Introduction to EXP2-1 Part 2

EXP2-1 Same or Different?

Semester B, Weeks 1-2

The **aim** of EXP2-1 is to develop tools and skills, in order to:

- Present and report data in the correct way

  (units, significant figures, error, and others)

- Determine whether the data are statistically significant

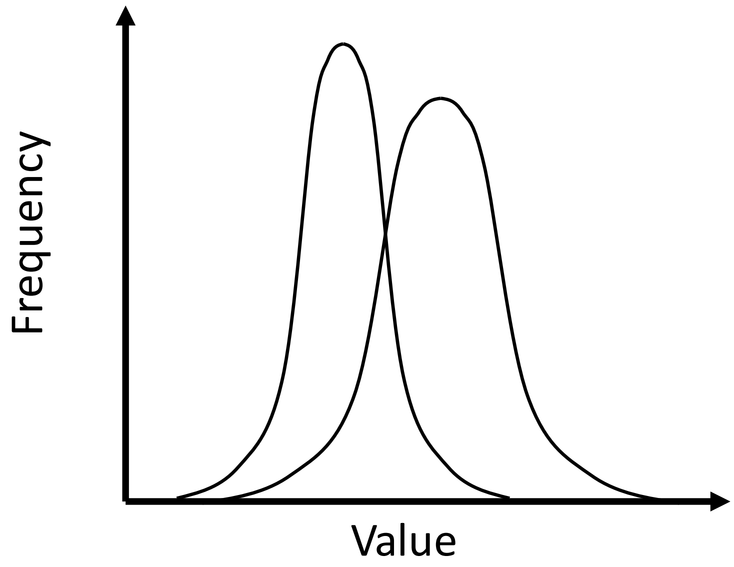
Introduction to EXP2-1 Part 2

Part 2 Significance tests

(comparing results, or samples)

Introduction to EXP2-1 Part 2

Significance tests



Are these 2 sets of data different from each other?

Significance tests

3 ways to determine differences between sets of data that vary randomly:

z-test - sample Vs. population

t-test - sample Vs. sample

F-test - samples have different variance from each other

The z-test

Is a sample different from a known population?

You can knowledge about a population from a previous sample, if you had enough data.

e.g. Are your measured values different from known values?

Used to compare new values to **a <u>large</u> existing data set**

Used in process control to see if a process has changed

Some definitions

Null hypothesis (H_0) : a type of hypothesis stating that there is no difference between two samples / sample and a population.

Alternative hypothesis (H1): The sample means / the sample and population mean are different. (The opposite region from H_0).

Significance level (α): the probability of rejecting the null hypothesis.

z standard score: a measure of how many standard deviations a sample mean is from the population mean.

α - significance level e.g. $\alpha = 0.05$, 5% level, i.e. $z = 1.96$

$\alpha = 0.01$, $z = 2.58$

$\alpha = 0.003$, $z = 3.00$

Example 1

The average weight of newborn babies is 3190 g in year 1989 with a standard deviation σ = 80. Now the average weight of 100 newborn babies from year 1990 are 3210 g.

1. Is there a significant difference in the weight of newborn between year 1989 and 1990?

μ = 3190 g (population)

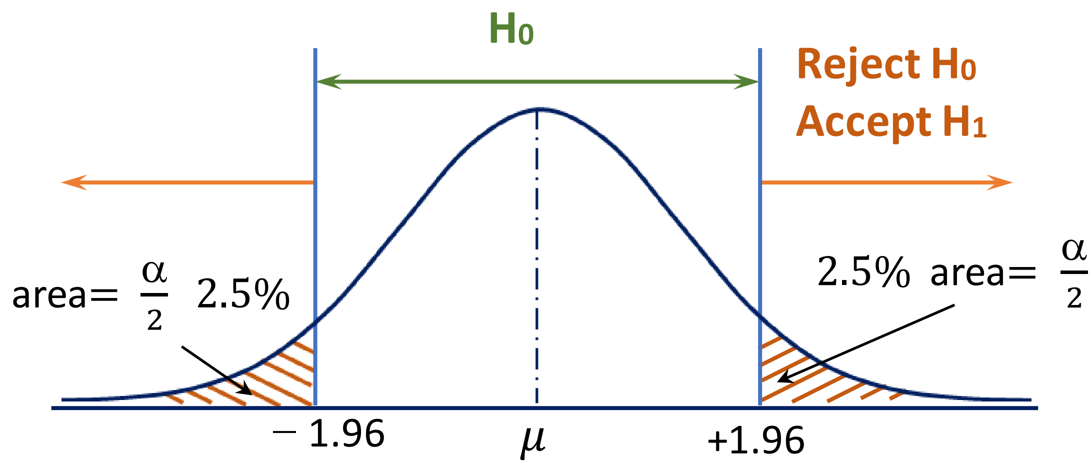
α - significance level

α α = 0.05, 5% level, i.e. z = 1.96

$\bar{X}$ **Vs.** $\boldsymbol{\mu}$

H$_0$: $\bar{X} = \mu = 3190$ g

H$_1$: $\bar{X} \neq \mu \neq 3190$ g



Two-tailed test at the 5% level

Example 1

The average weight of newborn babies is 3190 g in year 1989 with a standard deviation σ = 80. Now the average weight of 100 newborn babies from year 1990 are 3210 g.

2. Is the weight of newborn babies in year 1990 lower than that of 1989?

μ_0 = 3190 g (population)

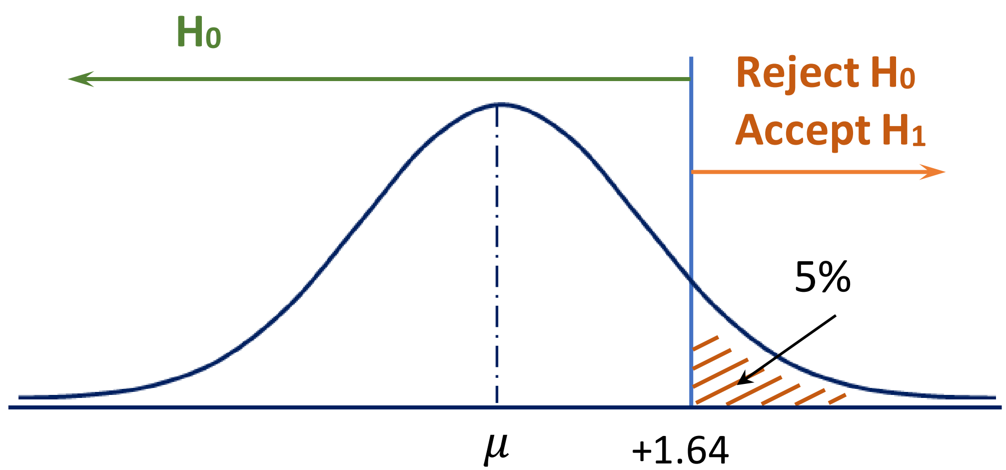
α - significance level

α α = 0.05, 5% level, i.e. z = 1.64

$\bar{X}$ **Vs.** μ

H_0: $\bar{X} < \mu$

H_1: $\bar{X} \geq \mu \neq 3190$ g



One-tailed test at the 5% level

Example 1

The average weight of newborn babies is 3190 g in year 1989 with a standard deviation σ = 80. Now the average weight of 100 newborn babies from year 1990 are 3210 g.

3. Is the weight of newborn babies in year 1990 higher than that of 1989?

μ_0 = 3190 g (population)

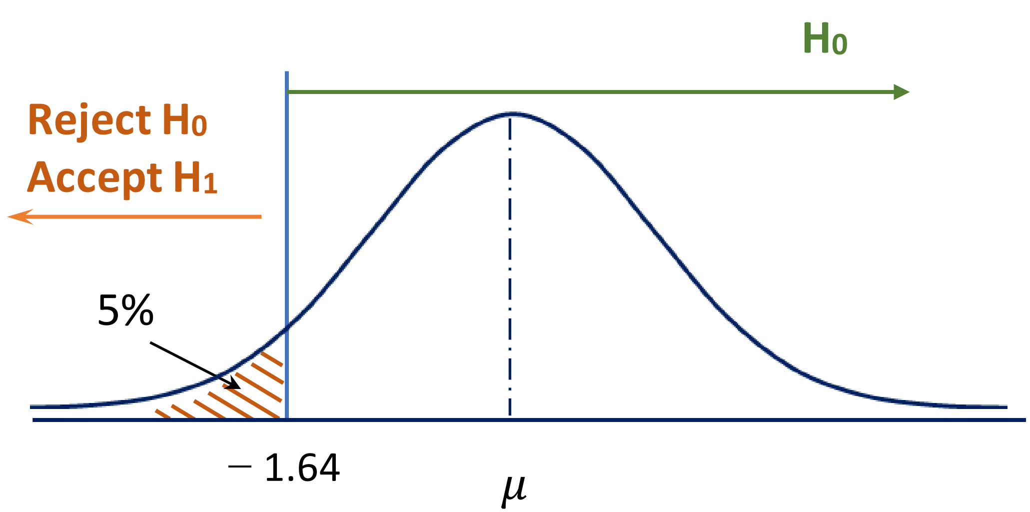
α - significance level

α α = 0.05, 5% level, i.e. z = 1.64

$\bar{X}$ **Vs.** $\boldsymbol{\mu}$

H$_0$: $\bar{X} > \mu$

H$_1$: $\mu \leq \mu_0 \neq 3190$ g



One-tailed test at the 5% level

Formula – z standard score

Mean of sample

Mean of population of comparison

z - standard score

$$z = \frac{\overline{X} - \mu}{s_{\overline{X}}}$$

SD sample x

z standard score: a measure of how many standard deviations a sample mean is from the population mean.

z > 0, sample mean > population mean;

z ~ 0, sample mean ~ population mean;

z < 0, sample mean < population mean

Example 1

The average weight of newborn babies is 3190 g in year 1989 with a standard deviation σ = 80. Now the average weight of 100 newborn babies from year 1990 are 3210 g.

1. Is there a significant difference in the weight of newborn between year 1989 and 1990?

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{3210 - 3190}{80/\sqrt{100}} = 2.5$$

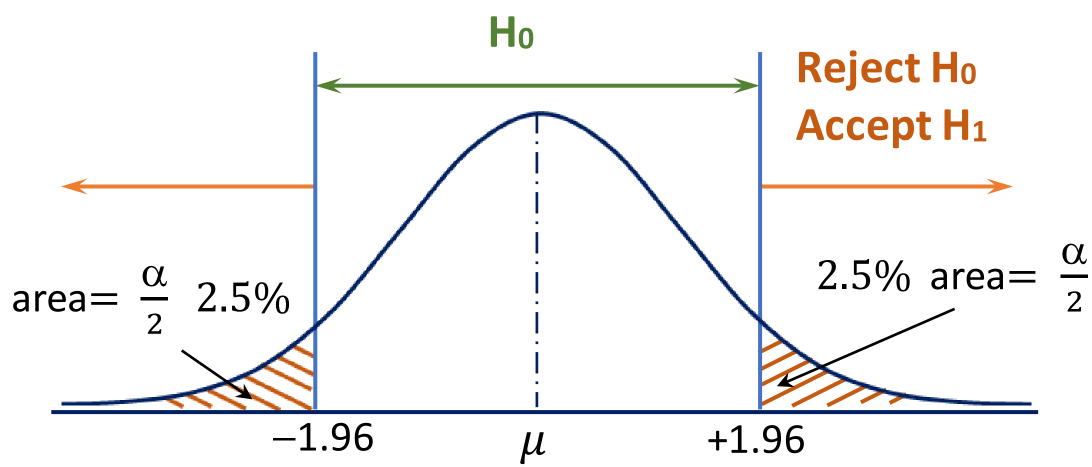
- α - significance level
- α = 0.05, 5% level, i.e. z = 1.96

$$|Z| > |Z_{\frac{\alpha}{2}}|$$

Reject H_0

Accept H_1

$\bar{X} \neq \mu \neq 3190$ g



Two-tailed test at the 5% level

Example 1

The average weight of newborn babies is 3190 g in year 1989 with a standard deviation σ = 80. Now the average weight of 100 newborn babies from year 1990 are 3210 g.

2. Is the weight of newborn babies in year 1990 lower than that of 1989?

α - significance level

α α = 0.05, 5% level, i.e. z = 1.64

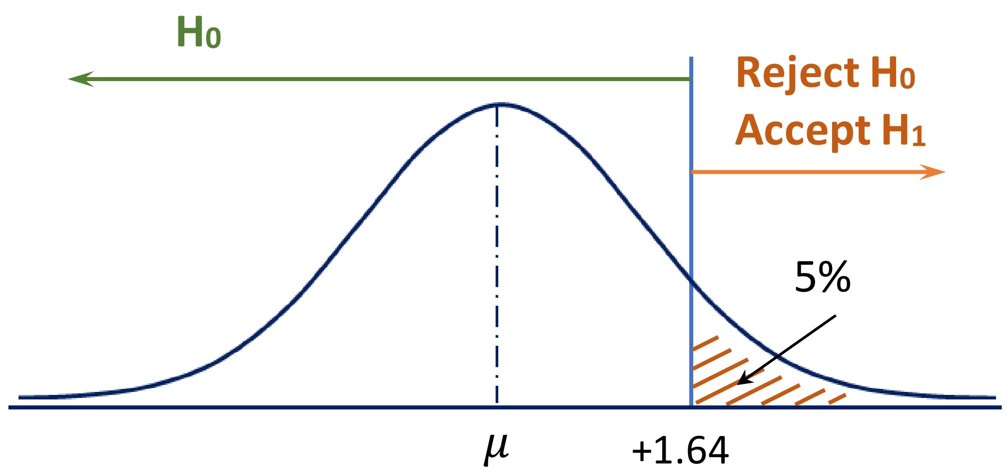
$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{3210 - 3190}{80/\sqrt{100}} = 2.5$$

$|Z| > |Z_\alpha|$

Reject H_0

Accept H_1

$\bar{X} \geq \mu \neq 3190$ g



One-tailed test at the 5% level

Example 1

The average weight of newborn babies is 3190 g in year 1989. Now the average weight of 100 newborn babies from year 1990 are 3210 g.

3. Is the weight of newborn babies in year 1990 higher than that of 1989?

$\mu_0 = 3190$ g (population)

α - significance level

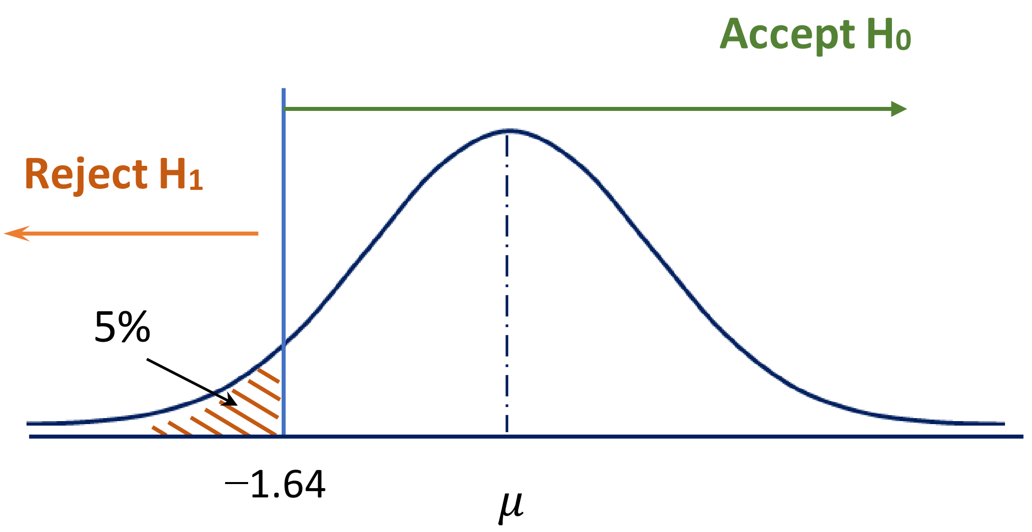
α $\alpha = 0.05$, 5% level, i.e. $z = 1.64$

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{3210 - 3190}{80/\sqrt{100}} = 2.5$$

$|Z| > |Z_\alpha|$

Accept H_0 $\bar{X} > \mu$

Reject H_1

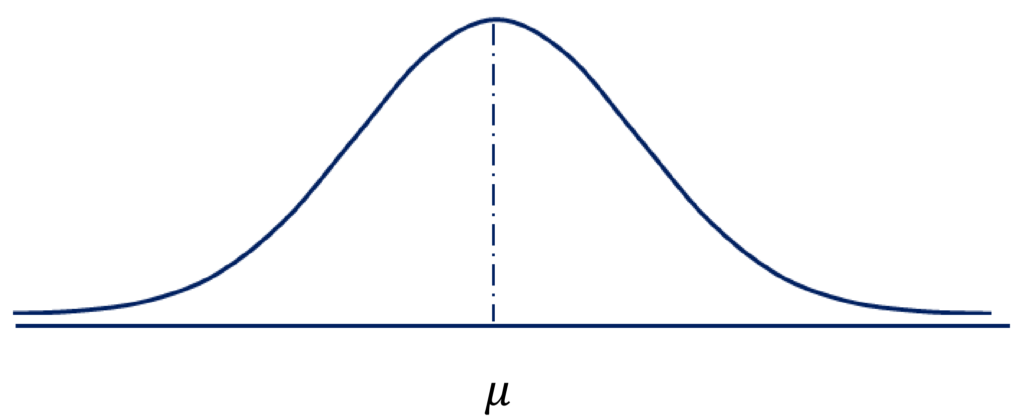


One-tailed test at the 5% level

Z-test conditions

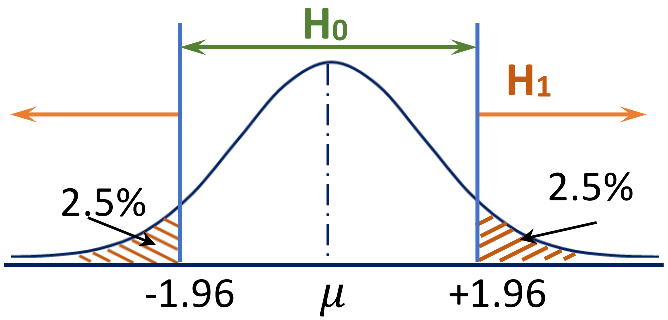
- Compare whether there is significant difference between a new data and a population when the SD of the population is known;
- Compare whether there is significant difference between the mean values of two samples (normally for large samples, n>30).

$\bar{X}$ **Vs.** $\boldsymbol{\mu}$?



Testing a null hypothesis

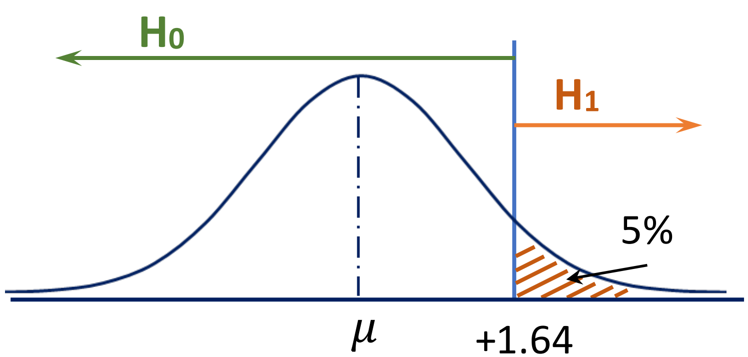
Two-tailed test



H$_0$: $\bar{X} = \mu$

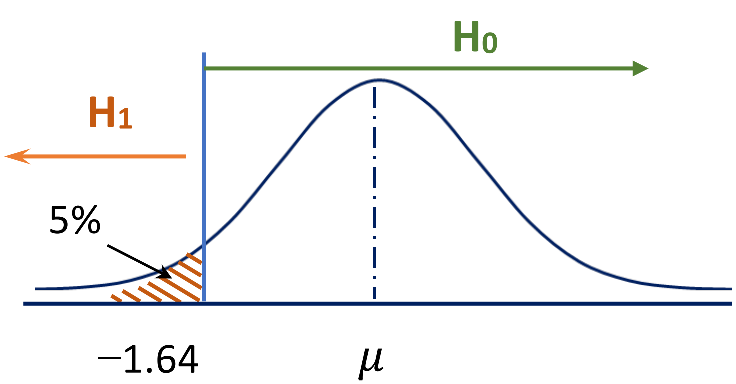
H$_1$: $\bar{X} \neq \mu$

One-tailed test



H$_0$: $\bar{X} < \mu$

H$_1$: $\bar{X} \geq \mu$



H$_0$: $\bar{X} > \mu$

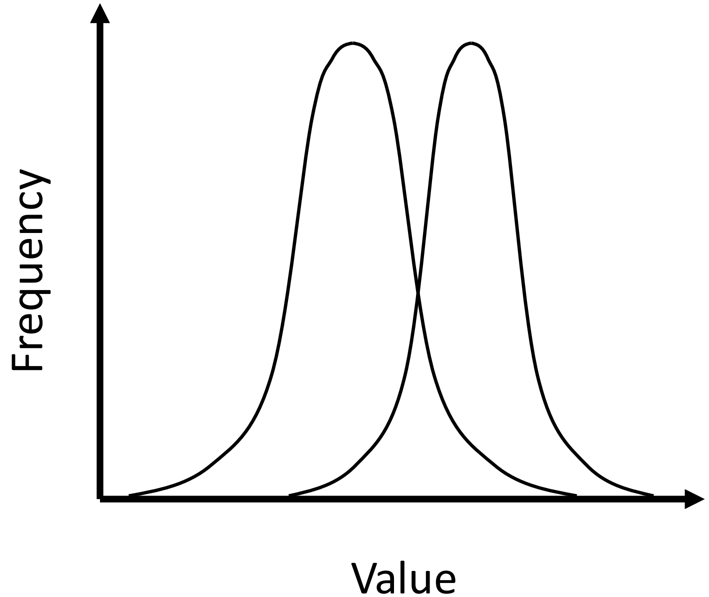
H$_1$: $\bar{X} \leq \mu_0$

$$z = \frac{\bar{X} - \mu}{s_{\bar{X}}}$$

Set a significance level, $\alpha = 0.05$, z = 1.96 (critical z)

Student’s t-test

Are two samples different from each other?



e.g. Are data from a control group and experimental group different?

Has a pharmaceutical drug had an effect?

Are men and women equally tall?

Tests

Comparing different types of variables

- One **numerical** variable (e.g. height)
- One **categorical** variable (groups A or B, women or men, etc.)
- Comparisons of means, medians and distributions
- **t tests**, Mann-Whitney, ANOVA

Two **categorical** variables:

- Cross-tabulation & proportions
- Chi-squared & McNemar tests

Two **numerical** variables:

- Scatter plots
- Regression & correlation

Student’s t-test

Research questions:

- Does the **typical** value of the **numerical variable** depend on the value of the **categorical variable**?
- Do the samples **differ** systematically in **mean** or **median** value?

There are MANY t-tests types, each one with their own equation: make sure you apply the one corresponding to the case you are studying

Two-sample t-test

Independent samples subjected to same measurement.

- **<u>Purpose</u>:** Test significance of **difference in means** of 2 samples with **numerical** variables. Samples refer to 2 different values of a **categorical** variable.

  e.g. Are men and women equally tall?

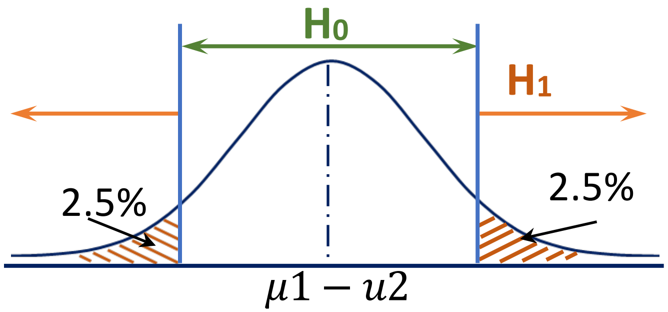
- **<u>Null hypothesis</u> (H_0):** There is no difference between the means of sample 1&2.

  There is **no** systematic **difference**. Sample means differ only by **chance**.

- **<u>Alternative hypothesis</u>** (**H_1**): The sample means are different.
  - Men and women are not equally tall
    (<u>but we're not saying who is taller</u>) (**2-tailed**)
  - Men are taller than women (**1-tailed**) OR
  - Women are taller than men (**1-tailed**)

1 and 2 tailed tests

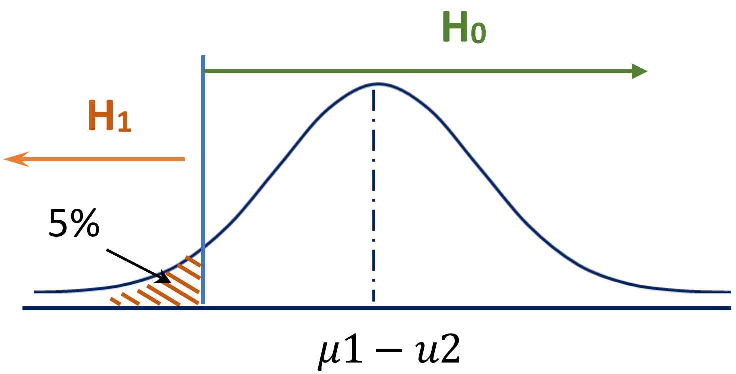
2-tailed test

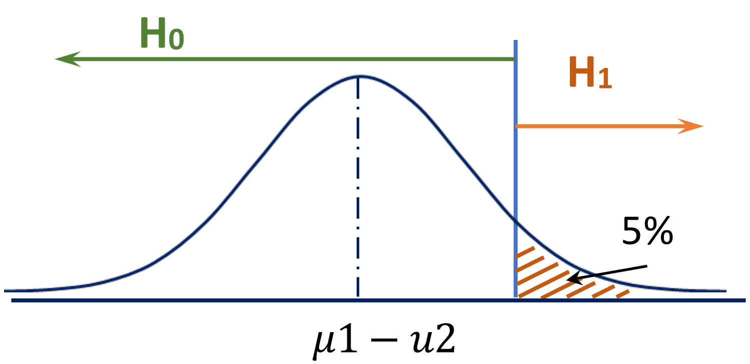


2-tailed: difference can be in either direction (or distribution tail)

$$H_0: \mu_1 = \mu_2$$

1-tailed test





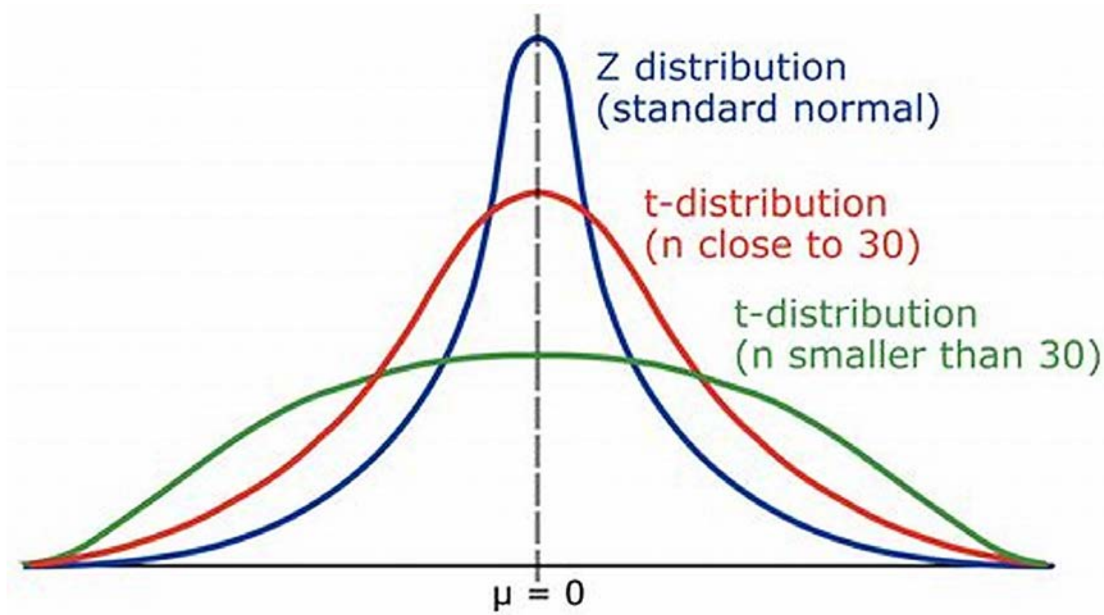
1-tailed: difference can only be in one direction (or distribution tail)

$$H_0: \mu_1 > \mu_2 \ or \ \mu_1 < \mu_2$$

Two-sample t-test

How 2-sample (independent) t-test works?

- Calculate test statistic $d/SE_{(d)}$ (**"sample t value"**) and compare with a critical t-value for the degree of freedom of the sample (related to number of measurements or data points)
- If $\mathbf{H_0}$ is true, this test statistic follows the t distribution (like a normal distribution but tails are slightly stretched out)



Formula for t statistic

Sample mean of sample 1

Sample mean of sample 2

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{(s_1^2/n_1) + (s_2^2/n_2)}}$$

Standard deviation of sample 1

Sample size of sample 1

Standard deviation of sample 2

Sample size of sample 2

Equation for t statistic: independent samples, equal variances

Some definitions

Significance level (α):

the probability of rejecting the null hypothesis.

Degree of freedom = $n1+n2-2$

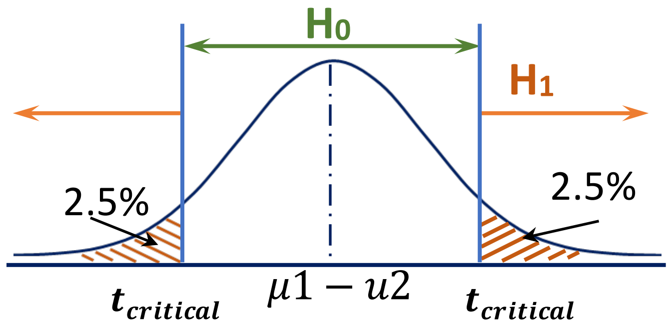
α - significance level e.g. $\alpha = 0.05$, 5% level, i.e. $z = 1.96$

$\alpha = 0.01$, $z = 2.58$

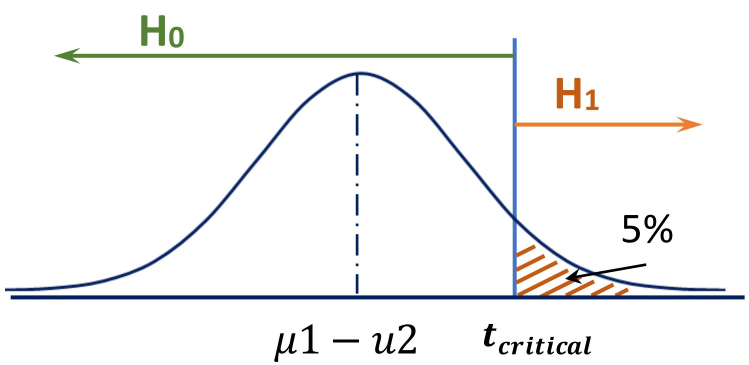
$\alpha = 0.003$, $z = 3.00$

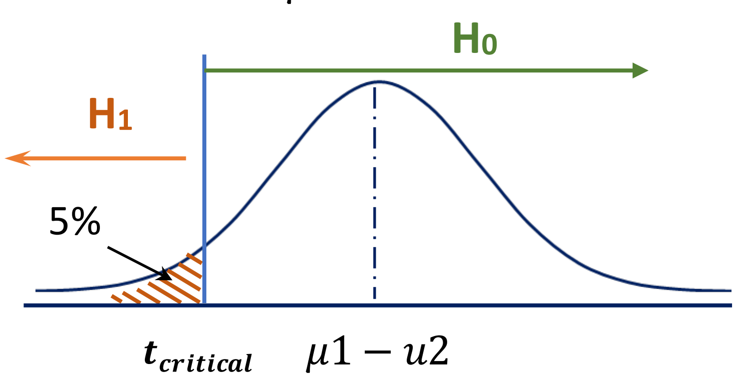
1 and 2 tailed tests

Two-tailed test



One-tailed test





$\boldsymbol{\alpha = 0.05,}$

$\boldsymbol{t = tcritical\ (95\%\ confidence)}$

2-tailed: difference can be in either direction (or distribution tail)

$\boldsymbol{Reject\ H_0\,, t > t_{critical}\ or\ t < t_{critical}}$

1-tailed: difference can only be in one direction (or distribution tail)

$\boldsymbol{Reject\ H_0\,, |t| > |tc_{ritical}|}$

Example 2

The authority would like to compare the English testing scores from two schools.

School 1

School 2

$\mu 1$ unknown

$\bar{X}1 = 86$

n1 = 46

S1 = 5.8

$\mu 2$ unknown

$\bar{X}2 = 78$

n2 = 33

S2 = 7.2

Example 2

1. Is the score of School 1 significantly different from school 2?

H$_0$: $\mu 1 = \mu 2$

H$_1$: $\mu 1 \neq \mu 2$

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{(s_1^2/n_1)+(s_2^2/n_2)}} = \frac{86-78}{\sqrt{(5.8^2\ /46)+(7.2^2\ /33)}} = 5.275$$

Df = 46+33-2 = 77

Critical values of Student's *t* distribution with *ν* degrees of freedom

Probability less than the critical value ($t_{1-\alpha,\nu}$)

v	0.90	0.95	0.975	0.99	0.995	0.999
1.	3.078	6.314	12.706	31.821	63.657	318.313
2.	1.886	2.920	4.303	6.965	9.925	22.327
3.	1.638	2.353	3.182	4.541	5.841	10.215
4.	1.533	2.132	2.776	3.747	4.604	7.173
5.	1.476	2.015	2.571	3.365	4.032	5.893
6.	1.440	1.943	2.447	3.143	3.707	5.208
7.	1.415	1.895	2.365	2.998	3.499	4.782
8.	1.397	1.860	2.306	2.896	3.355	4.499

Example 2

1. Is the score of School 1 significantly different from school 2?

H_0: $\mu 1 = \mu 2$ t =5.275

H_1: $\mu 1 \neq \mu 2$ Df = 46+33-2 = 77

2 tailed $\alpha = 0.05 = 5\%$

$\boldsymbol{Degrees\ of\ freedom\ 77}$

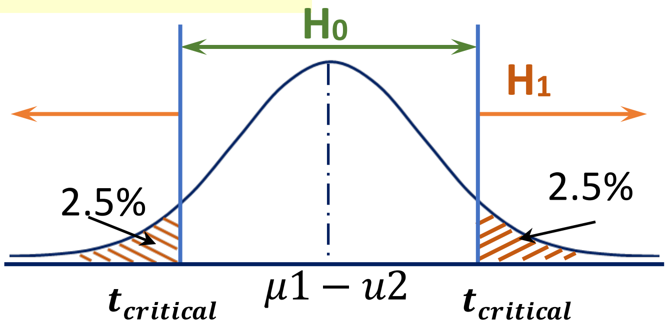
$\boldsymbol{2 - tailed\ t_{critical} = 1.991}$

Critical values of Student's *t* distribution with *ν* degrees of freedom

Probability less than the critical value ($t_{1-\alpha,\nu}$)

ν	0.90	0.95	0.975	0.99	0.995	0.999
77.	1.293	1.665	1.991	2.376	2.641	3.199
78.	1.292	1.665	1.991	2.375	2.640	3.198
79.	1.292	1.664	1.990	2.374	2.640	3.197
80.	1.292	1.664	1.990	2.374	2.639	3.195

$\boldsymbol{Reject\ H_0, t > t_{critical}}$



Example 2

2. Is the score of School 1 lower than that of school 2?

H_0: $\mu 1 > \mu 2$

H_1: $\mu 1 < \mu 2$

t =5.275

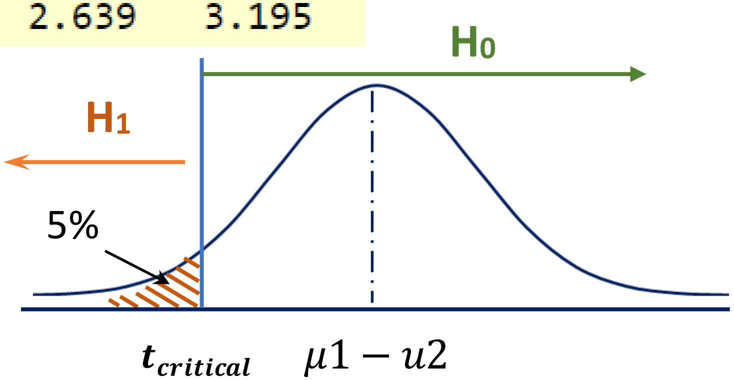
Df = 46+33-2 = 77

Critical values of Student's *t* distribution with *v* degrees of freedom

Probability less than the critical value ($t_{1-\alpha,v}$)

v	0.90	0.95	0.975	0.99	0.995	0.999
77.	1.293	$\boxed{1.665}$	1.991	2.376	2.641	3.199
78.	1.292	1.665	1.991	2.375	2.640	3.198
79.	1.292	1.664	1.990	2.374	2.640	3.197
80.	1.292	1.664	1.990	2.374	2.639	3.195

$Accept\ H_0, t > t_{critical}$



Example 2

3. Is the score of School 1 higher than that of school 2?

H$_0$: $\mu 1 < \mu 2$

H$_1$: $\mu 1 > \mu 2$

t =5.275

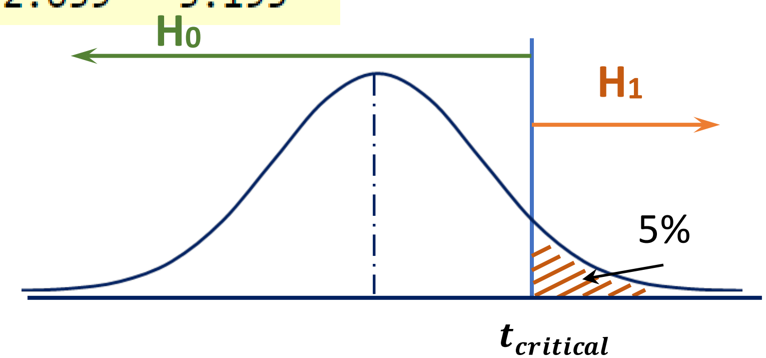
Df = 46+33-2 = 77

Critical values of Student's *t* distribution with *v* degrees of freedom

Probability less than the critical value ($t_{1-\alpha,v}$)

v	0.90	0.95	0.975	0.99	0.995	0.999
77.	1.293	$\boxed{1.665}$	1.991	2.376	2.641	3.199
78.	1.292	1.665	1.991	2.375	2.640	3.198
79.	1.292	1.664	1.990	2.374	2.640	3.197
80.	1.292	1.664	1.990	2.374	2.639	3.195

Reject H_0 , $t > t_{critical}$



Critical Values of the Student's t Distribution

Taken from: https://www.itl.nist.gov/div898/handbook/eda/section3/eda3672.htm

The table in following pages contains critical values of the Student's t distribution. The t distribution is symmetric so that **$t_{1-\alpha,\nu} = -t_{\alpha,\nu}$**

The t table can be used for both one-tailed (also called one-sided), either lower or upper, and two-tailed (also two-sided) tests using the appropriate value of α.

The difference lies in finding out whether you are conducting a one-tailed test, or two-tailed test. For a one-tailed test, use the 1-α column (for example "0.95" for 95% confidence, α=0.05); for a two-tailed test, use the 1-α/2 column (for example "0.975" for 95% confidence, α=0.05) vs **the degrees of freedom for equal variance (= number of events-2)** for the t-critical values.

Decide whether you will be conducting a one-tailed or two-tailed test.
Calculate the t statistic for your samples. Find the critical t value and compare.
If t > critical t, then you reject the null hypothesis H_0, gaining confidence in your alternative hypothesis.

Question: what's the relevance of negative and positive values of t?

Critical values of Student's *t* distribution with *v* degrees of freedom

Probability less than the critical value ($t_{1-\alpha,v}$)

v	0.90	0.95	0.975	0.99	0.995	0.999
1.	3.078	6.314	12.706	31.821	63.657	318.313
2.	1.886	2.920	4.303	6.965	9.925	22.327
3.	1.638	2.353	3.182	4.541	5.841	10.215
4.	1.533	2.132	2.776	3.747	4.604	7.173
5.	1.476	2.015	2.571	3.365	4.032	5.893
6.	1.440	1.943	2.447	3.143	3.707	5.208
7.	1.415	1.895	2.365	2.998	3.499	4.782
8.	1.397	1.860	2.306	2.896	3.355	4.499
9.	1.383	$\boxed{1.833}$	2.262	2.821	3.250	4.296
10.	1.372	1.812	2.228	2.764	3.169	4.143
11.	1.363	1.796	2.201	2.718	3.106	4.024
12.	1.356	1.782	2.179	2.681	3.055	3.929
13.	1.350	1.771	2.160	2.650	3.012	3.852
14.	1.345	1.761	2.145	2.624	2.977	3.787
15.	1.341	1.753	2.131	2.602	2.947	3.733
16.	1.337	1.746	2.120	2.583	2.921	3.686
17.	1.333	1.740	2.110	2.567	2.898	3.646
18.	1.330	1.734	2.101	2.552	2.878	3.610
19.	1.328	1.729	2.093	2.539	2.861	3.579
20.	1.325	1.725	2.086	2.528	2.845	3.552
21.	1.323	1.721	2.080	2.518	2.831	3.527
22.	1.321	1.717	2.074	2.508	2.819	3.505
23.	1.319	1.714	2.069	2.500	2.807	3.485

$\alpha = 0.05$

$Degrees\ of\ freedom\ 9$

$1 - tailed\ t_{critical} = 1.833$

Critical values of Student's *t* distribution with *v* degrees of freedom

Probability less than the critical value ($t_{1-\alpha,\nu}$)

v	0.90	0.95	0.975	0.99	0.995	0.999
24.	1.318	1.711	2.064	2.492	2.797	3.467
25.	1.316	1.708	2.060	2.485	2.787	3.450
26.	1.315	1.706	2.056	2.479	2.779	3.435
27.	1.314	1.703	2.052	2.473	2.771	3.421
28.	1.313	1.701	2.048	2.467	2.763	3.408
29.	1.311	1.699	2.045	2.462	2.756	3.396
30.	1.310	1.697	2.042	2.457	2.750	3.385
31.	1.309	1.696	2.040	2.453	2.744	3.375
32.	1.309	1.694	2.037	2.449	2.738	3.365
33.	1.308	1.692	2.035	2.445	2.733	3.356
34.	1.307	1.691	2.032	2.441	2.728	3.348
35.	1.306	1.690	2.030	2.438	2.724	3.340
36.	1.306	1.688	2.028	2.434	2.719	3.333
37.	1.305	1.687	2.026	2.431	2.715	3.326
38.	1.304	1.686	2.024	2.429	2.712	3.319
39.	1.304	1.685	2.023	2.426	2.708	3.313
40.	1.303	1.684	2.021	2.423	2.704	3.307
41.	1.303	1.683	2.020	2.421	2.701	3.301
42.	1.302	1.682	2.018	2.418	2.698	3.296
43.	1.302	1.681	2.017	2.416	2.695	3.291
44.	1.301	1.680	2.015	2.414	2.692	3.286
45.	1.301	1.679	2.014	2.412	2.690	3.281
46.	1.300	1.679	2.013	2.410	2.687	3.277
47.	1.300	1.678	2.012	2.408	2.685	3.273

v	0.90	0.95	0.975	0.99	0.995	0.999
48.	1.299	1.677	2.011	2.407	2.682	3.269
49.	1.299	1.677	2.010	2.405	2.680	3.265
50.	1.299	1.676	2.009	2.403	2.678	3.261
51.	1.298	1.675	2.008	2.402	2.676	3.258
52.	1.298	1.675	2.007	2.400	2.674	3.255
53.	1.298	1.674	2.006	2.399	2.672	3.251
54.	1.297	1.674	2.005	2.397	2.670	3.248
55.	1.297	1.673	2.004	2.396	2.668	3.245
56.	1.297	1.673	2.003	2.395	2.667	3.242
57.	1.297	1.672	2.002	2.394	2.665	3.239
58.	1.296	1.672	2.002	2.392	2.663	3.237
59.	1.296	1.671	2.001	2.391	2.662	3.234
60.	1.296	1.671	2.000	2.390	2.660	3.232
61.	1.296	1.670	2.000	2.389	2.659	3.229
62.	1.295	1.670	1.999	2.388	2.657	3.227
63.	1.295	1.669	1.998	2.387	2.656	3.225
64.	1.295	1.669	1.998	2.386	2.655	3.223
65.	1.295	1.669	1.997	2.385	2.654	3.220
66.	1.295	1.668	1.997	2.384	2.652	3.218
67.	1.294	1.668	1.996	2.383	2.651	3.216
68.	1.294	1.668	1.995	2.382	2.650	3.214
69.	1.294	1.667	1.995	2.382	2.649	3.213
70.	1.294	1.667	1.994	2.381	2.648	3.211
71.	1.294	1.667	1.994	2.380	2.647	3.209
72.	1.293	1.666	1.993	2.379	2.646	3.207
73.	1.293	1.666	1.993	2.379	2.645	3.206
74.	1.293	1.666	1.993	2.378	2.644	3.204
75.	1.293	1.665	1.992	2.377	2.643	3.202
76.	1.293	1.665	1.992	2.376	2.642	3.201

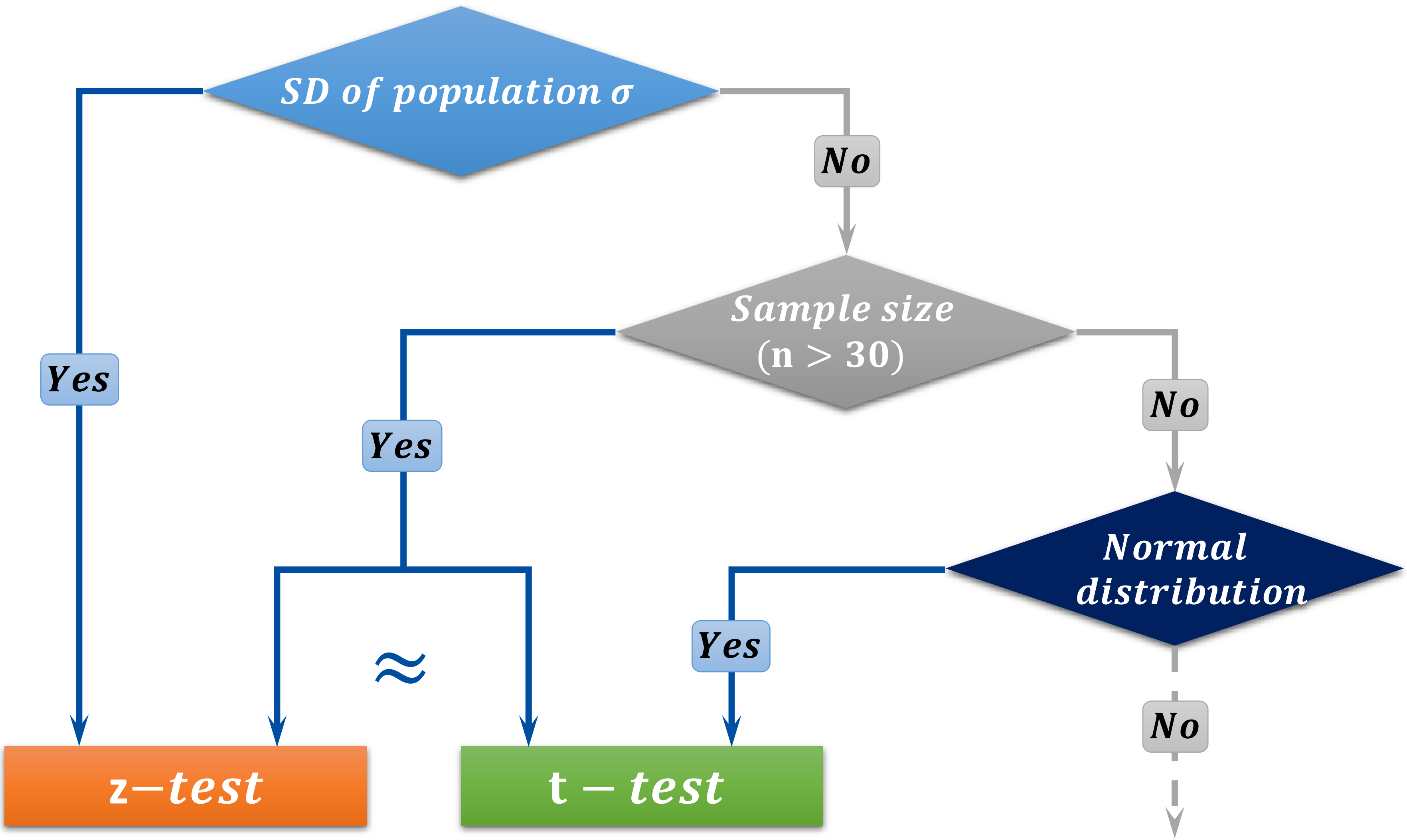
v	0.90	0.95	0.975	0.99	0.995	0.999
77.	1.293	1.665	1.991	2.376	2.641	3.199
78.	1.292	1.665	1.991	2.375	2.640	3.198
79.	1.292	1.664	1.990	2.374	2.640	3.197
80.	1.292	1.664	1.990	2.374	2.639	3.195
81.	1.292	1.664	1.990	2.373	2.638	3.194
82.	1.292	1.664	1.989	2.373	2.637	3.193
83.	1.292	1.663	1.989	2.372	2.636	3.191
84.	1.292	1.663	1.989	2.372	2.636	3.190
85.	1.292	1.663	1.988	2.371	2.635	3.189
86.	1.291	1.663	1.988	2.370	2.634	3.188
87.	1.291	1.663	1.988	2.370	2.634	3.187
88.	1.291	1.662	1.987	2.369	2.633	3.185
89.	1.291	1.662	1.987	2.369	2.632	3.184
90.	1.291	1.662	1.987	2.368	2.632	3.183
91.	1.291	1.662	1.986	2.368	2.631	3.182
92.	1.291	1.662	1.986	2.368	2.630	3.181
93.	1.291	1.661	1.986	2.367	2.630	3.180
94.	1.291	1.661	1.986	2.367	2.629	3.179
95.	1.291	1.661	1.985	2.366	2.629	3.178
96.	1.290	1.661	1.985	2.366	2.628	3.177
97.	1.290	1.661	1.985	2.365	2.627	3.176
98.	1.290	1.661	1.984	2.365	2.627	3.175
99.	1.290	1.660	1.984	2.365	2.626	3.175
100.	1.290	1.660	1.984	2.364	2.626	3.174
∞	1.282	1.645	1.960	2.326	2.576	3.090

T-test conditions

t-test **assumes** that the sample means follows a **normal distribution** when **H_0** is true:

- Valid if **population is normal distribution** (in which case samples will appear approximately normal)
- Also valid for **non-normal** populations (and samples) **if** sample size is **large** (> 30 per sample)
- Should **not** be used if samples are both small & non-normal.
- Unreliable if samples are too small to tell whether normally distributed.

z-test OR t-test?



Summary

z-test - sample different from a population

(are the new results different from the ‘norm’?)

t-test - samples different from each other

(are two samples different from the each other?)