_2^2)$,
- 4 and the sample 1 points are independent of the sample 2 points and $\boxed{\sigma_1^2 \approx \sigma_2^2.}$ 5

Then

5 assumptions to use

$$K = \frac{\overline{X}_1 - \overline{X}_2 - (\#)}{S_p\sqrt{(\frac{1}{n_1} + \frac{1}{n_2})}} \sim \boxed{t_{(n_1+n_2-2)}}$$

Hypothesis Testing

Null

Alternative

P-value

CI method

Matched Pairs

Two-sample

Small samples

$1 - \alpha$ confidence intervals (2-sided, 1-sided upper, and 1-sided lower, respectively) for $\mu_1 - \mu_2$ under these assumptions are of the form:

(let $\boxed{\nu = n_1 + n_2 - 2)}$

- *Two-sided* $100(1-\alpha)\%$ confidence interval for $\mu_1 - \mu_2$

$$(\bar{x}_1 - \bar{x}_2) \pm \boxed{t_{(\nu,1-\alpha/2)}} * \boxed{S_p\sqrt{(\frac{1}{n_1}+\frac{1}{n_2})}}$$

- *One-sided* $100(1-\alpha)\%$ confidence interval for $\mu_1 - \mu_2$ with a upper confidence bound

$$(-\infty\ ,\ (\bar{x}_1 - \bar{x}_2) + t_{(\nu,1-\alpha)} * S_p\sqrt{(\frac{1}{n_1}+\frac{1}{n_2})}$$

n_1+n_2-2

Hypothesis Testing

Null

Alternative

P-value

CI method

Matched Pairs

Two-sample

Small samples

- *One-sided* $100(1-\alpha)\%$ confidence interval for μ with a lower confidence bound

$$\left((\bar{x}_1 - \bar{x}_2) - t_{(\nu,1-\alpha)} * S_p \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)},\ +\infty\right)$$

$\nu = n_1 + n_2 - 2$

In general: CI:

estimate $\pm$ Population quantiles $\times$ SD(estimate)

$\mu = 0$

$\mu \neq 0$

large sample

: $\bar{x} \pm Z_{1-\alpha/2}\, S/\sqrt{n}$

* HW 9 posted (optional)
  Due Thursday Dec. 12.

* Final exam:
  - Wednesday, Dec 18
    9:45 – 11:45 (in-class)
  - Comprehensive with focus on materials after quiz 3.

Small samples

Example:[Springs]

The data of W. Armstrong on spring lifetimes (appearing in the book by Cox and Oakes) not only concern spring 2 longevity at a 950 N/ mm^2 stress level but also longevity at 1 a 900 N/ mm^2 stress level.

Let sample 1 be the 900 N/ mm^2 stress group and sample 2 be the 950 N/ mm^2 stress group.

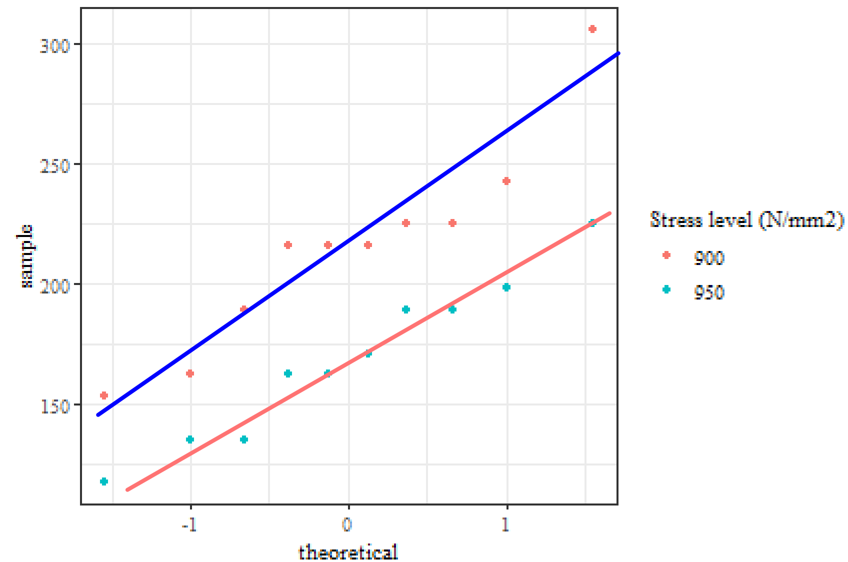
1 900 N/mm2 Stress	2 950 N/mm2 Stress
216, 162, 153, 216, 225, 216, 306, 225, 243, 189	225, 171, 198, 189, 189, 135, 162, 135, 117, 162

$n_1 = 10$ $n_2 = 10$

Small samples

Example:[Springs]

$\boxed{\text{Normal}}$ plots of Spring lifetime under two different stress level.



Let's do a hypothesis test to see if the sample 1 springs lasted significantly longer than the sample 2 springs. (For $\alpha = 0.05$)

Note: however the sample sizes are small, the data look pretty Normal.

1, $\begin{cases} H_0: \mu_1 - \mu_2 = 0 \\ H_a: \boxed{\mu_1 - \mu_2 > 0} \end{cases}$

2, $\alpha = 0.05$

3, If we assume that ① H_0 is true & ② sample 1 is iid $N(\mu_1, \sigma_1^2)$ ③ independent from sample 2 ④ iid $N(\mu_2, \sigma_2^2)$ and ⑤ $\sigma_1^2 \approx \sigma_2^2$, then the test statistic is

$$\rightarrow K = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \;\dot{\sim}\; t_{(n_1+n_2-2)}$$

4, calculates

$$\begin{cases} \bar{x}_1 = 215.1, & S_1 = 42.9, \quad S_1^2 = 1840.41 \\ \bar{x}_2 = 168.3, & S_2 = 33.1, \quad S_2^2 = 1095.61 \end{cases}$$

$$S_p = \sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}} = \sqrt{\frac{(10-1)1840.41 + (10-1)1095.61}{(10+10-2)}}$$

$$\Rightarrow S_p = 38.3$$

Then $K = \dfrac{(\bar{x}_1 - \bar{x}_2) - 0}{\boxed{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}} = \dfrac{(215.1 - 163.3)}{38.3 \sqrt{\frac{1}{10} + \frac{1}{10}}} = 2.7$

Recall:

We need to calculate p-value now.

* Corresponding to α: $t_{(n_1+n_2-2,\ 1-\alpha)} = t_{(18,\ 0.95)}$

by table $= 1.73$

(By the methods we learned about p-value and α, we just need to decide if p-value $\gtrless \alpha$)

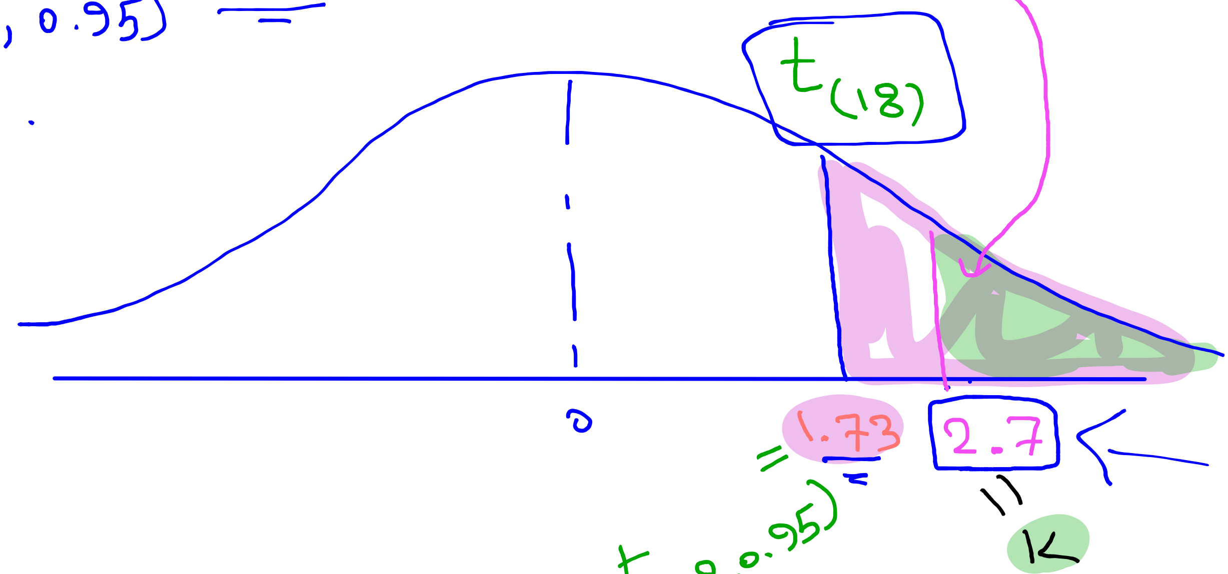
→ * Corresponding to p-value $\equiv k = 2.7$

Now:

one sided test

p-value $= P(T > K) = P(T > 2.7)$

$t_{(18,\, 0.95)} = 1.73$



5/ Since $K > 1.73 \; (= t_{(18,\, 0.95)})$

$\Rightarrow$ P-value $< \alpha$ $\Rightarrow$ Reject H_0

(area under the curve for $P(T > K)$ is smaller than α)

6, There is enough evidence to conclude that springs on average last longer subjected to $\underline{900}$ N/mm^2 of stress than $\underline{950}$ N/mm^2 of stress.

Hypothesis Testing

Null

Alternative

P-value

CI method

Matched Pairs

Two-sample

Small samples

Example:[Stopping distance]

Suppose μ_1 and μ_2 are true mean stopping distances (in meters) at 50 mph for cars of a certain type equipped with two different types of breaking systems.

Suppose $n_1 = n_2 = 6, \overline{x}_1 = 115.7, \overline{x}_2 = 129.3, s_1 = 5.08$, and $s_2 = 5.38$.

Use significance level $\boxed{\alpha = 0.01}$ to test
$\text{H}_0 : \mu_1 - \mu_2 = -10$ vs. $\text{H}_A : \mu_1 - \mu_2 < -10$.

Construct a 2-sided 99% confidence interval for the true difference in stopping distances.

① $\begin{cases} H_0: \mu_1-\mu_2 = -10 \\ H_a: \mu_1-\mu_2 < -10 \end{cases}$

② $\alpha = 0.01$

③ Under the assumptions that ① H_0 is true and

② sample 1 is iid $N(\mu_1, \sigma_1^2)$ ③ independent of

④ sample 2 iid $N(\mu_2, \sigma_2^2)$ and ⑤ $\sigma_1^2 \approx \sigma_2^2$ ✓

we use test statistic

$$K = \frac{\bar{x}_1 - \bar{x}_2 - (-10)}{S_p \sqrt{1/n_1 + 1/n_2}}$$

where $\overset{*}{K} \sim t_{(n_1+n_2-2)}$

(4) $$S_p = \sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}} = \sqrt{\frac{(6-1)5.08^2+(6-1)5.38^2}{6+6-2}}$$

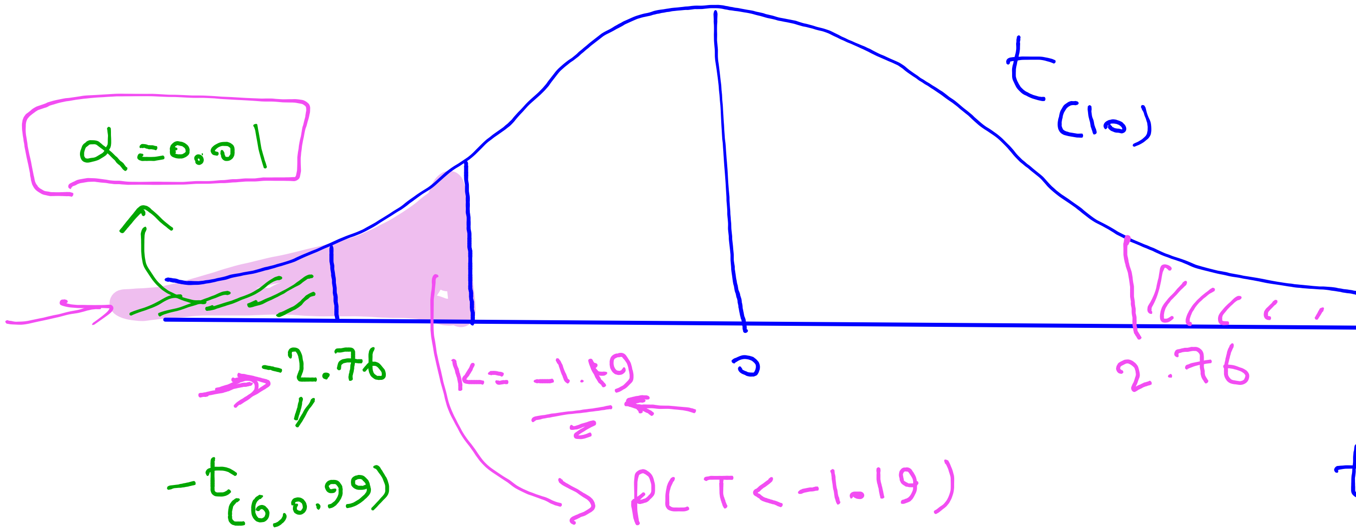
$$\rightarrow S_p = 5.23$$

$$\rightarrow K = \frac{(\bar{x}_1-\bar{x}_2)-(-10)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{(115.7-129.3)+10}{5.23\sqrt{\frac{1}{6}+\frac{1}{6}}} = -1.19$$

$$\begin{cases} H_0: \mu_1-\mu_2 = -10 \\ H_a: \mu_1-\mu_2 < -10 \end{cases} \rightarrow \text{p-value} = P(T < K) = P(T < -1.19)$$

$T \sim t_{(n_1+n_2-2)}$

by the table: $t_{(6+6-2,\ 1-\alpha)} = t_{(10,\ 0.99)} = \boxed{2.76}$



$$\text{since}\quad \underbrace{P(T<K) = P(T<-1.19)}_{p\text{-value}} > \underbrace{P(T<\overset{t_{(6,\,0.99)}}{-2.76})}_{\alpha}$$

we fail to reject H_0.

⑤ Since p-value > α, we fail to reject H_0

⑥ There is Not enough evidence to conclude that stopping distances for breaking system 1 are on average less than those of breaking system 2 by over 10 m.

Construct $\boxed{\text{2-sided}}$ 99% CI for $\mu_1 - \mu_2$ (the true difference of mean stopping distance):

using formula:

$$\rightarrow \left((\bar{x}_1 - \bar{x}_2) - t_{(n_1+n_2-2,\, 1-\alpha/2)} \, s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \,,\, (\bar{x}_1 - \bar{x}_2) + t_{(n_1+n_2-2,\, 1-\alpha/2)} \, s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right)$$

$\underbrace{\qquad}_{10}$ $\underbrace{\qquad}_{1-\frac{0.01}{2}=0.995}$

$$= \left((115.7 - 129.3) - 3.17 \cdot (5.23)\sqrt{\frac{1}{6}+\frac{1}{6}} \, , (115.7-129.3)+3.17(5.23)\sqrt{\frac{1}{6}+\frac{1}{6}} \right)$$

$$= \boxed{(-23.17, -4.03)}$$

we are 99% confident that the true mean stopping distance of system 1 is anywhere between 23.17 m to 4.03 m less than that of system 2.