

Biostatistics Week IX

Ege Ülgen, MD, PhD

1 December 2022



ACIBADEM
MEHMET ALİ AYDINLAR
ÜNİVERSİTESİ

Hypothesis Testing

- **Hypothesis:** an assumption that can be tested based on the evidence available
 - A novel drug is efficient in treating a certain disease
 - Regular smoking leads to lung cancer
 - Overweight individuals who (1) consume greasy food and (2) consume a low amount vegetables (1) have high levels of cholesterol and (2) have a higher risk of cardiovascular diseases
- **Hypothesis test:** investigation of the hypothesis using the sample
 - Assessing evidence provided by the data against the null claim (the claim which is to be assumed true unless enough evidence exists to reject it)

Null and Alternative Hypotheses

- H_0 – Null hypothesis
 - The mean of a variable is not different than c
 - There is no difference between the two groups' means
 - There is no difference compared to baseline
 - ...
- H_a or H_1 – Alternative hypothesis
 - There is a difference between the two groups' means
 - The mean in group A is higher than group B
 - ...

One- vs. Two-tailed Tests

- The coin is biased

Two-tailed

$$H_0: p = 0.5$$

$$H_a: p \neq 0.5$$

- The probability of heads is larger than 0.5

One-tailed

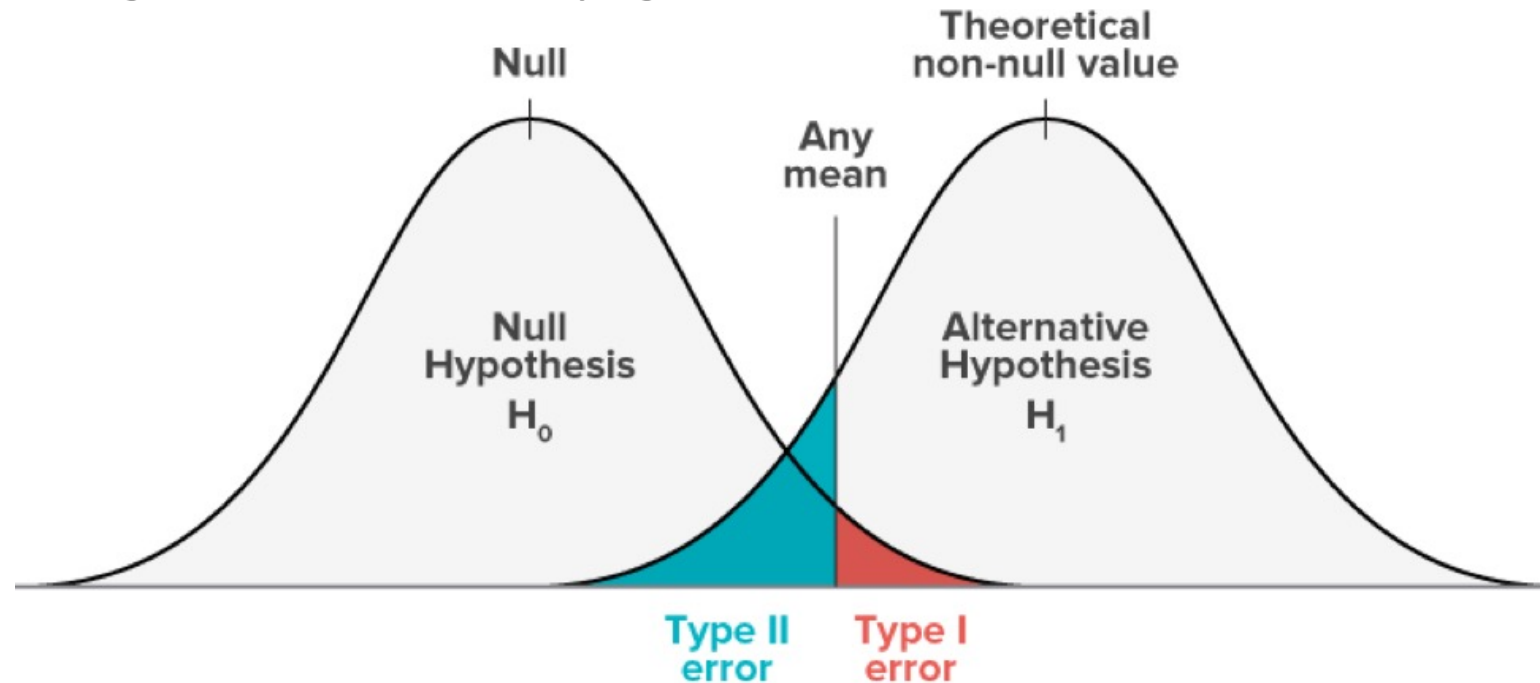
$$H_0: p \leq 0.5$$

$$H_a: p > 0.5$$

	Decision	
	Fail to reject	Reject
H_0		
True	Correct decision	Type I Error α
False	Type II Error β	Correct decision

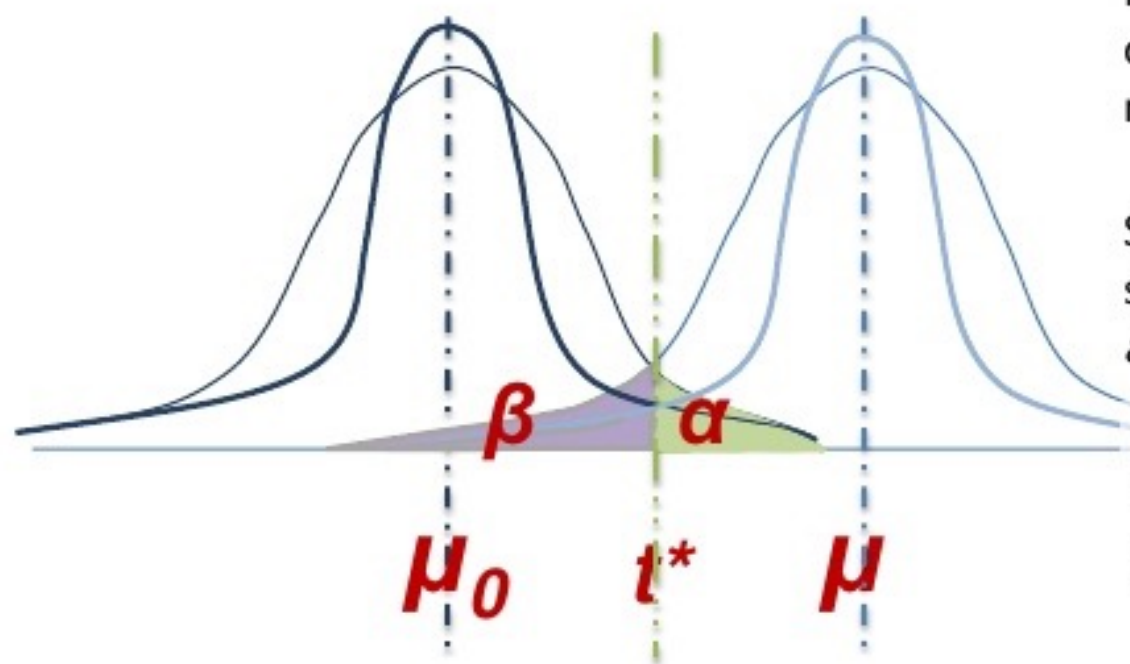
Hypothesis Testing

- $P(\text{Type 1 error}) = \alpha = P(\text{reject } H_0 \mid H_0 \text{ is true})$
- $P(\text{Type 2 error}) = \beta = P(\text{fail to reject } H_0 \mid H_0 \text{ is false})$
- As α gets larger β gets smaller, vice versa
- As n gets large, both α and β get smaller



Reducing the probabilities of errors

- We control the **spread** of our normal curves.



CLT

If our sample size increases, the centers don't move, and we reduce variability...

So by increasing n , our sample size, we've reduced *both* α and β .

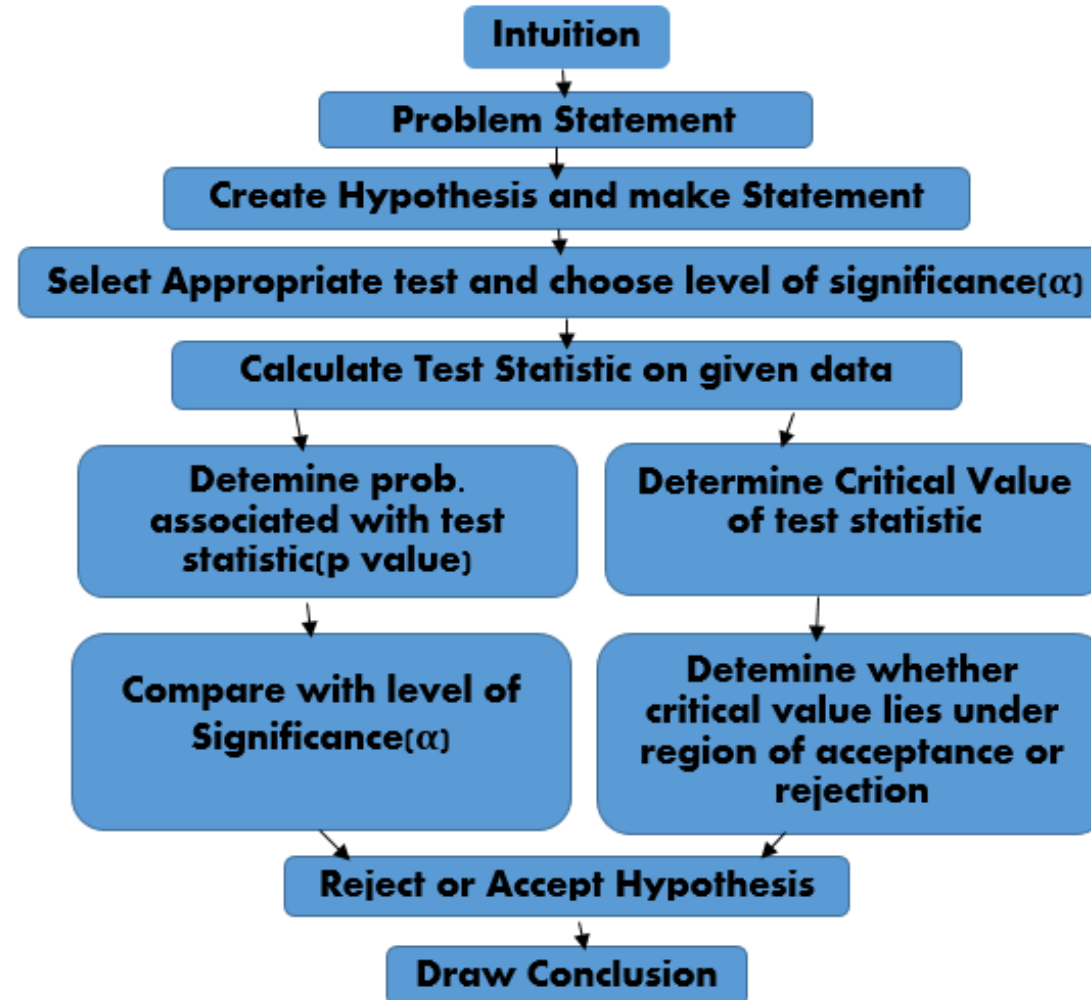
And if we've reduced β , we've increased the power ($1 - \beta$).

Hypothesis Testing

H_0	Decision	
	Fail to reject	Reject
True	Correct decision	Type I Error α
False	Type II Error β	Correct decision

- **Confidence level** = $1 - \alpha$
 - $P(\text{fail to reject } H_0 \mid H_0 \text{ is true})$
- **Statistical power** = $1 - \beta$
 - $P(\text{reject } H_0 \mid H_0 \text{ is false})$

Hypothesis Testing - Steps



Hypothesis Testing- Steps

1. Check assumptions, determine H_0 and H_a , choose α

- Assumptions differ based on the test
- The null hypothesis always contains equality (=)

2. Calculate the appropriate test statistic

- z , t , χ^2 , ...

3. Calculate critical values/p value

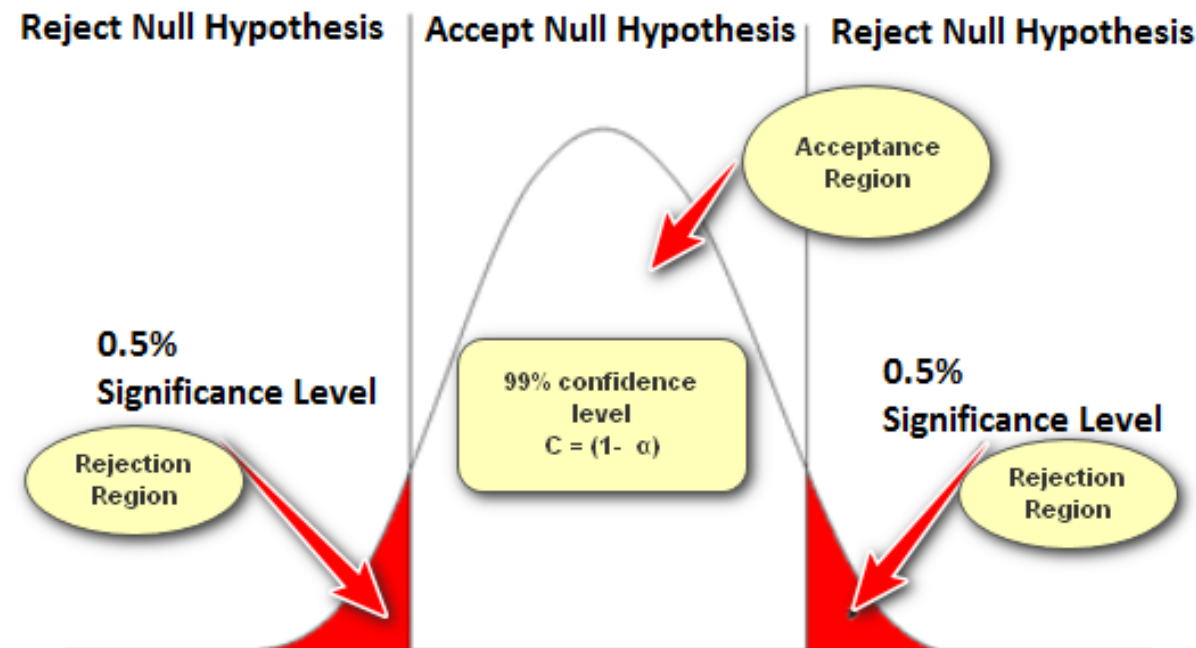
- With the aid of precalculated tables/software

4. Decide whether to reject/fail to reject H_0

- Reject if the statistic is within the critical region/ $p \leq \alpha$

Critical Value/Rejection Region

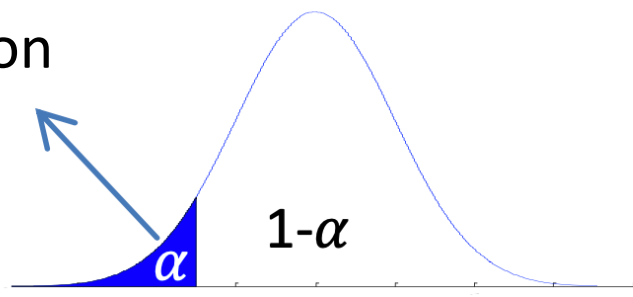
- We select α (**significance level**) prior to performing a hypothesis test
 - Some common values for α are 0.01, **0.05** and 0.10
- Based on the selected α , the critical values are calculated, and the rejection region is determined
 - the region where the null hypothesis is rejected



$$H_0: \mu = \mu_0$$

$$H_1: \mu < \mu_0$$

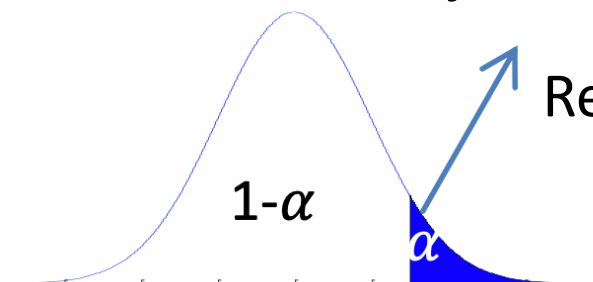
Rejection
region



$$H_0: \mu = \mu_0$$

$$H_1: \mu > \mu_0$$

Rejection region

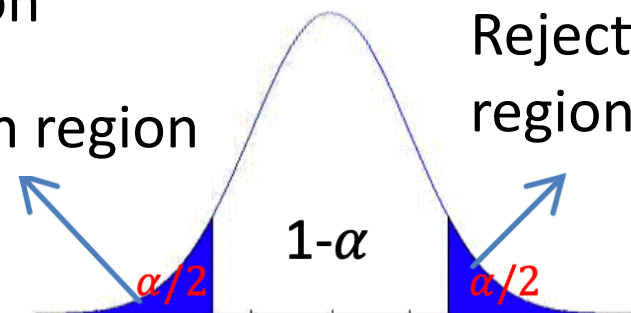


$$H_0: \mu = \mu_0$$

$$H_1: \mu \neq \mu_0$$

Rejection region

Rejection
region



Test Statistic

$$\text{test statistic} = \frac{\text{estimator} - \text{null value}}{\text{standard error of estimator}}$$

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

Commonly Used Hypothesis Tests

- Continuous (comparing means)
 - One sample
 - z-Test
 - Student's t-Test
 - Two samples
 - Student's t-Test
 - Equal variance
 - Unequal variance
 - Paired t-Test
 - >2 samples
 - Analysis of Variance (ANOVA)
- Categorical
 - Chi-squared (χ^2) Test

One-sample z-Test for μ

- **Assuming we know the true value of the population variance σ^2**
- If X_1, \dots, X_n are i.i.d. RVs $\sim N(\mu, \sigma^2)$, then:

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$
$$z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

One-Sample z-Test for μ – Example

- It is claimed that the post-treatment tumor volume of glioblastoma patients subject to a novel treatment is different than 5 cm³
- The mean tumor volume of 41 randomly-selected patients is 5.9 cm³
- **Population standard deviation** is 1.74

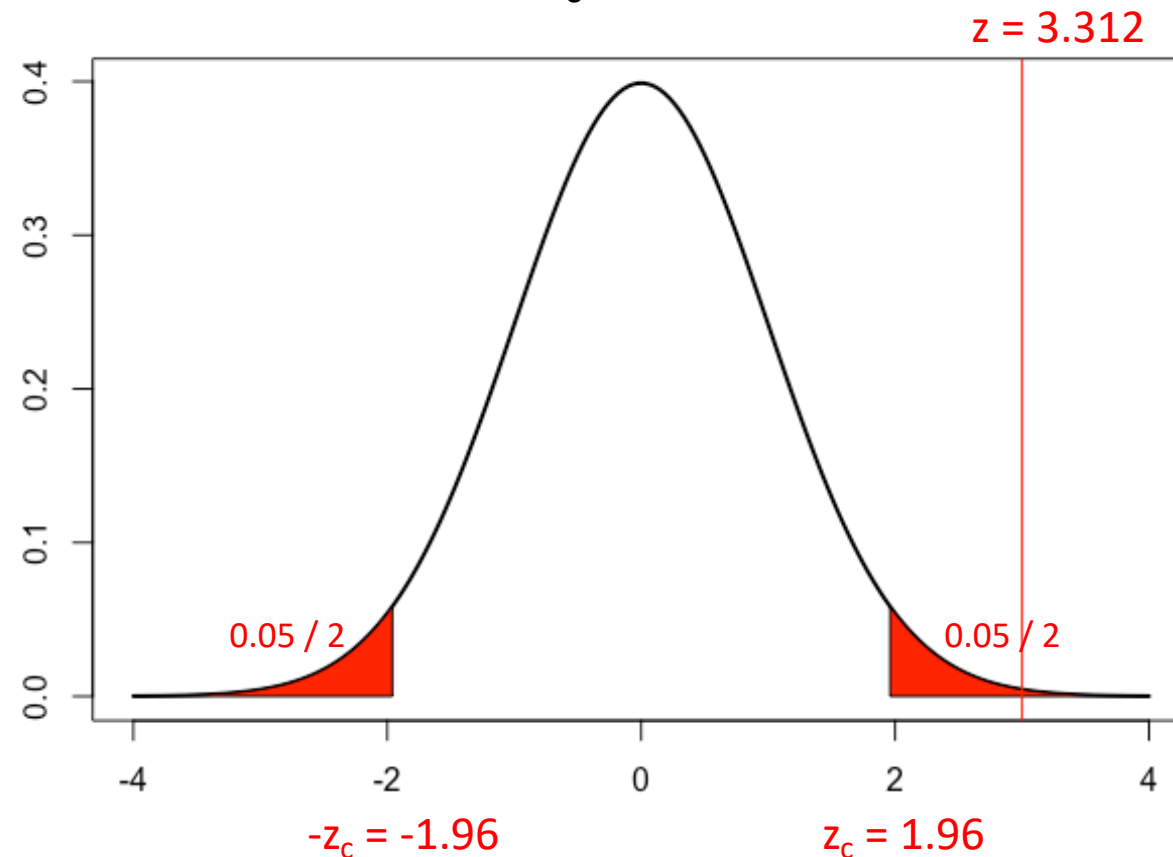
One-Sample z-Test for μ – Example (cont.)

1. Check assumptions, determine H_0 and H_a , choose α
 - Normality of the variable is checked, **Population variance is known**
 - $H_0: \mu = 5$ $H_a: \mu \neq 5$
 - $\alpha = 0.05$
2. Calculate the appropriate test statistic

$$z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{5.9 - 5}{1.74/\sqrt{41}} = 3.312$$

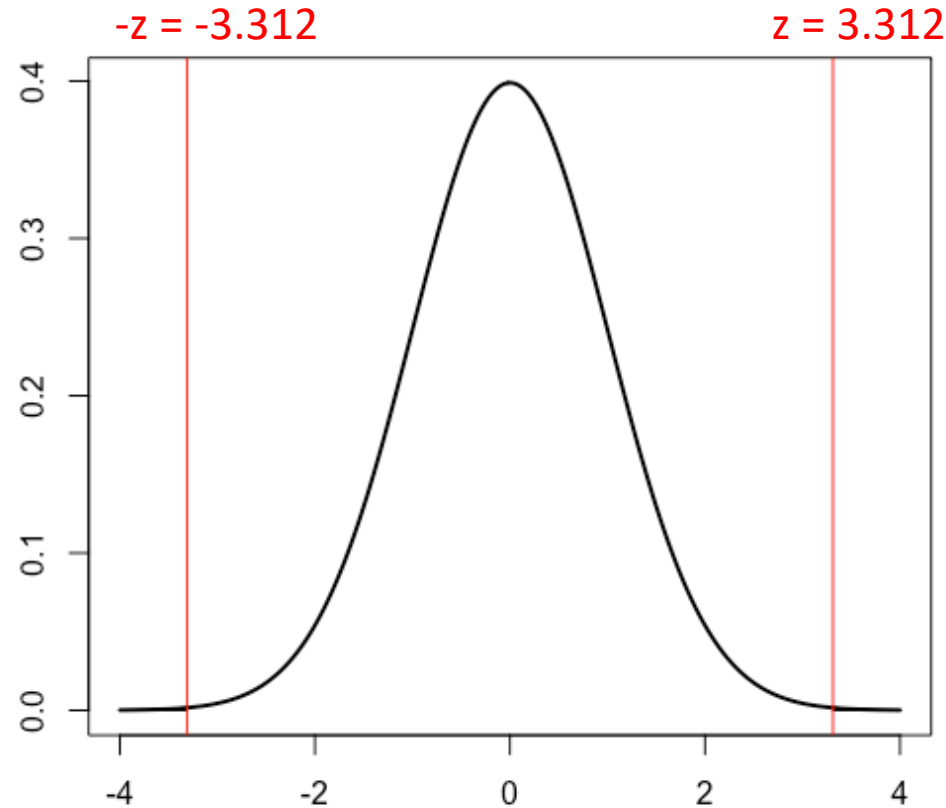
One-Sample z-Test for μ – Example (cont.)

3. Calculate **critical values**/p value
4. Decide whether to reject/fail to reject H_0



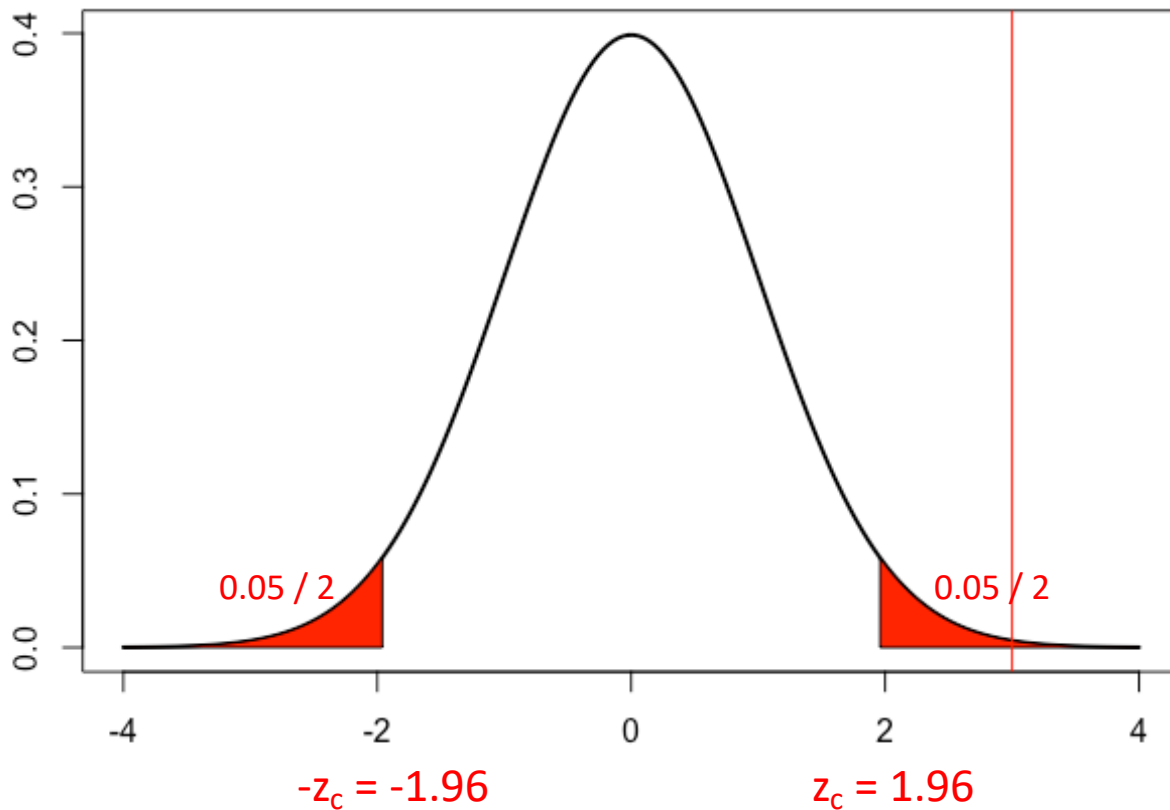
One-Sample z-Test for μ – Example (cont.)

3. Calculate critical values/**p value**
4. Decide whether to reject/fail to reject H_0



$p < 0.001$

Confidence Interval



$$P \left(\bar{X} - Z_c \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + Z_c \frac{\sigma}{\sqrt{n}} \right) = 0.95$$

$$\bar{X} \mp Z_c \frac{\sigma}{\sqrt{n}} = 5.9 \mp 1.96 \frac{1.74}{\sqrt{41}}$$

95% Confidence Interval for μ
[5.367, 6.433]

One-Sample z-Test for μ – Example (cont.)

5. State a conclusion:

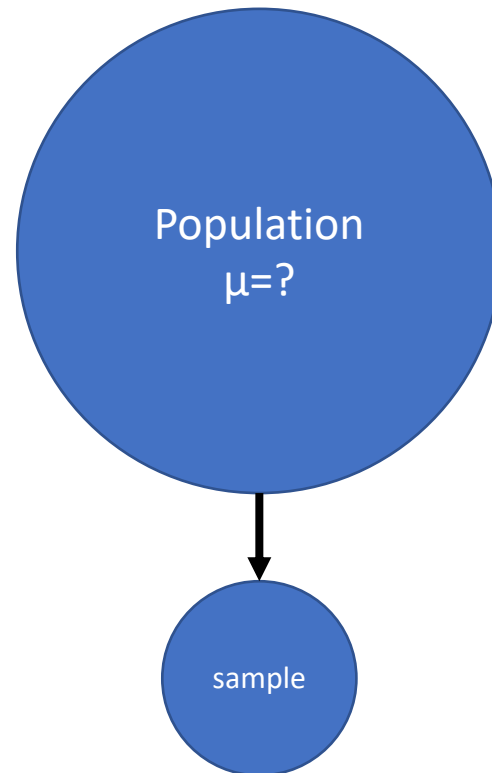
With 95% confidence, we can conclude that there is enough evidence to say that post-treatment tumor volume of glioblastoma patients subject to a novel treatment is different than 5 cm³

Post-treatment tumor volume of glioblastoma patients subject to a novel treatment was found to be different than 5 cm³ (z-test, $p < 0.001$)

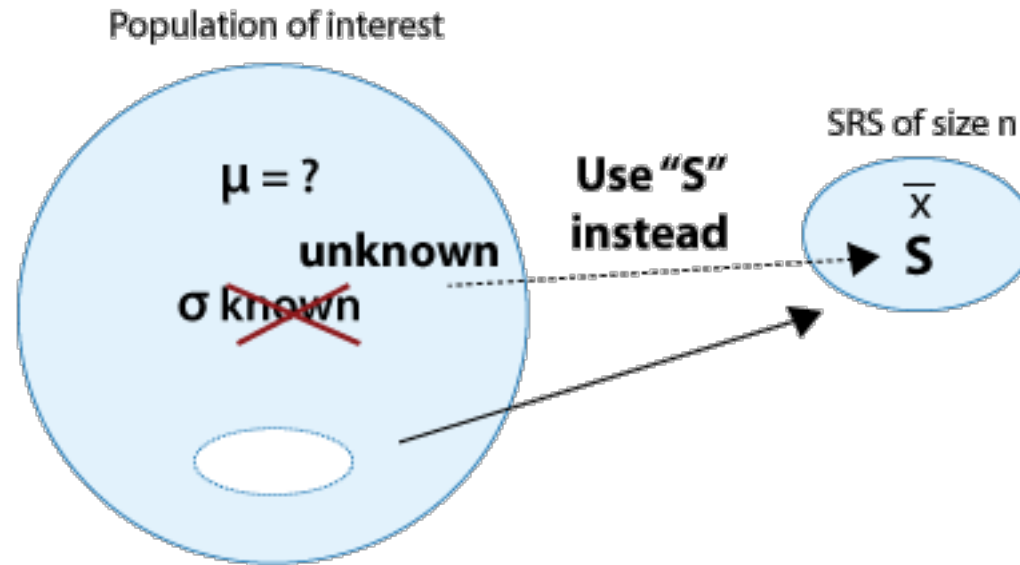
Post-treatment tumor volume of glioblastoma patients subject to a novel treatment was found to be different than 5 cm³ (z-test, 95% CI = 5.367-6.433)

One-Sample t-Test

- a statistical hypothesis test used to determine whether an unknown population mean is different from a specific value



One-Sample t-Test



One-Sample t-Test – Example I

- It is claimed that the post-treatment tumor volume of glioblastoma patients subject to a novel treatment is different than 5 cm^3
- The mean tumor volume of 41 randomly-selected patients is 5.9 cm^3
- **Population variance is unknown**
- Sample standard deviation is 1.74

One-Sample t-Test – Example I (cont.)

1. Check assumptions, determine H_0 and H_a , choose α

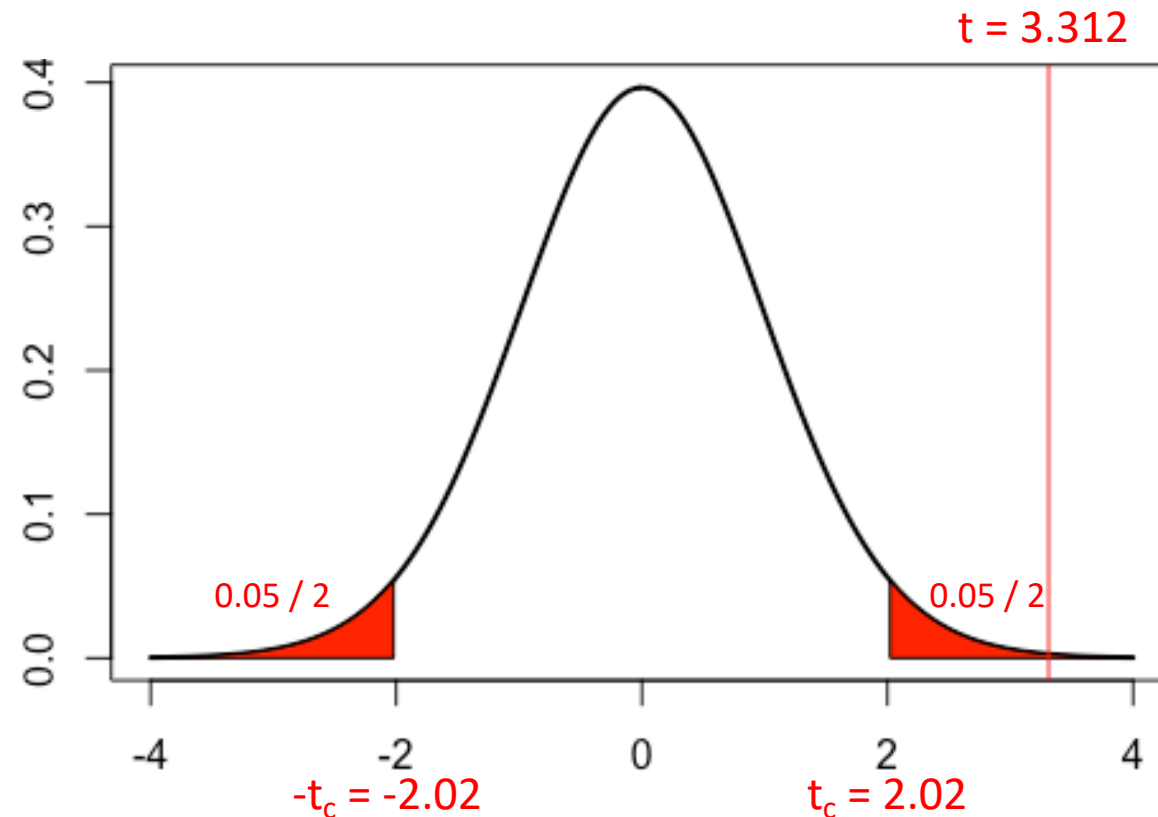
- Normality of the variable is checked
- $H_0: \mu = 5$ $H_a: \mu \neq 5$
- $\alpha = 0.05$

2. Calculate the appropriate test statistic

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}} = \frac{5.9 - 5}{1.74/\sqrt{41}} = 3.312 \quad (\sim t_{n-1} = t_{40})$$

One-Sample t-Test – Example I (cont.)

3. Calculate **critical values**/p value
4. Decide whether to reject/fail to reject H_0



One-Sample t-Test – Example II (cont.)

5. **State a conclusion:**

With 95% confidence, we can conclude that there is enough evidence to say that post-treatment tumor volume of glioblastoma patients subject to a novel treatment is different than 5 cm³.

One-Sample t-Test – Example II

id	week_1	cd4_1	week_2	cd4_2	perc_benefit
361	0	26	7.43	3	-11.905994
1017	0	13	7.00	10	-3.296703
519	0	3	8.14	5	8.190008
1147	0	65	33.00	97	1.491841
1216	0	36	8.00	31	-1.736111
52	0	16	9.43	31	9.941676
660	0	34	8.43	32	-0.697788
1145	0	41	8.00	71	9.146341
697	0	33	8.00	45	4.545455
560	0	21	8.00	27	3.571429

- Mean percentage benefit is 1.925015
- Is it due to chance? Or does it indicate positive impact of the novel treatment?
 - What would be the value of mean percentage benefit what if you selected another set of 10 patients?

One-Sample t-Test – Example II (cont.)

1. Check assumptions, determine H_0 and H_a , choose α
 - Normality of the variable is checked (Quantile-quantile plot)
 - $H_0: \mu = 0$ $H_a: \mu \neq 0$
 - $\alpha = 0.05$

One-Sample t-Test – Example II (cont.)

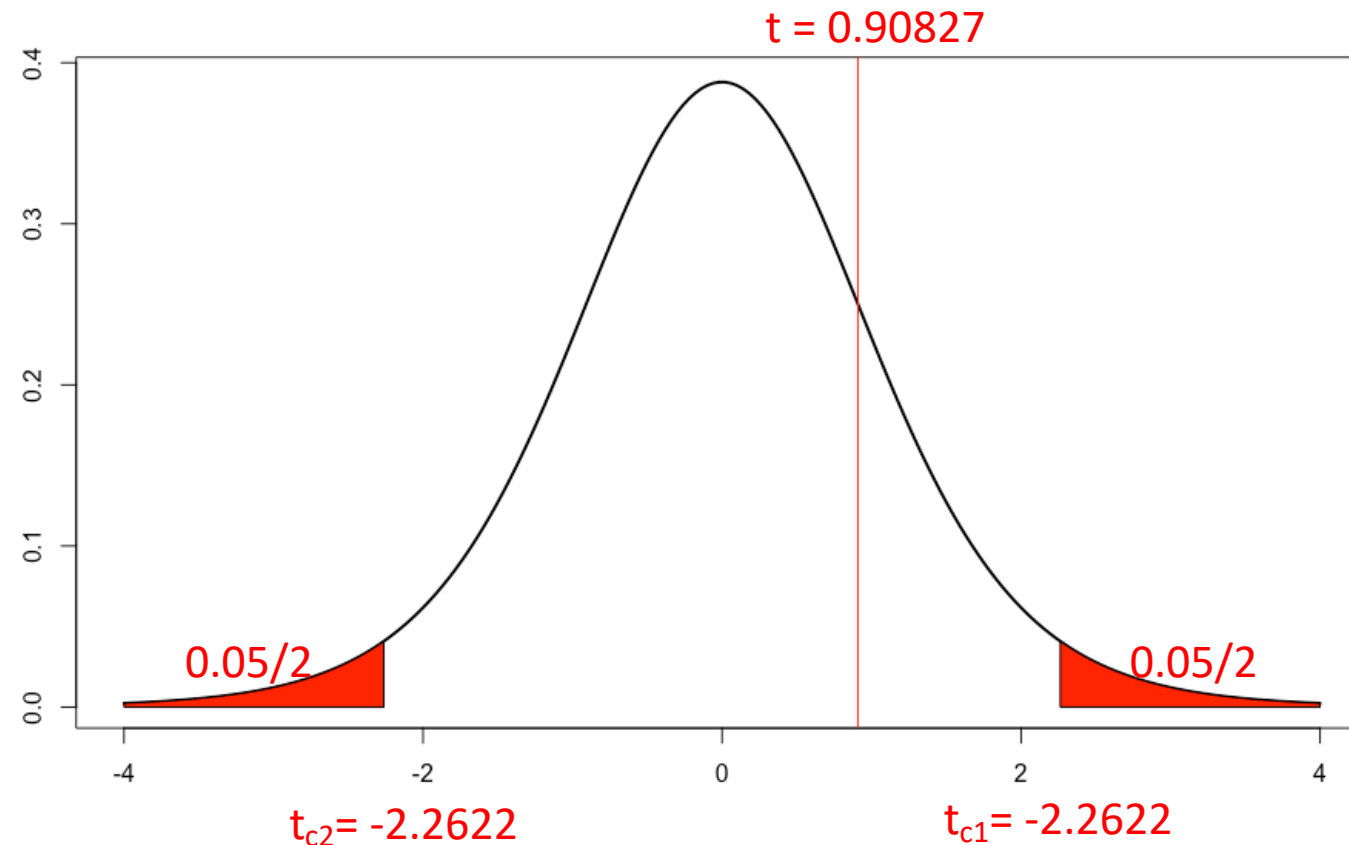
2. Calculate the appropriate test statistic

- Mean percentage benefit is 1.925015
- Standard deviation is 6.702202
- Sample size is 10

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}} = \frac{1.925015 - 0}{6.702202/\sqrt{10}} = 0.9082736 \quad (\sim t_{n-1} = t_9)$$

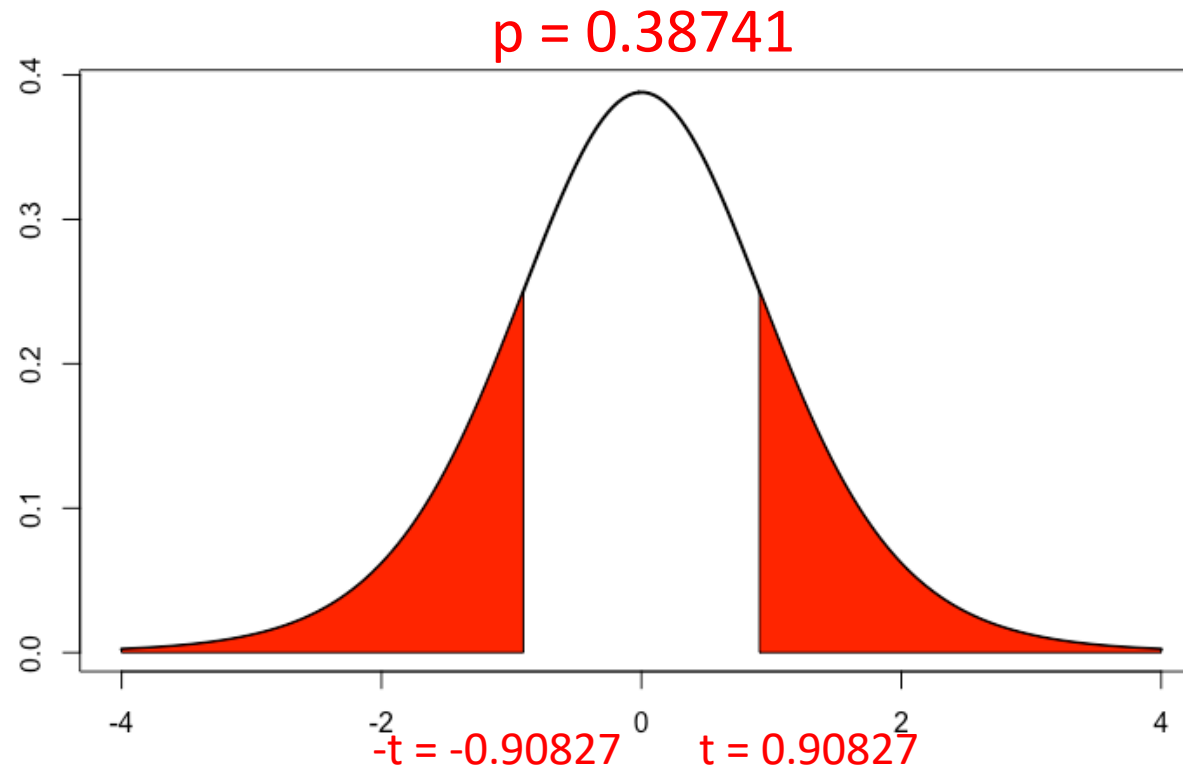
One-Sample t-Test – Example II (cont.)

3. Calculate **critical values**/p value
4. Decide whether to reject/fail to reject H_0



One-Sample t-Test – Example II (cont.)

3. Calculate critical values/**p value**
4. Decide whether to reject/fail to reject H_0



One-Sample t-Test – Example III

- It is claimed that:
- A novel drug reduces the recovery time of patients to less than 10 days
- Recovery time for 7 randomly-selected patients:
2, 4, 11, 3, 4, 6, 8 ($\bar{X} = 5.43$, $s = 3.15$)
- Test the hypothesis using $\alpha = 0.01$

One-Sample t-Test – Example III((cont.)

1. Check assumptions, determine H_0 and H_a , choose α

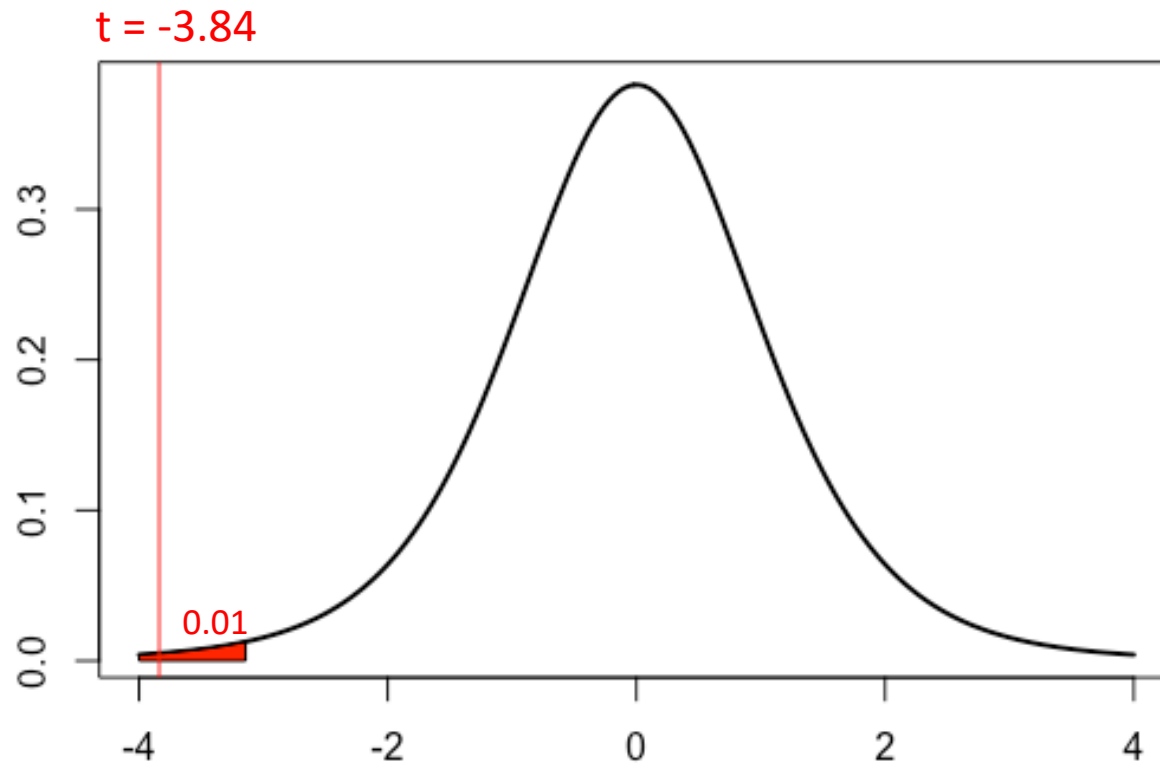
- Normality of the variable is checked
- $H_0: \mu \geq 10$ $H_a: \mu < 10$
- $\alpha = 0.01$

2. Calculate the appropriate test statistic

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}} = \frac{5.43 - 10}{3.15/\sqrt{7}} = -3.84 \quad (\sim t_{n-1} = t_6)$$

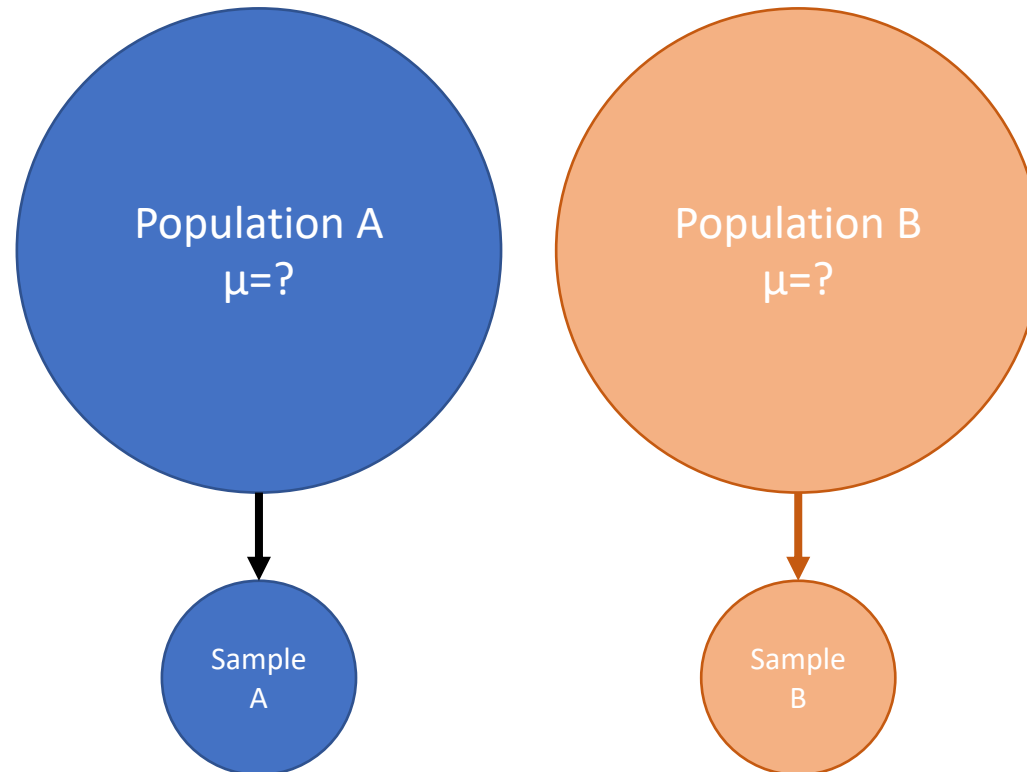
One-Sample t-Test – Example III (cont.)

3. Calculate **critical values**/p value
4. Decide whether to reject/fail to reject H_0



Two-Sample t-Test

- The **two-sample t-test** (also known as the **independent samples t-test**) is a method used to test whether the unknown population means of two groups are equal or not



Two-sample t-Test

$$H_0: \mu_X = \mu_Y$$

$$H_a: \mu_X \neq \mu_Y$$

or

$$\mathbf{H_0: \mu_X - \mu_Y = 0}$$

$$\mathbf{H_a: \mu_X - \mu_Y \neq 0}$$

Two-sample t-Test

$$\sigma_X^2 = \sigma_Y^2$$

$$T = \frac{\bar{X} - \bar{Y}}{s_P \sqrt{\frac{1}{n_X} + \frac{1}{n_Y}}} \sim t(n_X + n_Y - 2)$$
$$s_P = \frac{(n_X - 1)s_x^2 + (n_Y - 1)s_Y^2}{n_X + n_Y - 2}.$$

$$\sigma_X^2 \neq \sigma_Y^2$$

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{s_X^2}{n_X} + \frac{s_Y^2}{n_Y}}} \sim t(m),$$

$$m = \frac{(w_X + w_Y)^2}{\left(\frac{w_X^2}{n_X - 1} + \frac{w_Y^2}{n_Y - 1}\right)}$$

$$w_X = s_X^2/n_X, \quad w_Y = s_Y^2/n_Y$$

Can be checked via F-test

Two-sample t-Test – Example I

id	treatment	perc_benefit	id	treatment	perc_benefit
158	trt1	37.2549020	15	trt2	10.0978368
392	trt1	-4.3864459	143	trt2	0.5048635
457	trt1	-5.1075269	470	trt2	-0.8156940
487	trt1	36.7043369	536	trt2	50.0000000
723	trt1	5.1303099	549	trt2	-3.0303030
832	trt1	3.1806616	750	trt2	-2.8977108
894	trt1	-3.9062500	891	trt2	26.3872135
1104	trt1	5.9443608	997	trt2	4.3651179
1283	trt1	-0.8601855	1000	trt2	2.3582125
1288	trt1	-3.1674208	1209	trt2	8.9702189

- Mean percentage benefit is 7.0787 for treatment arm 1, and 9.5940 for treatment arm 2
- Is the difference a significant one?

F-test for Variance Equality

$$\begin{aligned} H_0: \sigma_X^2 &= \sigma_Y^2 \\ H_1: \sigma_X^2 &\neq \sigma_Y^2 \end{aligned}$$

or

$$H_0: \frac{\sigma_X^2}{\sigma_Y^2} = 1$$

F-test (cont.)

1. Check assumptions, determine H_0 and H_1 , choose α

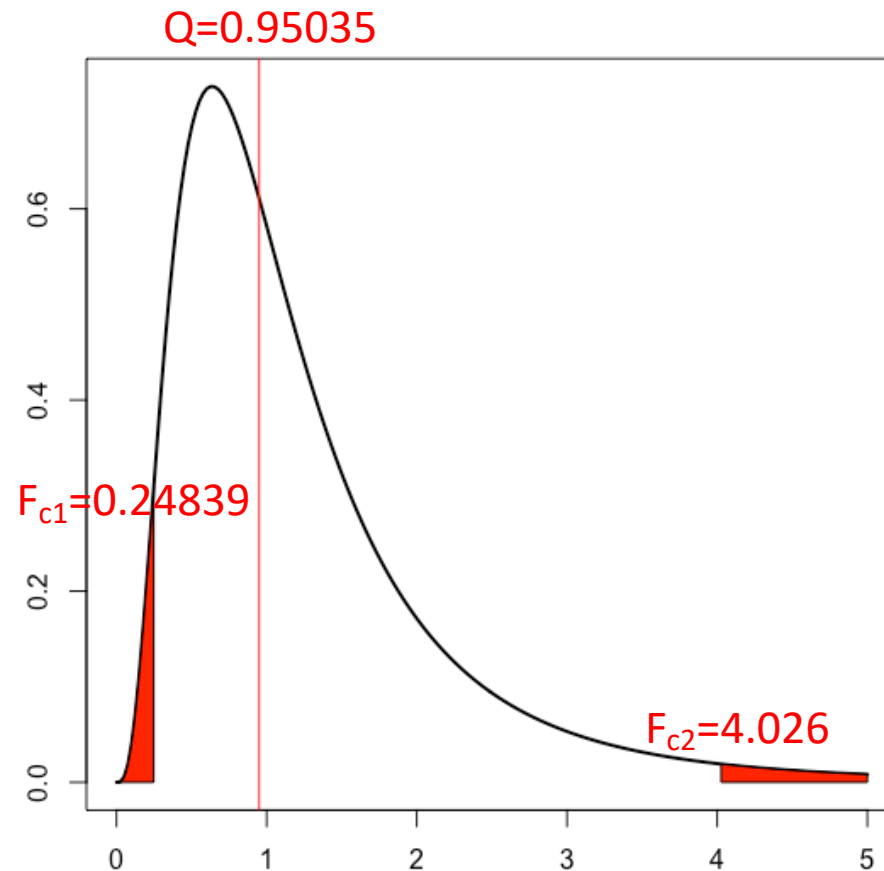
- $H_0: \frac{\sigma_X^2}{\sigma_Y^2} = 1$ $H_a: \frac{\sigma_X^2}{\sigma_Y^2} \neq 1$
- $\alpha = 0.05$

2. Calculate the appropriate test statistic

$$Q = \frac{s_X^2}{s_Y^2} = \frac{264.13}{277.93} = 0.95035 \quad \sim F_{n_X-1, n_Y-1} = F_{9,9}$$

F-test (cont.)

3. Calculate **critical region**/p value
4. Decide whether to reject/fail to reject H_0

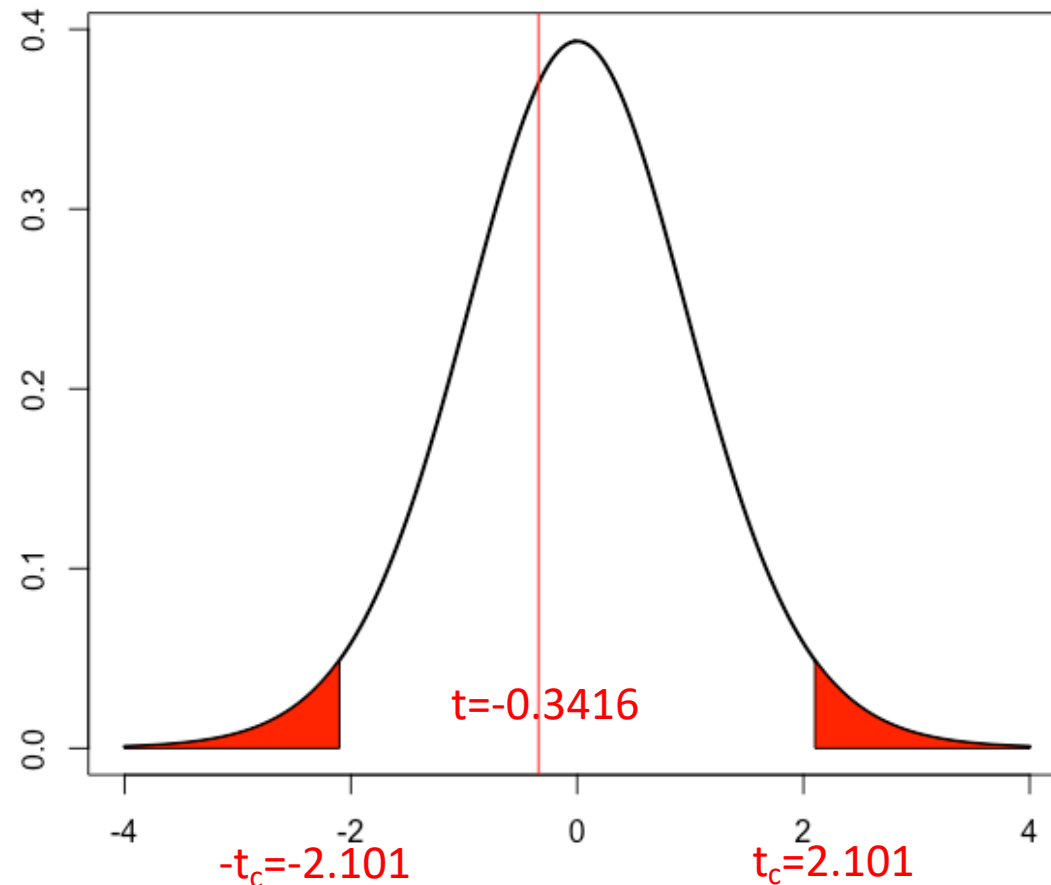


Two-sample t-Test - Example I (cont.)

1. Check assumptions, determine H_0 and H_a , choose α
 - We check that the variables are normally distributed
 - We have decided that the variances are equal
 - $H_0: \mu_1 = \mu_2$ $H_a: \mu_1 \neq \mu_2$
 - $\alpha = 0.05$
2. Calculate the appropriate test statistic
 $t = -0.3416 (\sim t_{17.98834})$

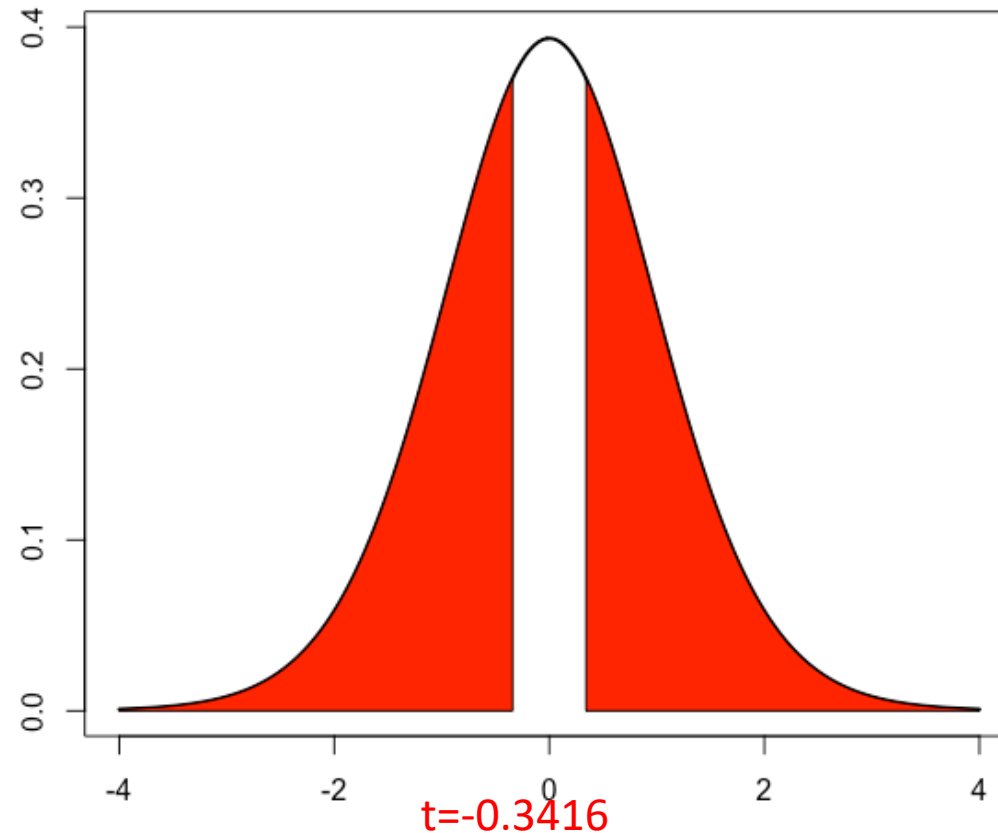
Two-sample t-Test - Example I (cont.)

3. Calculate **critical values**/p value
4. Decide whether to reject/fail to reject H_0

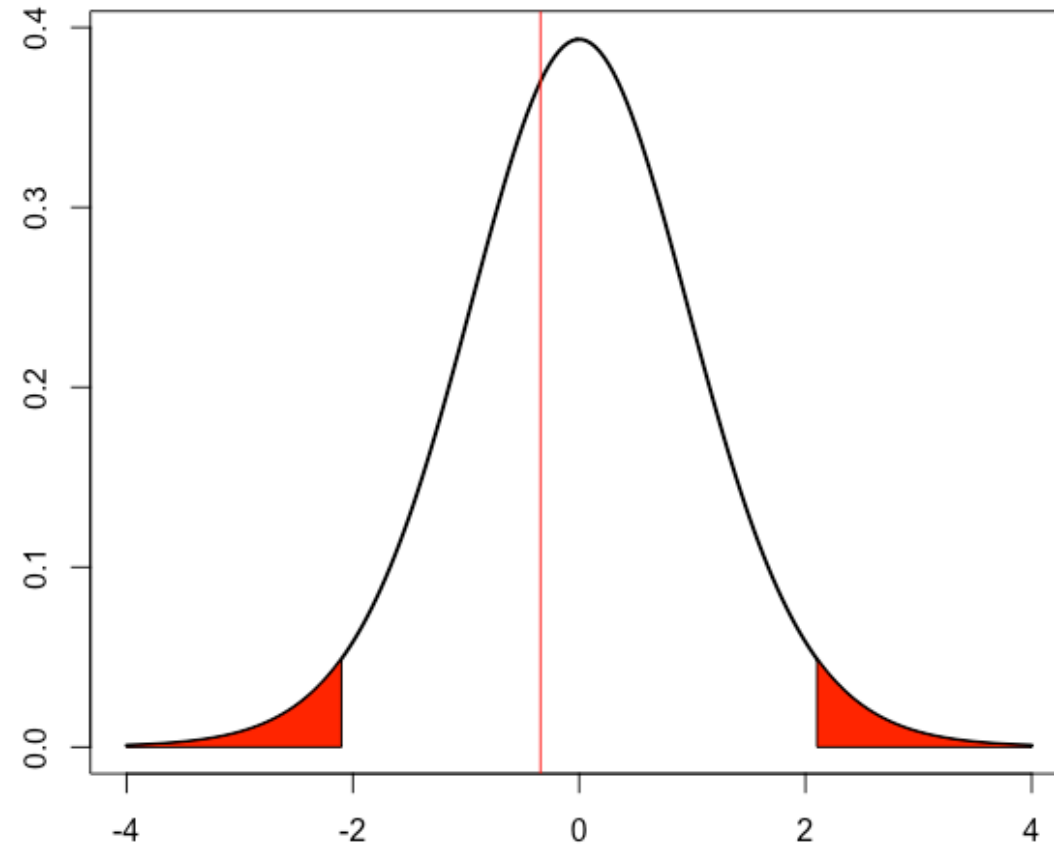


Two-sample t-Test - Example I (cont.)

3. Calculate critical values/**p value**
4. Decide whether to reject/fail to reject H_0



Two-sample t-Test - Example I (cont.)



95% confidence interval for $\mu_1 - \mu_2 = [-17.98, 12.95]$

Two-sample t-Test - Example I (cont.)

- there is not enough evidence to say mean percentage benefit for treatment 1 and treatment 2 are significantly different

Two-sample t-Test - Example II

- In a study,
 - The sedimentation rate of 12 arthritis patients was measured: $\bar{X}_1 = 82.79$ mm and $s_1 = 18.4$ mm
 - The sedimentation rate of 15 healthy controls was measured: $\bar{X}_2 = 69.03$ mm and $s_2 = 21.4$ mm
- Is there a difference between the mean sedimentation rates of the two groups?

Two-sample t-Test - Example II (cont.)

1. Check assumptions, determine H_0 and H_a , choose α

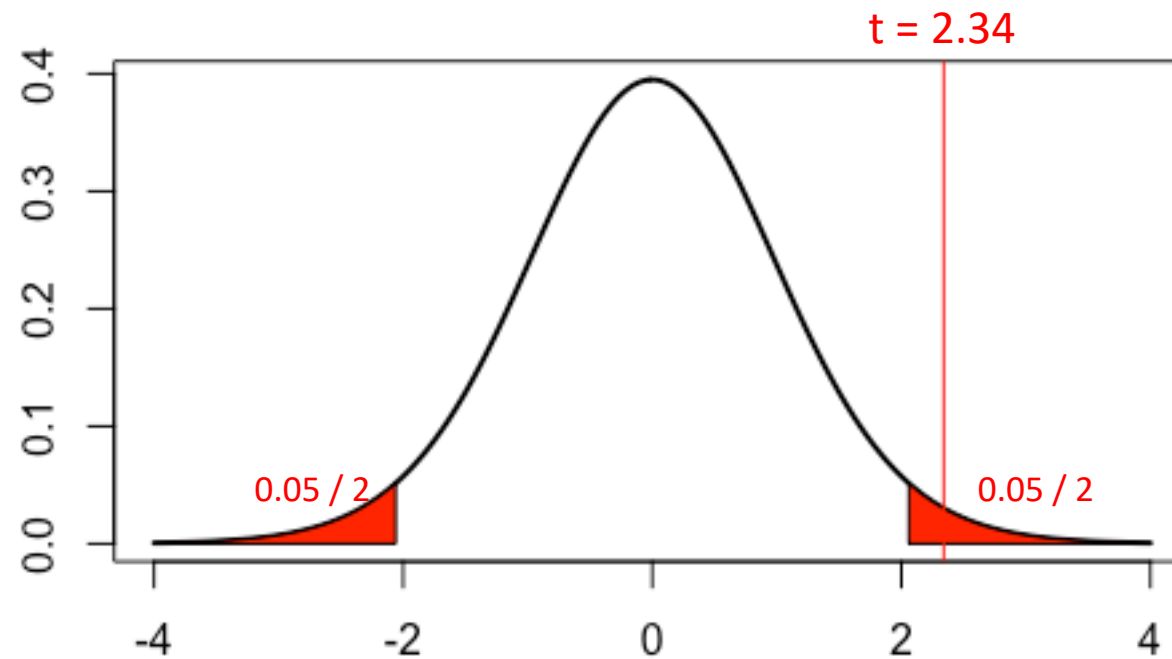
- We check that the variables are normally distributed
- $H_0: \mu_1 = \mu_2$ $H_a: \mu_1 \neq \mu_2$
- $\alpha = 0.05$

2. Calculate the appropriate test statistic

$$t = 2.34 \quad (\sim t_{25})$$

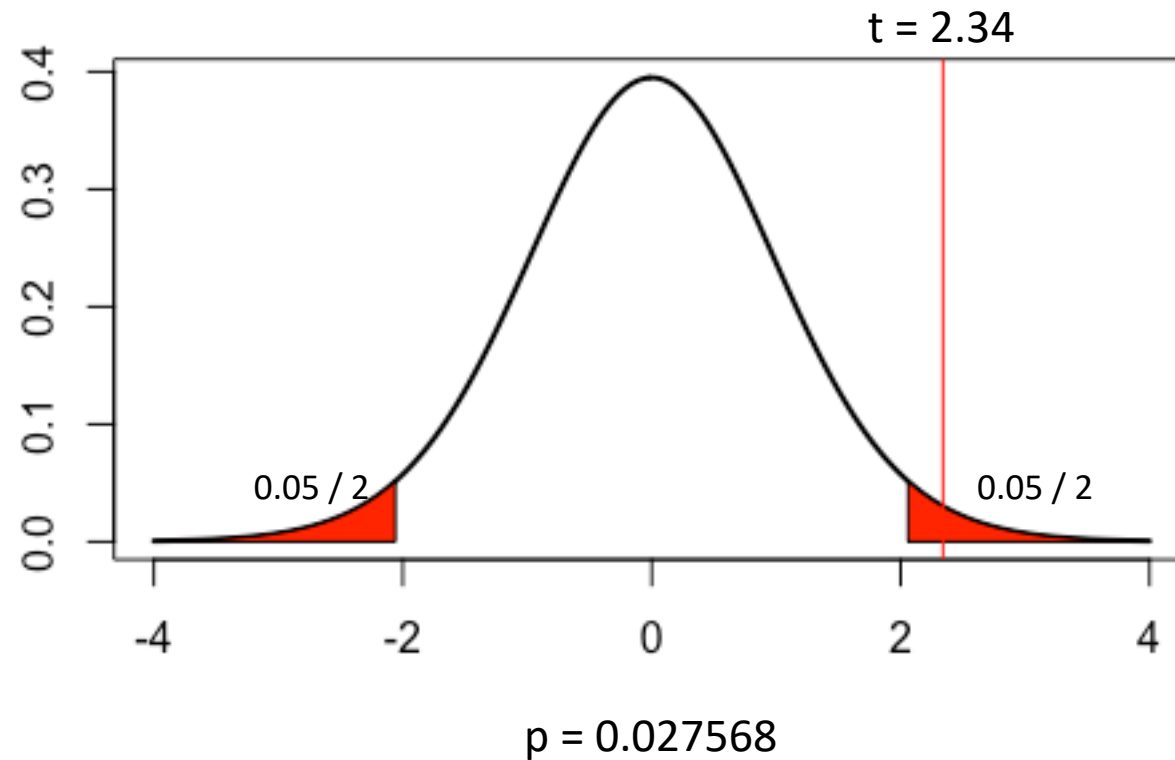
Two-sample t-Test - Example II (cont.)

3. Calculate critical values/p value
4. Decide whether to reject/fail to reject H_0



$$p = 0.027568$$

Two-sample t-Test - Example II (cont.)

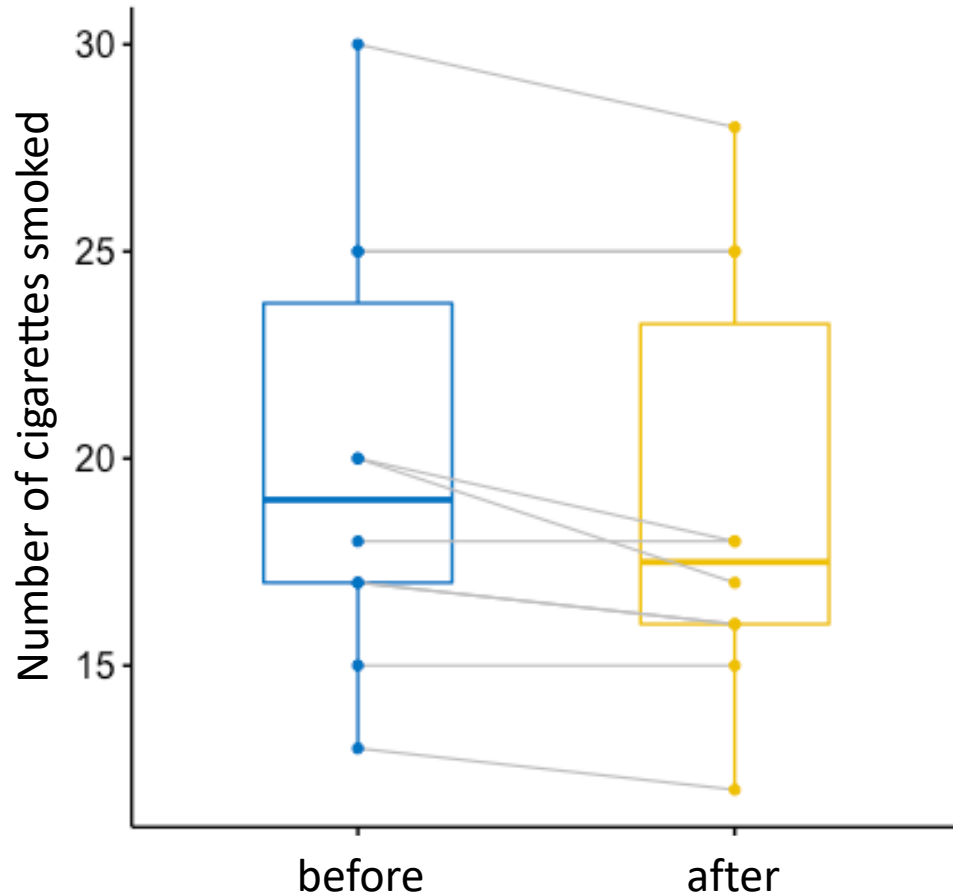


95% confidence interval for $\mu_1 - \mu_2 = [3.52, 33]$

Two-sample t-Test - Example II (cont.)

- With 95% confidence, there is enough evidence to say that there is a difference between the mean sedimentation rates of the two groups

Paired/matched t-test



$$t_h = \frac{\bar{D} - \mu_D}{S_D / \sqrt{n}} \sim t_{n-1}$$

$$D = (X_{11} - X_{21}), (X_{12} - X_{22}), \dots, (X_{1n} - X_{2n})$$

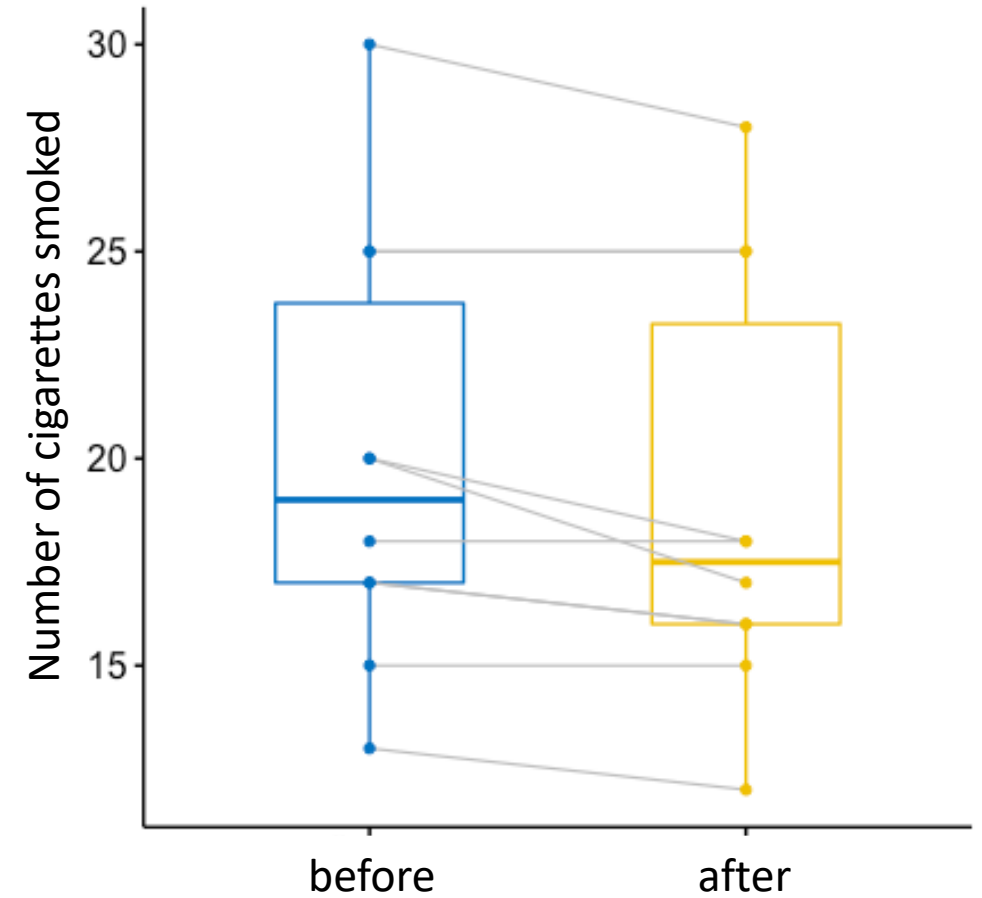
$$\bar{D} = \frac{\sum_{i=1}^n D_i}{n}, S_D = \sqrt{\frac{\sum_{i=1}^n D_i^2 - \frac{(\sum_{i=1}^n D_i)^2}{n}}{n-1}}$$

Paired t-test

- In a study investigates a seminar that aims to reduce the number of cigarettes smoked
- For 10 randomly-selected smokers,
 - The number of cigarettes smoked per day before the seminar (X_1)
 - The number of cigarettes smoked per day after the seminar (X_2)are recorded
- Can it be claimed that the seminar reduces the number of cigarettes smoked per day?

Paired t-test

X_{1i}	X_{2i}	$D_j = X_1 - X_2$
30	28	2
25	25	0
25	25	0
20	18	2
20	17	3
18	18	0
17	16	1
17	16	1
15	15	0
13	12	1



Paired t-test

1. Check assumptions, determine H_0 and H_a , choose α

- $H_0: \mu_D \leq 0$ $H_a: \mu_D > 0$ $D = X_1 - X_2$
- $\alpha = 0.05$

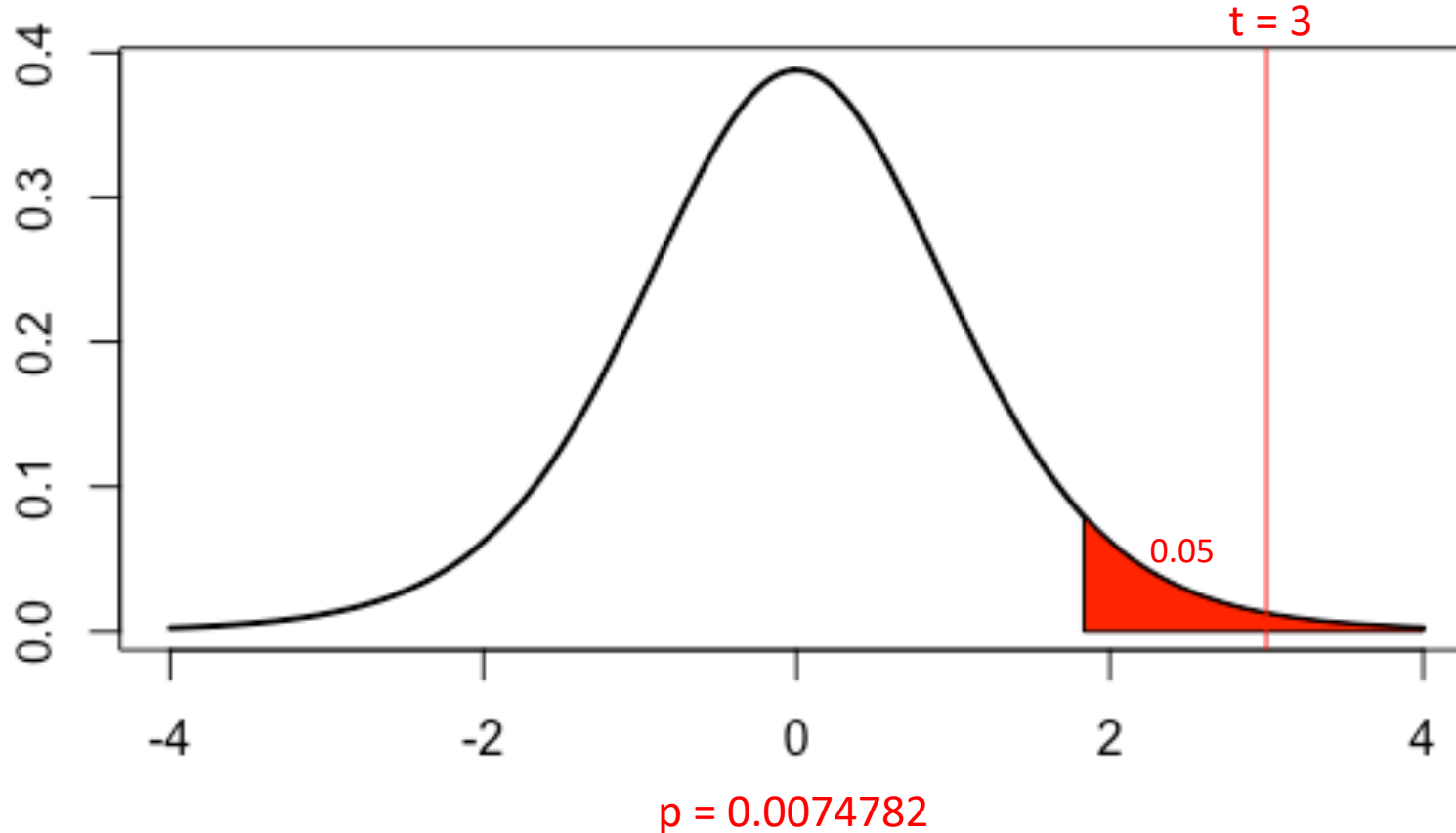
2. Calculate the appropriate test statistic

$$t = 3 \quad (\sim t_9)$$

$$D = X_1 - X_2$$

Paired t-test

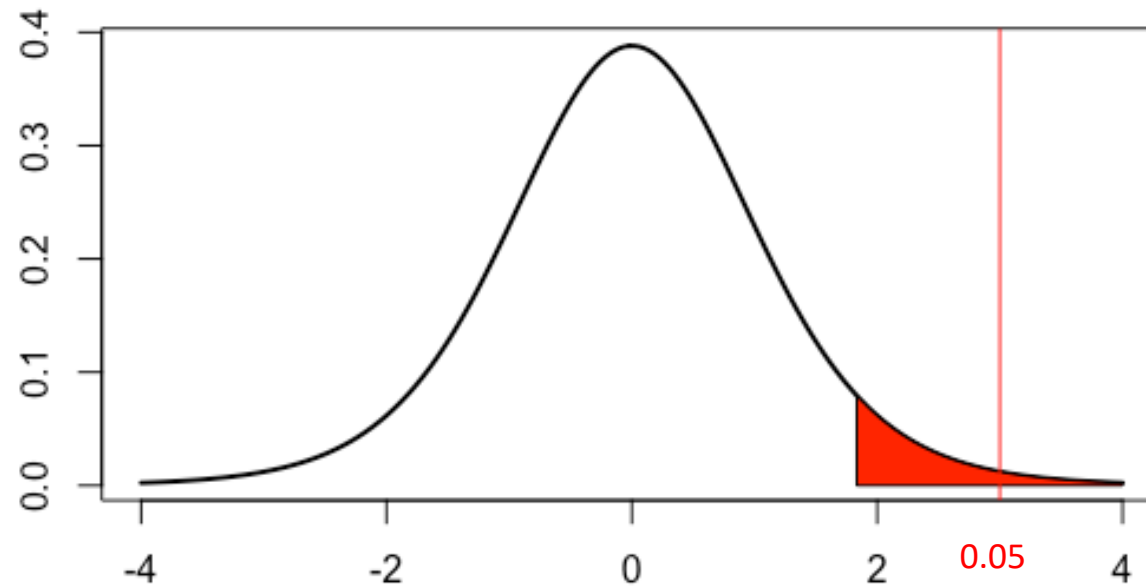
3. Determine the **critical region** / calculate the **p value** / calculate the confidence interval
4. Decide whether the null hypothesis can be rejected or not



$$D = X_1 - X_2$$

Paired t-test

3. Determine the critical region / calculate the p value / calculate **the confidence interval**
4. Decide whether the null hypothesis can be rejected or not



95% confidence interval for $\bar{D} = [0.246, \infty]$

Brief Summary

