LECTURE 12
HYPOTHESIS TESTING: PAIRED-SAMPLES t-TEST

PSY2002

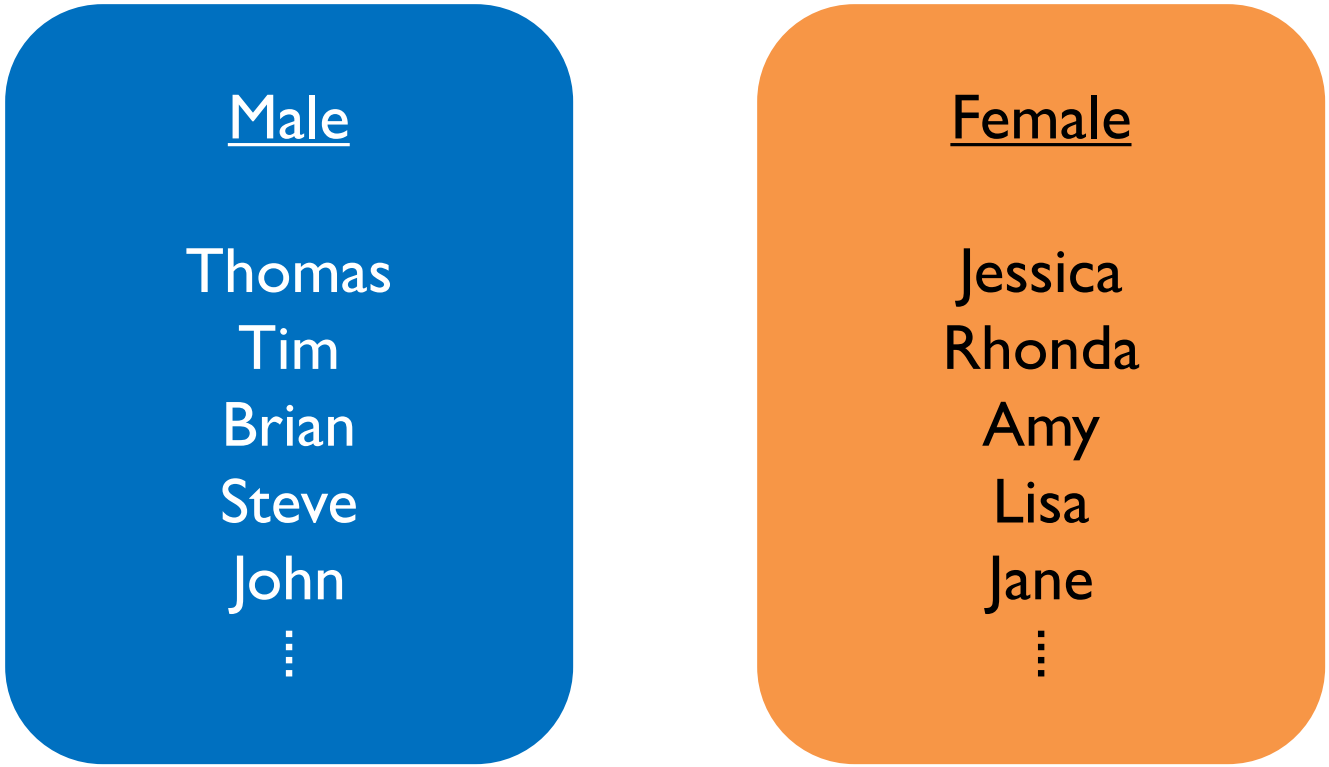
Hye Won Suk

- One-sample tests
  - One-sample *Z*-test
    - used when the population standard deviation (σ) is known
  - One-sample *t*-test
    - used when the population standard deviation (σ) is unknown

- Two-samples tests
  - Independent-samples *t*-test
    - used when two groups of data are collected by independent-groups design
  - Paired-samples *t*-test
    - used when two groups of data are collected by repeated-measures design

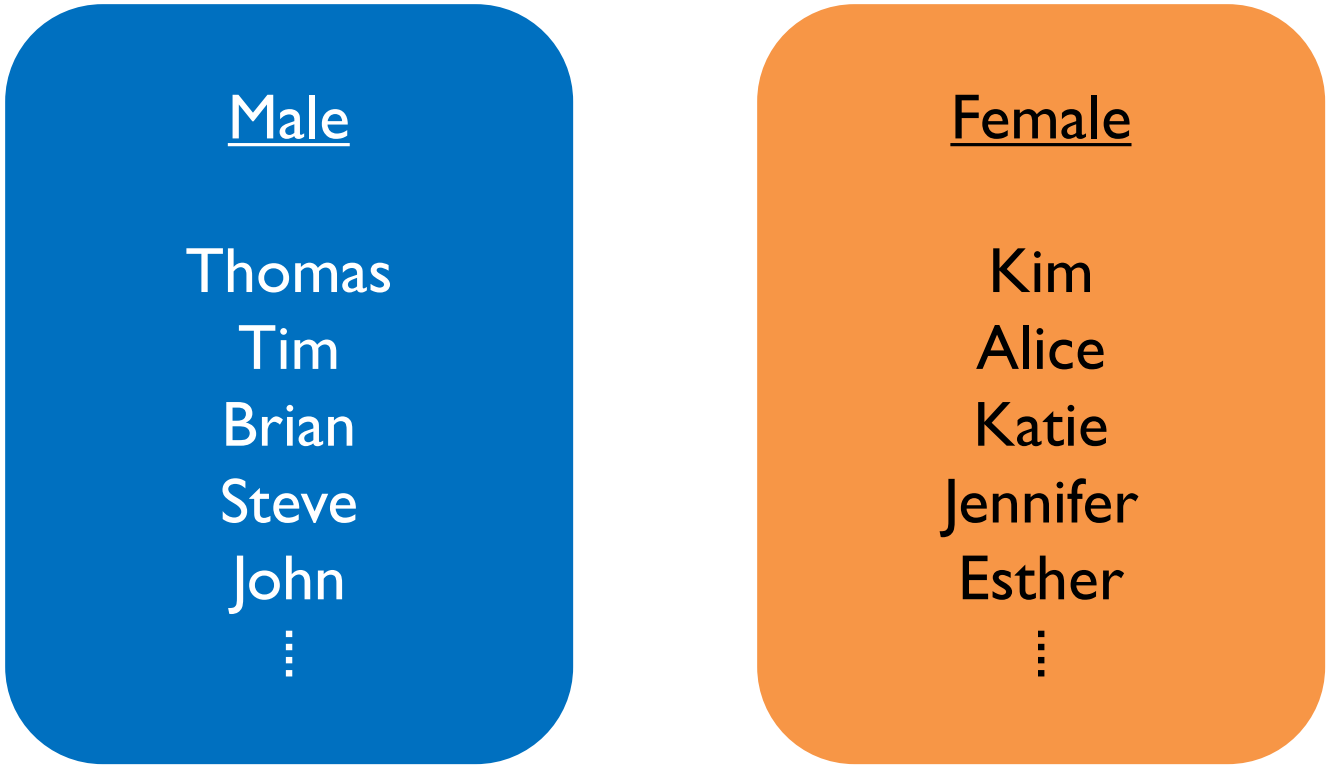
INDEPENDENT-GROUPS DESIGN

- The two samples are considered independent when the selection of people for one group is in no way affected by the selection of people for the other group.



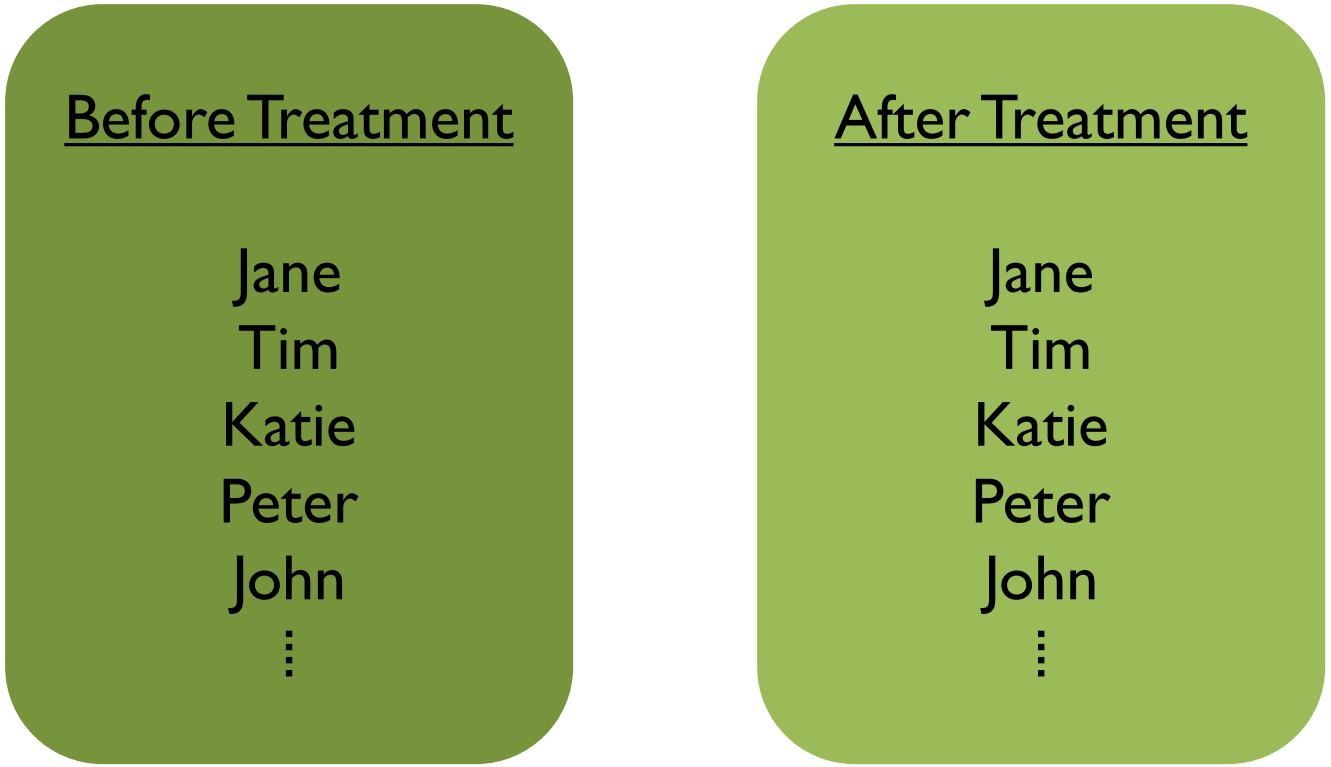
INDEPENDENT-GROUPS DESIGN

- In other words, having Thomas, Tim, Brian, Steve, and John in the male group does not imply that we should have Jessica, Rhonda, Amy, Lisa, and Jane in the female group. We could have selected other people like Kim, Alice, Katie, Jennifer, and Esther.



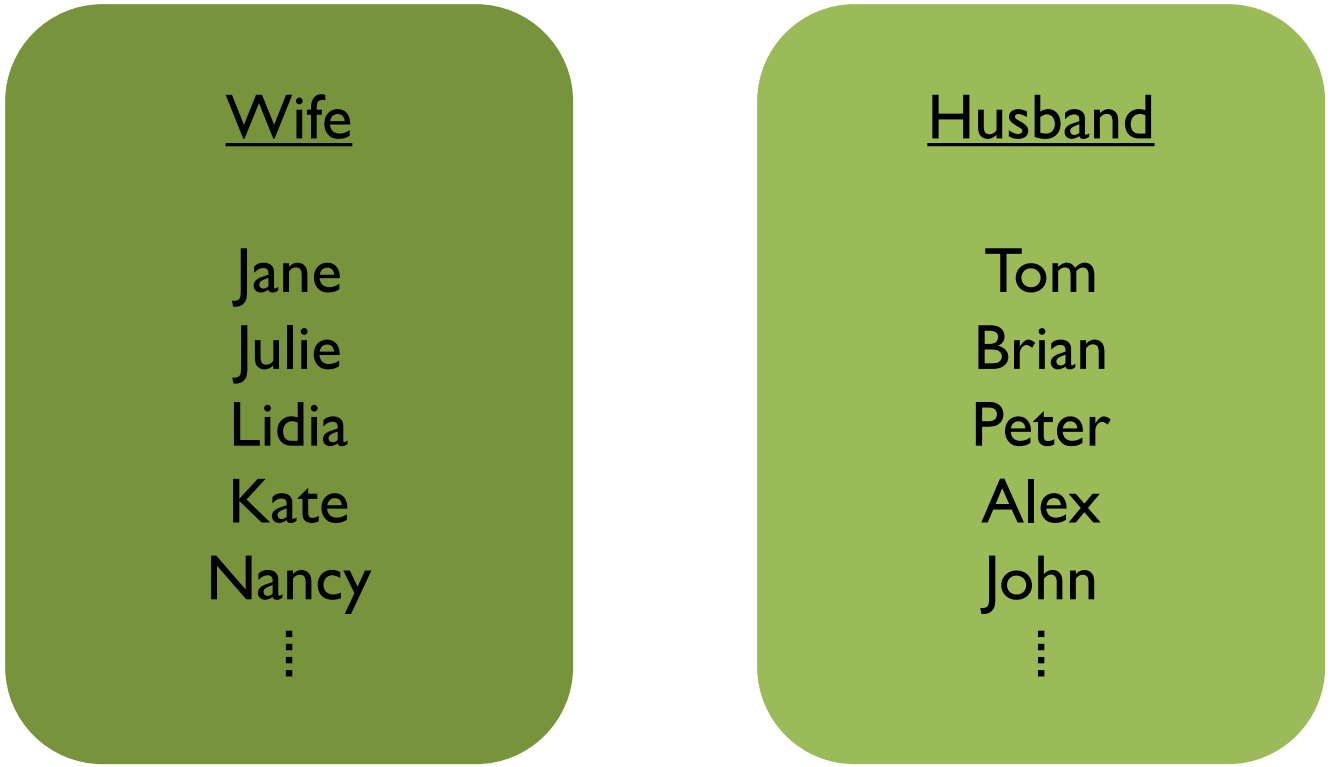
REPEATED-MEASURES DESIGN

- Each individual in a single sample is measured two (or more) times. The same group of participants is used in all of the treatment conditions.



REPEATED-MEASURES DESIGN

- In a broader sense, a matched (or paired) subject design is also considered repeated-measured design. In a matched-subject design, each individual in one sample is matched with an individual in the other sample.



- One such example is wife-husband pair.

REPEATED-MEASURES DESIGN

- In a broader sense, a matched (or paired) subject design is also considered repeated-measured design. In a matched-subject design, each individual in one sample is matched with an individual in the other sample.

Mother	Child
Jane	Elsa
Julie	Judy
Lidia	Sally
Kate	Mike
Nancy	Angelina
⋮	⋮

- Another example is mother-child pair.

RESEARCH DESIGNS

- According to which design we used to collect data, we should use different statistical tests to compare the two groups.

  - Independent-groups design or between-subjects design

    ➔ Independent-samples *t*-test

  - Repeated-measures design or within-subject design

    ➔ Paired-samples *t*-test

ASSUMPTIONS

- To discuss independent-samples *t*-test and paired-samples *t*-test, we will assume the following two things.

  - First, we will assume that the original scores form a normal distribution.

  - Second, we will assume that the population standard deviation (σ) is unknown.

PAIRED SAMPLES *t*-TEST

- Previous studies have shown that the number of minutes of sleep during class is a good indicator of lack of interest in class materials (the number of minutes of sleep during class and lack of interest in class materials are positively and strongly correlated). Use the sample data collected from six students who attended both Calculus III and Psych Stats to test the theory that the class materials in the two classes are not equally interesting.

(from Introduction to Psychological Statistics, coursepack by Prof. Rhonda Amsel)

PAIRED SAMPLES *t*-TEST

	Calculus III	Psych Stats
Student 1	6	2
Student 2	13	10
Student 3	8	7
Student 4	6	8
Student 5	19	10
Student 6	10	7

PAIRED SAMPLES *t*-TEST

	Calculus III	Psych Stats	Difference
Student 1	6	2	6 – 2 = 4
Student 2	13	10	13 – 10 = 3
Student 3	8	7	8 – 7 = 1
Student 4	6	8	6 – 8 = -2
Student 5	19	10	19 – 10 = 9
Student 6	10	7	10 – 7 = 3

Difference
= Calculus III – Psych Stats

- We can perform a one-sample *t*-test by using these difference scores.
- If the course materials for the two classes are equally interesting, the population mean of the difference scores will be zero.

PAIRED SAMPLES *t*-TEST

- Again, follow the five steps of hypothesis testing.
    - Step 1: State the hypotheses
    - Step 2: Set the criteria for a decision
    - Step 3: Collect data and compute test statistics
        - *t*-test: calculate a *t*-statistic
    - Step 4: Make a decision
    - Step 5: State a conclusion
- Let's do the example!

PAIRED SAMPLES *t*-TEST

- Step 1: State the hypotheses.

  - Null hypothesis
    - H_0: The class materials for the two classes, Calculus III and Psych Stats, are equally interesting; The population mean of the difference scores will be zero. ($\mu_D = 0$).

  - Alternative hypothesis
    - H_1: The class materials for the two classes, Calculus III and Psych Stats, are not equally interesting; The population mean of the difference scores will be different from zero. ($\mu_D \neq 0$).

*The null and alternative hypotheses of a paired-samples *t*-test are identical to those of a one sample *t*-test except that the population mean of the difference scores is always assumed to be ZERO in the null hypothesis.

PAIRED SAMPLES *t*-TEST

- Step 2: Set the criteria for a decision.
  - $\alpha = 0.05$
  - The alpha level (or level of significance) is a probability value that is used to define the concept of "very unlikely" in a hypothesis test.
  - By convention, we use $\alpha = 0.05$ unless otherwise specified.

PAIRED SAMPLES *t*-TEST

- Step 3: Compute a test statistic.

	Calculus III	Psych Stats	Difference
Student 1	6	2	4
Student 2	13	10	3
Student 3	8	7	1
Student 4	6	8	-2
Student 5	19	10	9
Student 6	10	7	3

- First, calculate the difference scores: Calculus III – Psych Stats.
- It would be equally good to calculate Psych Stats – Calculus III.

PAIRED SAMPLES *t*-TEST

- Step 3: Compute a test statistic.

	Calculus III	Psych Stats	Difference
Student 1	6	2	4
Student 2	13	10	3
Student 3	8	7	1
Student 4	6	8	-2
Student 5	19	10	9
Student 6	10	7	3

- Then ignore the original scores. Use the difference scores only.
- Perform a one-sample *t*-test using these difference scores as if they were the original scores.

PAIRED SAMPLES *t*-TEST

- Step 3: Compute a test statistic.

$$t = \frac{\bar{X} - \mu_{\bar{X}}}{s_{\bar{X}}} = \frac{\bar{X} - \mu}{s / \sqrt{n}}$$

- This is the formula that we used in a one-sample *t*-test.
- In this formula *X* indicates an original score.
- However, in a paired-samples t-test, this indicates a difference score.

$$t = \frac{\bar{D} - \mu_{\bar{D}}}{s_{\bar{D}}} = \frac{\bar{D} - \mu_{D}}{s_{D} / \sqrt{n}}$$

- Therefore, we will replace *X* by *D* to make it clear that *D* indicates a difference score.

PAIRED SAMPLES *t*-TEST

	Calculus	Psych Stats	Difference (D)	$D - \bar{D}$
Student 1	6	2	4	1
Student 2	13	10	3	0
Student 3	8	7	1	-2
Student 4	6	8	-2	-5
Student 5	19	10	9	6
Student 6	10	7	3	0

$$\bar{D} = \frac{\sum D}{n} = \frac{4+3+1+(-2)+9+3}{6} = \frac{18}{6} = 3$$

$$s_D^2 = \frac{\sum (D-\bar{D})^2}{n-1} = \frac{1^2+0^2+(-2)^2+(-5)^2+6^2+0^2}{5} = \frac{66}{5} = 13.2$$

$$s_D = \sqrt{s_D^2} = \sqrt{13.2} = 3.63$$

PAIRED SAMPLES *t*-TEST

- Step 3: Compute a test statistic.

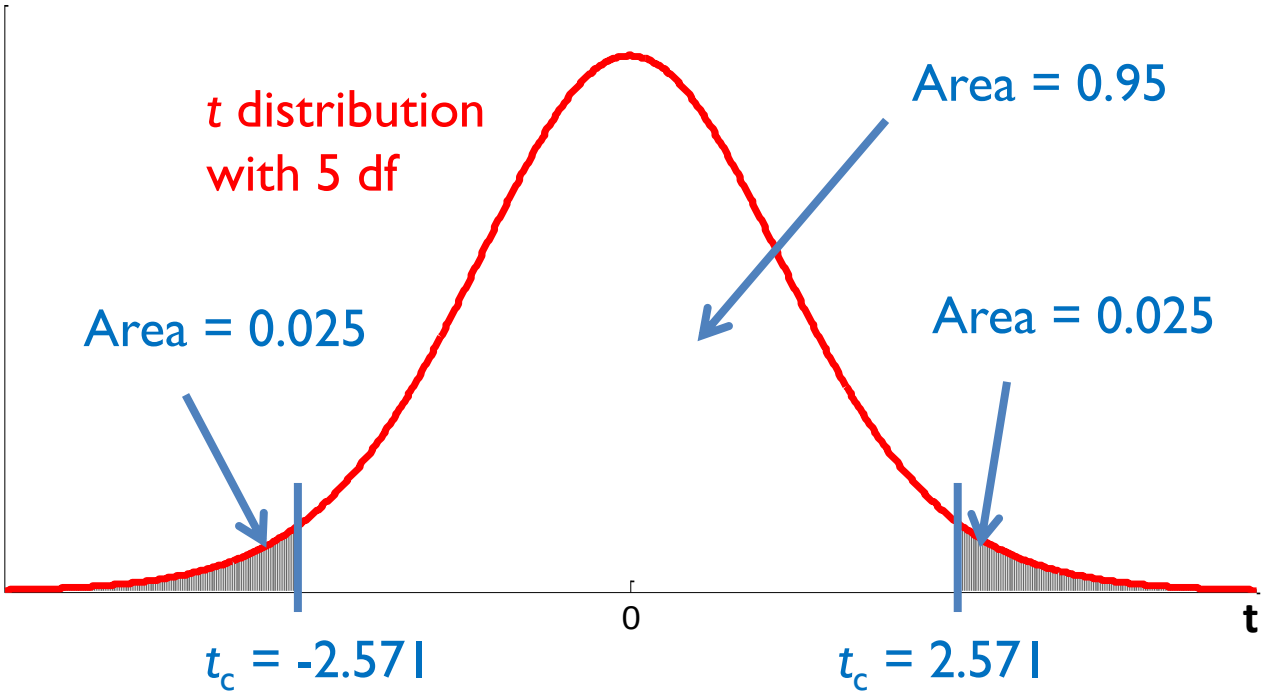
$$t = \frac{\bar{D} - \mu_{\bar{D}}}{s_{\bar{D}}} = \frac{\bar{D} - \mu_D}{s_D / \sqrt{n}} = \frac{3 - 0}{3.63 / \sqrt{6}} = \frac{3}{3.63 / 2.45} = \frac{3}{1.48} = 2.02$$

$$df = n - 1 = 6 - 1 = 5$$

In this case, *n* indicates the number of pairs.

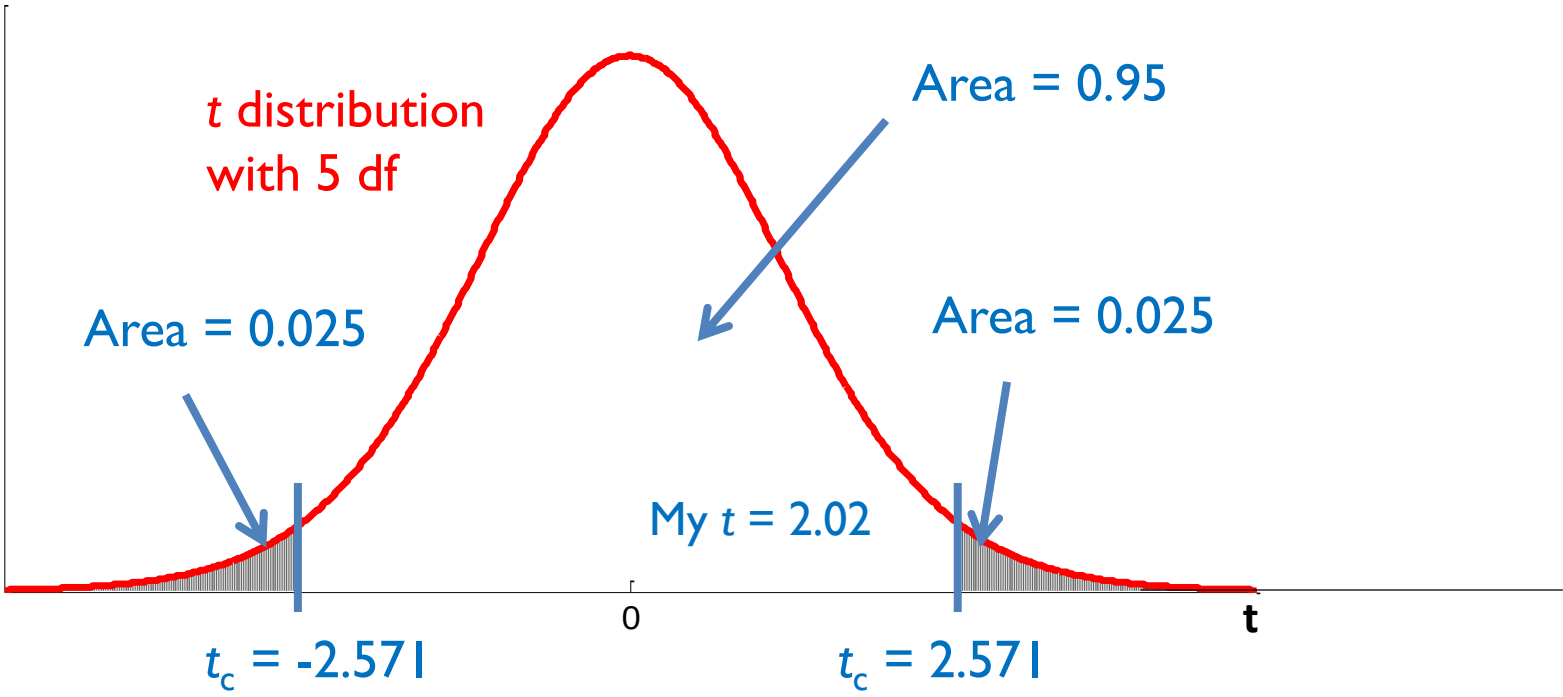
PAIRED SAMPLES *t*-TEST

- Step 4: Make a decision
  - Look up the *t* critical value in the table.
  - $\alpha = 0.05$ (two-tailed), df = 5
  - The *t* critical value is 2.571.



PAIRED SAMPLES *t*-TEST

- Step 4: Make a decision
  - |My *t*-value| < t_c ; 2.02 < 2.571
  - My *t*-value is not in the extreme zone (it does not fall within the critical region) ➔ Fail to reject the null



- Another equivalent way to make a decision would be to calculate the *p*-value and compare it to the level of significance ($\alpha = 0.05$).

- Like in the one-sample *t*-test, it is impossible to obtain the *p*-value using the *t*-table; the *t*-table provides the critical values only.

- SAS will provide the *p*-value. The exact *p*-value for the example is .0990.

  - $p < \alpha$ ➔ reject the null hypothesis.
  - $p > \alpha$ ➔ fail to reject the null hypothesis. (This is the case in this example).

PAIRED SAMPLES *t*-TEST

- Step 5: State a conclusion.

  - There is no significant difference in sleeping time between the two classes ($t(5) = 2.02$, $p > .05$). This implies that the class materials for the two classes, Calculus III and Psych Stats, are equally interesting.

  - If you obtain the exact *p*-value (using SAS), state a conclusion as follows.
  - There is no significant difference in sleeping time between the two classes ($t(5) = 2.02$, $p = .0990$). This implies that the class materials for the two classes, Calculus III and Psych Stats, are equally interesting.

SAS OUTPUT

The TTEST Procedure

Difference: calc3 - psychstats

N	Mean	Std Dev	Std Err	Minimum	Maximum
6	3.0000	3.6332	1.4832	-2.0000	9.0000

Mean	95% CL Mean		Std Dev	95% CL Std Dev	
3.0000	-0.8128	6.8128	3.6332	2.2679	8.9108

DF	t Value	Pr > |t|
5	2.02	0.0990

df = 5

t-value was 2.02 in hand calculation.

p-value = .0990

SUMMARY

- A paired-samples *t*-test is used when two samples of data are collected using the repeated-measures design (or within-subject design) and the researcher wants to examine the population mean difference between the two groups (or conditions).

- A paired-samples *t*-test is identical to the one-sample *t*-test except the followings:
  - Difference scores are used;
  - The null hypothesis states that the mean difference score in the population is 0.