Hypothesis testing

E. Papageorgiou, G. Katsouleas

University of West Attica

June 17, 2024

Null and alternative hypotheses

- Hypothesis testing is a data-based decision procedure that can produce a conclusion about some system.
- A statistical hypothesis is an assertion or conjecture concerning one or more populations.
- The null hypothesis, denoted by H_0, is a tentative assumption about some population parameter.
- The alternative hypothesis, denoted by H_1, is the opposite of what is stated in the null hypothesis. (usually set to what the test is attempting to establish):

$$\text{2-sided} \begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta \neq \theta_0 \end{cases} , \quad \text{left-sided} \begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta < \theta_0 \end{cases} , \quad \text{right-sided} \begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta > \theta_0 \end{cases} ,$$

- The conclusion that the research hypothesis is true is based on sample data that contradict the null hypothesis and is not reached with absolute certainty.

Type I and Type II errors

- Because hypothesis tests are based on sample data, we must allow for the possibility of errors.
- A Type I error is rejecting H_0 when it is true. The probability of making a Type I error when the null hypothesis is true as an equality is called level of significance:

$$\alpha = P(\text{reject } H_0 | H_0 \text{ true})$$

- Applications of hypothesis testing that only control the Type I error are often called significance tests.
- A Type II error is accepting β_0 when it is false:

$$\beta = P(\text{accept } H_0 | H_0 \text{ false})$$

- It is difficult to control for the probability of making a Type II error. Statisticians avoid the risk of making a Type II error by using "do not reject H_0" and not "accept H_0".
- Power of a test is defined by

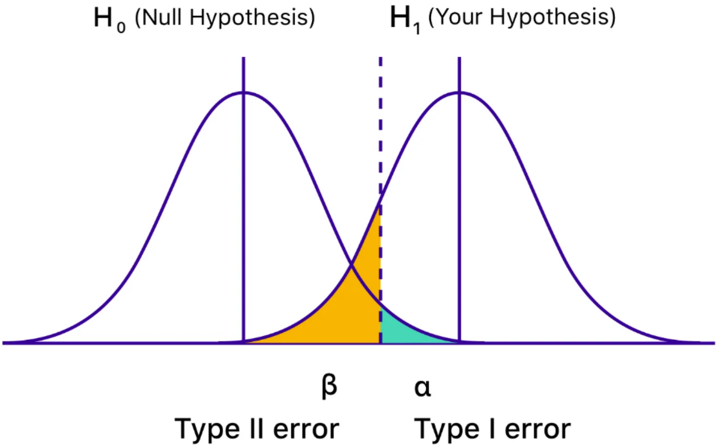
$$\gamma = 1 - \beta = P(\text{reject } H_0 | H_0 \text{ false})$$

- Computation of β and γ depends on the actual value of a population parameter.

Possible decisions of an analyst

	Null hypothesis is TRUE	Null hypothesis is FALSE
Reject null hypothesis	Type I Error (False positive)	Correct outcome! (True positive)
Fail to reject null hypothesis	Correct outcome! (True negative)	Type II Error (False negative)

Representation of Type I and Type II errors



Steps in Hypothesis Testing

- Develop the null and alternative hypotheses.
- Specify the level of significance α (typically, $\alpha = 0.01, 0.05$ or 0.10).
- Collect the sample data and compute the test statistic. This is function of the data and should have a known distribution under H_0.
- Critical value approach:
  - Use the level of significance α to specify the rejection region.
  - H_0 is rejected whenever the test statistic falls within the rejection region.
- *p*-value approach:
  - Use the value of the test statistic to compute the corresponding *p*-value. This is the probability of observing such an extreme test statistic value as the one obtained by the sample data.
    Hence, for a Z-test and computed statistic z^*, the corresponding p–value is given by:

$$p - \text{value} = \begin{cases} P(Z > |z^*| | H_0 \text{ true}), & \text{for 1-sided tests} \\ 2 \cdot P(Z > |z^*| | H_0 \text{ true}), & \text{for 2-sided tests} \end{cases}$$

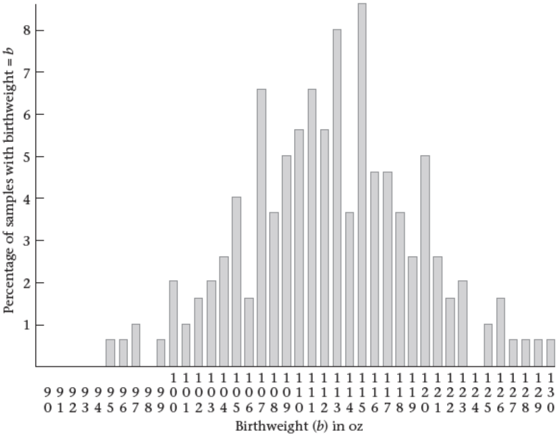
  - H_0 is rejected whenever $p - \text{value} \leq \alpha$.

Sampling distribution of mean (birthweights)

000–019	116	124	119	100	127	103	140	82	107	132	100	92	76	129	138	128	115	133	70	121
020–039	114	114	121	107	120	123	83	96	116	110	71	86	136	118	120	110	107	157	89	71
040–059	98	105	106	52	123	101	111	130	129	94	124	127	128	112	83	95	118	115	86	120
060–079	106	115	100	107	131	114	121	110	115	93	116	76	138	126	143	93	121	135	81	135
080–099	108	152	127	118	110	115	109	133	116	129	118	126	137	110	32	139	132	110	140	119
100–119	109	108	103	88	87	144	105	138	115	104	129	108	92	100	145	93	115	85	124	123
120–139	141	96	146	115	124	113	98	110	153	165	140	132	79	101	127	137	129	144	126	155
140–159	120	128	119	108	113	93	144	124	89	126	87	120	99	60	115	86	143	97	106	148
160–179	113	135	117	129	120	117	92	118	80	132	121	119	57	126	126	77	135	130	102	107
180–199	115	135	112	121	89	135	127	115	133	64	91	126	78	85	106	94	122	111	109	89
200–219	99	118	104	102	94	113	124	118	104	124	133	80	117	112	112	112	102	118	107	104
220–239	90	113	132	122	89	111	118	108	148	103	112	128	86	111	140	126	143	120	124	110
240–259	142	92	132	128	97	132	99	131	120	106	115	101	130	120	130	89	107	152	90	116
260–279	106	111	120	198	123	152	135	83	107	55	131	108	100	104	112	121	102	114	102	101
280–299	118	114	112	133	139	113	77	109	142	144	114	117	97	96	93	120	149	107	107	117
300–319	93	103	121	118	110	89	127	100	156	106	122	105	92	128	124	125	118	113	110	149
320–339	98	98	141	131	92	141	110	134	90	88	111	137	67	95	102	75	108	118	99	79
340–359	110	124	122	104	133	98	108	125	106	128	132	95	114	67	134	136	138	122	103	113
360–379	142	121	125	111	97	127	117	122	120	80	114	126	103	98	108	100	106	98	116	109
380–399	98	97	129	114	102	128	107	119	84	117	119	128	121	113	128	111	112	120	122	91
400–419	117	100	108	101	144	104	110	146	117	107	126	120	104	129	147	111	106	138	97	90
420–439	120	117	94	116	119	108	109	106	134	121	125	105	177	109	109	109	79	118	92	103
440–459	110	95	111	144	130	83	93	81	116	115	131	135	116	97	108	103	134	140	72	112
460–479	101	111	129	128	108	90	113	99	103	41	129	104	144	124	70	106	118	99	85	93
480–499	100	105	104	113	106	88	102	125	132	123	160	100	128	131	49	102	110	106	96	116
500–519	128	102	124	110	129	102	101	119	101	119	141	112	100	105	155	124	67	94	134	123
520–539	92	56	17	135	141	105	133	118	117	112	87	92	104	104	132	121	118	126	114	90
540–559	109	78	117	165	127	122	108	109	119	98	120	101	96	76	143	83	100	128	124	137
560–579	90	129	89	125	131	118	72	121	91	113	91	137	110	137	111	135	105	88	112	104
580–599	102	122	144	114	120	136	144	98	108	130	119	97	142	115	129	125	109	103	114	106
600–619	109	119	89	98	104	115	99	138	122	91	161	96	138	140	32	132	108	92	118	58
620–639	158	127	121	75	112	121	140	80	125	73	115	120	85	104	95	106	100	87	99	113
640–659	95	146	126	58	64	137	69	90	104	124	120	62	83	96	126	155	133	115	97	105
660–679	117	78	105	99	123	86	126	121	109	97	131	133	121	125	120	97	101	92	111	119
680–699	117	80	145	128	140	97	126	109	113	125	157	97	119	103	102	128	116	96	109	112
700–719	67	121	116	126	106	116	77	119	119	122	109	117	127	114	102	75	88	117	99	136
720–739	127	136	103	97	130	129	128	119	22	109	145	129	96	128	122	115	102	127	109	120
740–759	111	114	115	112	146	100	106	137	48	110	97	103	104	107	123	87	140	89	112	123
760–779	130	123	125	124	135	119	78	125	103	55	69	83	106	130	98	81	92	110	112	104
780–799	118	107	117	123	138	130	100	78	146	137	114	61	132	109	133	132	120	116	133	133
800–819	86	116	101	124	126	94	93	132	126	107	98	102	135	59	137	120	119	106	125	122
820–839	101	119	97	86	105	140	89	139	74	131	118	91	98	121	102	115	115	135	100	90
840–859	110	113	136	140	129	117	117	129	143	88	105	110	123	87	97	99	128	128	110	132
860–879	78	128	126	93	148	121	95	121	127	80	109	105	136	141	103	95	140	115	118	117
880–899	114	109	144	119	127	116	103	144	117	131	74	109	117	100	103	123	93	107	113	144
900–919	99	170	97	135	115	89	120	106	141	137	107	132	132	58	113	102	120	98	104	108
920–939	85	115	108	89	88	126	122	107	68	121	113	116	94	85	93	132	146	98	132	104

Sampling distribution of mean (birthweights) - 2

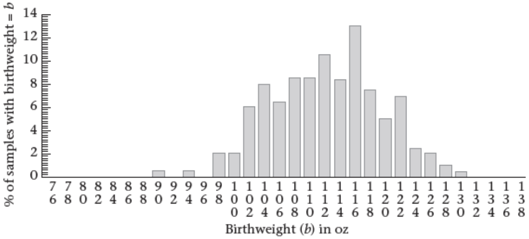
Sampling distribution of $\overline{X}$ over 200 samples of size 10 selected from the population of 1000 birthweights given in Table 6.2 (100 = 100.0–100.9, etc.)



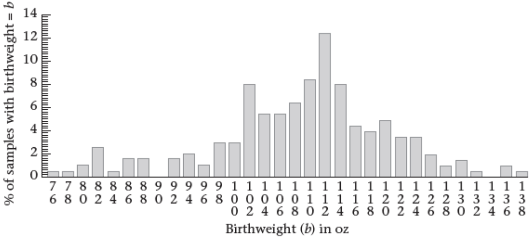
$$\overline{X} \sim \mathcal{N}(\mu_{\overline{X}}, \sigma^{\mathbf{2}}_{\overline{X}}), \text{ where } \mu_{\overline{X}} = \mu, \sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$$

with μ , σ the relevant population parameters and n sample size.

Sampling distributions of median and average of extreme observations have larger variance

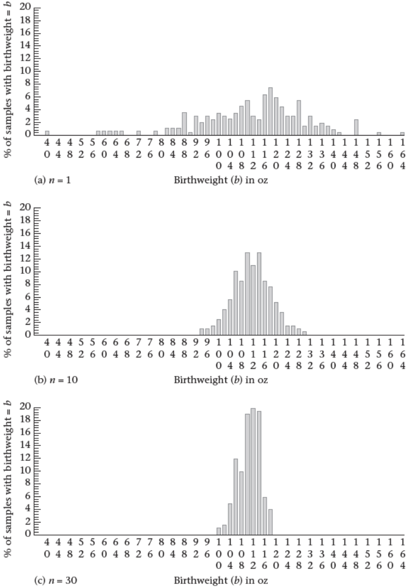


(b) Sampling distribution of the sample median



(c) Sampling distribution of the average of the smallest and largest observations

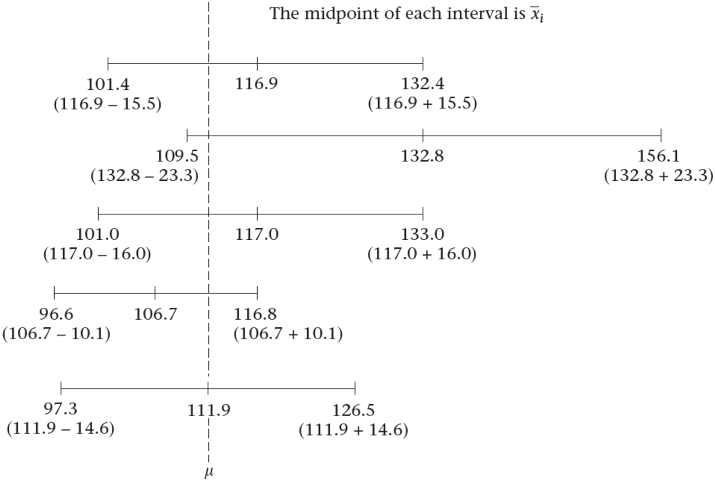
Sampling distributions of mean for increasing sample size



Five random samples of size 10 from the population of infants

	Sample				
Individual	1	2	3	4	5
1	97	177	97	101	137
2	117	198	125	114	118
3	140	107	62	79	78
4	78	99	120	120	129
5	99	104	132	115	87
6	148	121	135	117	110
7	108	148	118	106	106
8	135	133	137	86	116
9	126	126	126	110	140
10	121	115	118	119	98
$\bar{x}$	116.90	132.80	117.00	106.70	111.90
s	21.70	32.62	22.44	14.13	20.46

Collection of 95%-C.I.'s for the mean μ computed from repeated samples



Single sample tests

Single sample mean test: Variance known case

Variance known

Hypotheses	Test statistic	Reject H_0, whenever
$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$	$Z = \frac{\overline{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}}$	$\lvert Z \rvert > z_{\alpha/2}$
$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 \end{cases}$	$Z = \frac{\overline{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}}$	$Z > z_{\alpha}$
$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu < \mu_0 \end{cases}$	$Z = \frac{\overline{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}}$	$Z < -z_{\alpha}$

Critical values for the most usual significance levels:

$z_{0.005}$	2.58
$z_{0.01}$	2.33
$z_{0.025}$	1.96
$z_{0.05}$	1.645
$z_{0.1}$	1.285

Single sample with known σ: Power computation application

Application: For a sample of 16 early and late stage head and neck cancer patients, Ki-67 values (a measure of cell proliferation) are obtained from tissue biopsies. The goal is to compare this sample to a reference population, one with a mean Ki-67, μ_{Ki-67}, of 42% and a known standard deviation, σ_{Ki-67}, of 3%.

1. Compute the power of the left-tailed test with $\alpha = 0.05$ when the true value of the population mean is 40%.
2. What is the conclusion of the left-tailed test for the following sample?

Ki-67 values	38%	40%	42%	45%	43%	47%	45%	38%
	37%	39%	41%	39%	38%	38%	41%	40%

3. If the power of test in (1) is deemed too low, how should the sample size be modified, so as to obtain a power level of 0.90?

Solution:

1. We calculate the power of the test

$$\begin{cases} H_0 : \mu = 42\%; \\ H_1 : \mu < 42\% \end{cases}$$

with with significance $\alpha = 0.05$ when the population mean is in fact $\mu = 0.40\%$:

$$\begin{aligned}
\beta(0.40) &= P(\text{Accept } H_0|\mu = 0.40) = P\left(Z > -z_\alpha|\mu = 0.40\right) \\
&= P\left(\frac{\overline{x} - 0.42}{\sigma/\sqrt{n}} > -z_\alpha|\mu = 0.40\right) = P\left(\overline{x} > 0.42 - z_\alpha\frac{\sigma}{\sqrt{n}}|\mu = 0.40\right) \\
&= P\left(\underbrace{\frac{\overline{x} - 0.40}{\sigma/\sqrt{n}}}_{\sim\mathcal{N}(0,1)} > \frac{0.42 - z_\alpha\frac{\sigma}{\sqrt{n}} - 0.40}{\sigma/\sqrt{n}} = \frac{0.02}{\sigma/\sqrt{n}} - z_\alpha\right) \\
&= 1 - \Phi\left(\frac{0.02}{\sigma/\sqrt{n}} - z_\alpha\right) = 1 - \Phi\left(\frac{0.02}{0.03/\sqrt{16}} - z_{0.05}\right) = 1 - \Phi\left(1.021\right)
\end{aligned}$$

$$\Rightarrow \text{Power}(0.40) = 1 - \beta(0.40) = \Phi\left(1.021\right) = 0.8465$$

Single sample with known σ: Power computation application (2)

Solution:

2 For the test

$$\begin{cases} H_0 : \mu = 42\%; \\ H_1 : \mu < 42\%, \end{cases}$$

we compute

$$Z = \frac{\overline{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{0.4069 - 0.42}{\frac{0.03}{\sqrt{16}}} = -1.75$$

The corresponding rejection region is $(-\infty, -z_\alpha) = (-\infty, -1.645)$ such that $-1.75 \in (-\infty, -1.645)$.

Hence, the null hypothesis is rejected at the significance level of $\alpha = 0.05$.

Single sample with known σ: Power computation application (3)

Solution:

3 We need

$$\text{Power}(0.40) = 0.90 \Rightarrow 1 - \beta(0.40) = 0.90 \Rightarrow \beta(0.40) = 0.10.$$

On the other hand, we have seen

$$\beta(0.40) = 1 - \Phi\left(\frac{0.02}{\sigma/\sqrt{n}} - z_\alpha\right) = 1 - \Phi\left(\frac{0.02}{0.03}\sqrt{n} - z_{0.05}\right).$$

Combining the previous expressions, we have

$$1 - \Phi\left(\frac{0.02}{0.03}\sqrt{n} - 1.645\right) = 0.10 \Rightarrow \Phi\left(\underbrace{\frac{2}{3}\sqrt{n} - 1.645}_{=z_{\mathbf{0.10}}}\right) = 0.90$$

Hence, $\sqrt{n} = (1.645 + z_{0.10})\,\frac{3}{2} = 4.38 \Rightarrow \sqrt{n} = 4.3862 \Rightarrow n = 19.27.$

Note: For power computations (including power curve generation) for the analogue left-/right-/two-sided Z-tests of this example, refer to the uploaded excel file.

Two sided hypotheses tests and confidence intervals

Hypotheses	Test statistic	Reject H_0, whenever
$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$	$Z = \frac{\overline{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$	$\lvert Z \rvert > z_{\alpha/2}$

Hence, H_0 is accepted whenever

$$-z_{\alpha/2} < Z < z_{\alpha/2}$$

$$-z_{\alpha/2} < \frac{\overline{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} < z_{\alpha/2}$$

$$\overline{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu_0 < \overline{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Hence,

$H_0 : \mu = \mu_0$ is accepted at a significance level of α

precisely when the hypothesised value

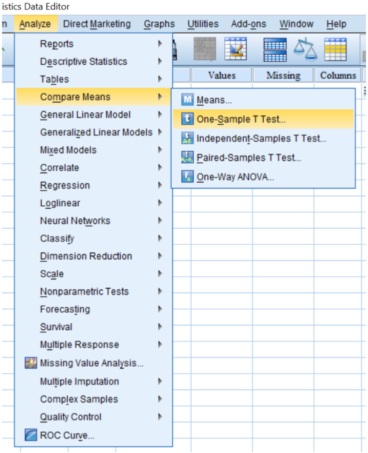
$$\mu_0 \in (1 - \alpha)\% - \text{C.I.}$$

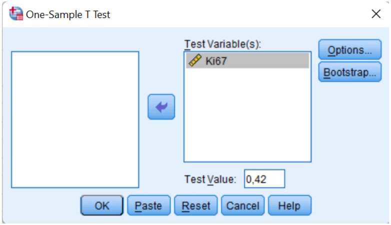
Variance unknown

Hypotheses	Test statistic	Reject H_0, whenever
$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$	$t = \frac{\overline{x} - \mu_0}{\frac{s}{\sqrt{n}}}$	$\lvert t \rvert > t_{\alpha/2;n-1}$
$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 \end{cases}$	$t = \frac{\overline{x} - \mu_0}{\frac{s}{\sqrt{n}}}$	$t > t_{\alpha;n-1}$
$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu < \mu_0 \end{cases}$	$t = \frac{\overline{x} - \mu_0}{\frac{s}{\sqrt{n}}}$	$t < -t_{\alpha;n-1}$

Single sample test in SPSS

Using the data in Slide #10:





One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Ki67	16	,4069	,02983	,00746

One-Sample Test

	Test Value = 0.42					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Ki67	-1,760	15	,099	-,01312	-,0290	,0028

Note that in this instance, for the left–sided test, we have the critical value

$$t_{0.05;15} = -1,75.$$

(Corresponding p–value for one-sided test is $0.099/2 \overset{\sim}{=} 0.05$.)

Independent samples tests

Introduction

- The previous analysis may be extended to two groups.
- Denote μ_1 and μ_2 the respective means of the two populations.
- Note a difference between the two populations exists whenever $\mu_1 - \mu_2 \neq 0$. This difference is the parameter to be estimated.
- We select two independent samples from the two populations: say n_1 cases from Group 1 and n_2 cases from Group 2.

Independent samples

- We select two independent samples from the two populations: say n_1 cases from Group 1 and n_2 cases from Group 2.
  - 1st Sample: $x_1^{(1)}, x_2^{(1)}, \ldots, x_{n_1}^{(1)}$ (from a population of mean μ_1 and standard deviation σ_1)
  - 2nd Sample: $x_1^{(2)}, x_2^{(2)}, \ldots, x_{n_2}^{(2)}$ (from a population of mean μ_2 and standard deviation σ_2)
  - After data collection, compute:

	Sample 1	**Sample 2**
Sample size	n_1	n_2
Sample mean	$\overline{x}_1$	$\overline{x}_2$
Sample standard deviation	s_1	s_2

  - The quantity $\overline{x}_1 - \overline{x}_2$ is an unbiased estimator of $\mu_1 - \mu_2$
  - The variance of this estimator is $\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$

Independent samples with known σ_1, σ_2

Hypotheses	Test statistic	Reject H_0, whenever
$\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 \neq 0 \end{cases}$	$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	$\lvert Z \rvert > z_{\alpha/2}$
$\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 > 0 \end{cases}$	$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	$Z > z_\alpha$
$\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 < 0 \end{cases}$	$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	$Z < -z_\alpha$

Critical values for the most usual α levels:

$z_{0.005}$	2.58
$z_{0.01}$	2.33
$z_{0.025}$	1.96
$z_{0.05}$	1.645
$z_{0.1}$	1.285

Large samples with unknown σ_1, σ_2

- As the standard deviations of the two populations are unknown, they are estimated by the sample standard deviations s_1 και s_2.

Hypotheses	Test statistic	Reject H_0, whenever
$\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 \neq 0 \end{cases}$	$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$	$\lvert Z \rvert > z_{\alpha/2}$
$\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 > 0 \end{cases}$	$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$	$Z > z_\alpha$
$\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 < 0 \end{cases}$	$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$	$Z < -z_\alpha$

Small samples with unknown σ_1, σ_2

- Consider small samples from two populations; i.e., $n_1 < 30$ and/or $n_2 < 30$.
- We assume that both populations are normally distributed.
- Whenever σ_1, σ_2 are unknown, these are estimated by sample standard deviations s_1 και s_2.
- We assume that the variances of the two populations are equal: $\sigma_1^2 = \sigma_2^2 = \sigma^2$.
- Under the previous assumptions, the quantity $\overline{x}_1 - \overline{x_2}$ is normally distributed, independently of sample sizes.
- Combining data from both samples, we obtain the "pooled" estimator of σ^2: $s_p^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$
- Note that for $n_1 = n_2$, we have $s_p^2 = \frac{s_1^2+s_2^2}{2}$

Hypotheses	**Test statistic**	**Reject H_0, whenever**
$\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 \neq 0 \end{cases}$	$t = \frac{\overline{x}_1 - \overline{x}_2}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$	$\lvert t \rvert > t_{\alpha/2;n_1+n_2-2}$
$\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 > 0 \end{cases}$	$t = \frac{\overline{x}_1 - \overline{x}_2}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$	$t > t_{\alpha;n_1+n_2-2}$
$\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 < 0 \end{cases}$	$t = \frac{\overline{x}_1 - \overline{x}_2}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$	$t < -t_{\alpha;n_1+n_2-2}$

C.I. for mean difference in small samples with unknown but equal $\sigma_1 = \sigma_2$

Hypotheses	Test statistic	Reject H_0, whenever
$\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 \neq 0 \end{cases}$	$t = \frac{\overline{x}_1 - \overline{x}_2}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$	$\lvert t \rvert > t_{\alpha/2;n_1+n_2-2}$
$\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 > 0 \end{cases}$	$t = \frac{\overline{x}_1 - \overline{x}_2}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$	$t > t_{\alpha;n_1+n_2-2}$
$\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 < 0 \end{cases}$	$t = \frac{\overline{x}_1 - \overline{x}_2}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$	$t < -t_{\alpha;n_1+n_2-2}$

- $(1-\alpha)$%-C.I. for the mean difference $\mu_1 - \mu_2$:

$$(\overline{x}_1 - \overline{x_2}) \pm s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\,t_{\alpha/2;n_1+n_2-2}$$

- Note that whenever 0 does not lie within the $(1-\alpha)$%-C.I. for the mean difference, then the null hypothesis in the 2–tailed test is rejected at a significance level of α.

C.I. for mean difference in small samples with unknown and different σ_1, σ_2

Hypotheses	Test statistic	Reject H_0, whenever
$\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 \neq 0 \end{cases}$	$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$	$\lvert t \rvert > t_{\alpha/2;df}$
$\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 > 0 \end{cases}$	$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$	$t > t_{\alpha;df}$
$\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 < 0 \end{cases}$	$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$	$t < -t_{\alpha/2;df}$

- Critical values that define the above rejection regions are found under t-distributions with degrees of freedom given by:

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right)^2}{\frac{1}{n_1 - 1} \left(\frac{s_1^2}{n_1} \right)^2 + \frac{1}{n_2 - 1} \left(\frac{s_2^2}{n_2} \right)^2}$$

Mean comparison in independent (small) samples with unknown but equal $\sigma_1 = \sigma_2$: Application

Application:

- A new antipyretic is being tested on two groups; Group 1 (20 cases) is given the new drug, Group 2 (18 cases) is given an older medication.
- Temperature data on the 38 patients are given in the next slide.
- Test whether the mean temperature in the two groups coincide or not ($\alpha = 0.05$).

Temperature data

Ομάδα A	Ομάδα B
38,40	40,90
36,80	39,50
40,00	39,40
39,80	38,20
38,60	39,70
39,10	38,90
38,90	38,60
36,80	39,90
40,40	41,30
39,40	38,10
38,00	39,60
38,60	37,10
40,10	39,50
38,10	40,30
37,20	41,50
39,50	39,30
37,30	37,60
39,10	40,60
39,90	
37,80	

Mean comparison in independent (small) samples with unknown but equal $\sigma_1 = \sigma_2$: Application solution

- 2-tailed hyptheses: $\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 \neq 0, \end{cases}$ where μ_1, μ_2 mean temperatures in the two populations.
- Test statistic:

$$t = \frac{\overline{x}_1 - \overline{x}_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}},$$

where

$$s_p{}^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}.$$

- Hence, we have to compute the quantities

$$\overline{x}_1 = \frac{\sum_{j=1}^{n_1} x_j^{(1)}}{n_1}, \quad s_1^2 = \frac{1}{n_1 - 1}\left(\sum_{j=1}^{n_1}\left(x_j^{(1)}\right)^2 - \frac{\left(\sum_{j=1}^{n_1}\left(x_j^{(1)}\right)\right)^2}{n_1}\right)$$

and

$$\overline{x}_2 = \frac{\sum_{j=1}^{n_2} x_j^{(2)}}{n_2}, \quad s_2^2 = \frac{1}{n_2 - 1}\left(\sum_{j=1}^{n_2}\left(x_j^{(2)}\right)^2 - \frac{\left(\sum_{j=1}^{n_2}\left(x_j^{(2)}\right)\right)^2}{n_2}\right).$$

Computations for application

	x_j(1)	x_j^(1)*x_j^(1)		x_j(2)	x_j^(2)*x_j^(2)
	38,40	1.474,56		40,90	1.672,81
	36,80	1.354,24		39,50	1.560,25
	40,00	1.600,00		39,40	1.552,36
	39,80	1.584,04		38,20	1.459,24
	38,60	1.489,96		39,70	1.576,09
	39,10	1.528,81		38,90	1.513,21
	38,90	1.513,21		38,60	1.489,96
	36,80	1.354,24		39,90	1.592,01
	40,40	1.632,16		41,30	1.705,69
	39,40	1.552,36		38,10	1.451,61
	38,00	1.444,00		39,60	1.568,16
	38,60	1.489,96		37,10	1.376,41
	40,10	1.608,01		39,50	1.560,25
	38,10	1.451,61		40,30	1.624,09
	37,20	1.383,84		41,50	1.722,25
	39,50	1.560,25		39,30	1.544,49
	37,30	1.391,29		37,60	1.413,76
	39,10	1.528,81		40,60	1.648,36
	39,90	1.592,01	ΑΘΡΟΙΣΜΑ	710,00	28.031,00
	37,80	1.428,84			
ΑΘΡΟΙΣΜΑ	773,80	29.962,20			

Mean comparison in independent (small) samples with unknown but equal $\sigma_1 = \sigma_2$: Application solution

- For Group 1, we have

$$\overline{x}_1 = \frac{\sum_{j=1}^{n_1} x_j^{(1)}}{n_1} = \frac{773.80}{20} = 38.69$$

and

$$s_1^2 = \frac{1}{n_1 - 1}\left(\sum_{j=1}^{n_1}\left(x_j^{(1)}\right)^2 - \frac{\left(\sum_{j=1}^{n_1}\left(x_j^{(1)}\right)\right)^2}{n_1}\right) = \frac{1}{19}\left(29962.2 - \frac{773.8^2}{20}\right) = 1.257$$

- For Group 2, we have

$$\overline{x}_2 = \frac{\sum_{j=1}^{n_2} x_j^{(2)}}{n_2} = \frac{710.0}{18} = 39.44$$

and

$$s_2^2 = \frac{1}{n_2 - 1}\left(\sum_{j=1}^{n_2}\left(x_j^{(2)}\right)^2 - \frac{\left(\sum_{j=1}^{n_2}\left(x_j^{(2)}\right)\right)^2}{n_2}\right) = \frac{1}{17}\left(28031 - \frac{710^2}{18}\right) = 1.497$$

Mean comparison in independent (small) samples with unknown but equal $\sigma_1 = \sigma_2$: Application solution

- 2-tailed hypotheses: $\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 \neq 0, \end{cases}$ where μ_1, μ_2 the mean temperatures in the two populations.
- Test statistic:

$$t = \frac{\overline{x}_1 - \overline{x}_2}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{38.69 - 39.44}{\sqrt{1.37}\sqrt{\frac{1}{20} + \frac{1}{18}}} = -2.00,$$

where

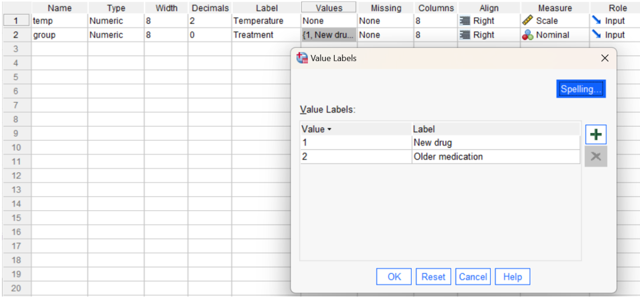
$${s_p}^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{(20 - 1)1.257 + (18 - 1)1.497}{20 + 18 - 2} = 1.37.$$

- Rejection region: $|t| > t_{\alpha/2;n_1+n_2-2} = t_{0.025;36} = 2.028$
- Evidently,

$$|t| = 2.00 < t_{0.025;36} = 2.028.$$

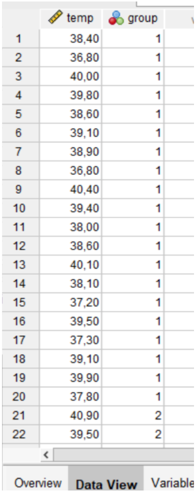
Thus, the test statistic does not lie in the rejection region and we cannot reject H_0 at a level of significance of 0.05.

Independent samples test in SPSS: Variables & Data



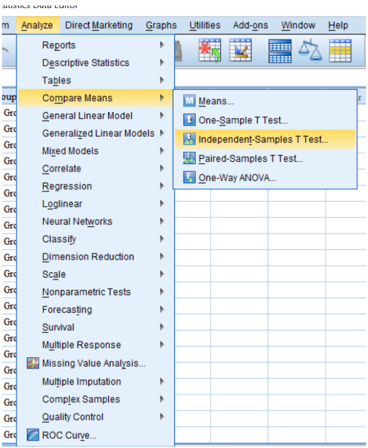
	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	temp	Numeric	8	2	Temperature	None	None	8	Right	Scale	Input
2	group	Numeric	8	0	Treatment	{1, New dru...	None	8	Right	Nominal	Input

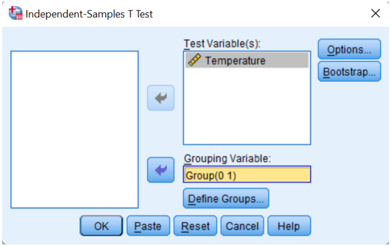
Value	Label
1	New drug
2	Older medication

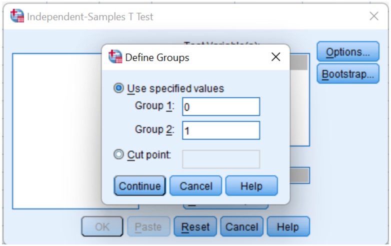


	temp	group
1	38,40	1
2	36,80	1
3	40,00	1
4	39,80	1
5	38,60	1
6	39,10	1
7	38,90	1
8	36,80	1
9	40,40	1
10	39,40	1
11	38,00	1
12	38,60	1
13	40,10	1
14	38,10	1
15	37,20	1
16	39,50	1
17	37,30	1
18	39,10	1
19	39,90	1
20	37,80	1
21	40,90	2
22	39,50	2

Independent samples test in SPSS







Independent samples test in SPSS: Output

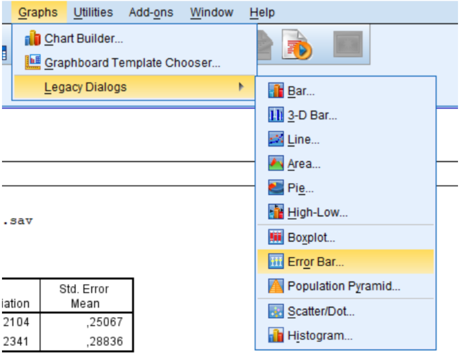
Group Statistics

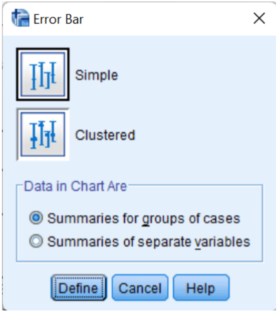
	Group	N	Mean	Std. Deviation	Std. Error Mean
Temperature	Group 1	20	38,6900	1,12104	,25067
	Group 2	18	39,4444	1,22341	,28836

Independent Samples Test

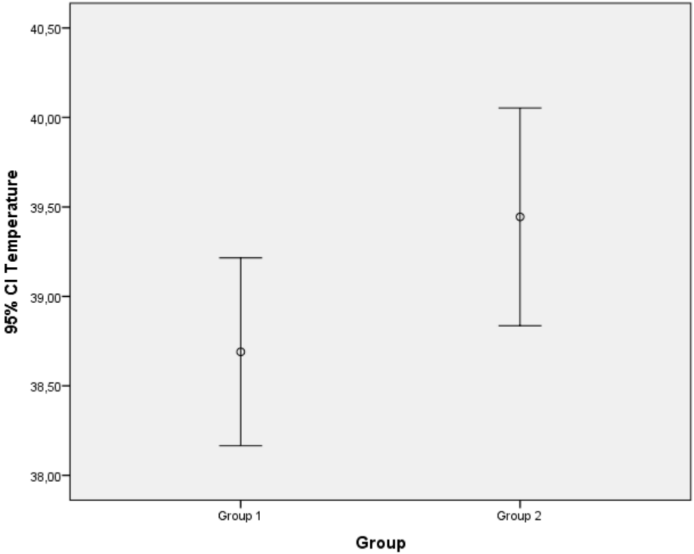
		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Temperature	Equal variances assumed	,000	,994	-1,984	36	,055	-,75444	,38029	-1,52570	,01681
	Equal variances not assumed			-1,975	34,681	,056	-,75444	,38208	-1,53037	,02148

C.I. Visualization in SPSS

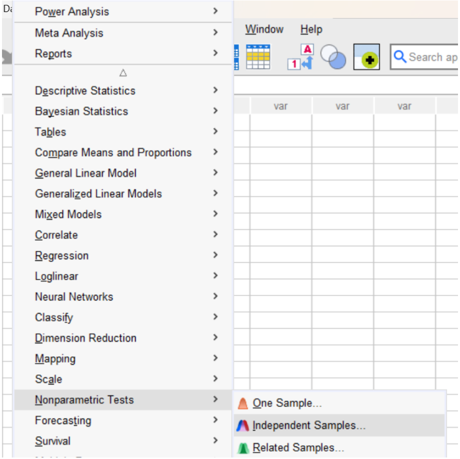




C.I. Visualization in SPSS



Nonparametric alternative in SPSS: Mann-Whitney U



- An essential assumption for the independent samples t-test is that the values of the scale variable should be distributed according to the normal distribution in each of the two independent samples. (Partition of the data in the respective sets is easily accomplished, using the Explore procedure of SPSS.)
- To assess normality, consider the various visual and statistical procedures, namely:
  - Superposition of the normal curve on histograms in the two samples.
  - PP/QQ plots.
  - The distribution should be mesokyrtic and not skewed (i.e., corresponding skewness/kyrtosis statistics, normalized by their respective standard errors, should not exceed 2 in absolute value).
  - Kolmogorov–Smirnov test, Shapiro–Wilk test.
- Should these fail, a non–parametric alternative provided in SPSS is Mann–Whitney's U.

Nonparametric Tests

[DataSet1]

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.[a,b]	Decision
1	The distribution of Temperature is the same across categories of Treatment.	Independent-Samples Mann-Whitney U Test	,082[c]	Retain the null hypothesis.

a. The significance level is ,050.

b. Asymptotic significance is displayed.

c. Exact significance is displayed for this test.

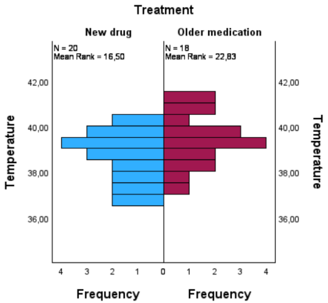
Independent-Samples Mann-Whitney U Test

Temperature across Treatment

Independent-Samples Mann-Whitney U Test Summary

Total N	38
Mann-Whitney U	240,000
Wilcoxon W	411,000
Test Statistic	240,000
Standard Error	34,179
Standardized Test Statistic	1,755
Asymptotic Sig.(2-sided test)	,079
Exact Sig.(2-sided test)	,082

Independent-Samples Mann-Whitney U Test



Paired means comparison

Paired means comparison

- With paired samples, the observations in the two groups are matched in a meaningful way. These are also known as dependent samples.
- Most often this occurs when data are collected twice from the same participants, called repeated measures ("prior" and "post").
- Paired data does not always need to involve two measurements on the same subject; it can also involve taking one measurement on each of two related subjects. For example, we may study husband-wife pairs, mother-son pairs, or pairs of twins.
- Using data from a sample with n cases for the two scale variables X, Y, we let $Z = X - Y$ and compute $\overline{z} = \overline{x} - \overline{y}$ and s_z.

Hypotheses	Test statistic	Reject H_0, whenever
$\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 \neq 0 \end{cases}$	$t = \frac{\overline{z}}{\frac{s_z}{\sqrt{n}}}$	$\lvert t \rvert > t_{\alpha/2;n-1}$
$\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 > 0 \end{cases}$	$t = \frac{\overline{z}}{\frac{s_z}{\sqrt{n}}}$	$t > t_{\alpha;n-1}$
$\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 < 0 \end{cases}$	$t = \frac{\overline{z}}{\frac{s_z}{\sqrt{n}}}$	$t < -t_{\alpha;n-1}$

Here μ_1, μ_2 denote population means of X, Y, respectively.

Paired means comparison: Application

Application.

- Blood pressure in 10 patients before (X) and after (Y) medicine administration is given in the following Table:

X	13	15	18	14	12	13	15	16	18	19
Y	12	13	15	15	14	13	13	14	14	13

- Check at a level of significance of 5% whether this medication is effective in decreasing pressure in patients.

Computations for Example

- Consider the one–tailed test:

$$\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 > 0, \end{cases}$$

where μ_1, μ_2 denote population means of blood pressure before and following medication administration, respectively.

- Define the differences $z_i = x_i - y_i$ $(i = 1, 2, \ldots, 10)$.
- Need to compute

$$\bar{z} = \frac{\sum_{j=1}^{10} z_j}{10} = \frac{17}{10} = 1.7$$

and

$$s_z^2 = \frac{1}{n-1}\left(\sum_{j=1}^{n}(z_j)^2 - \frac{\left(\sum_{j=1}^{n}(z_j)\right)^2}{n}\right) = \frac{1}{9}\left(79 - \frac{17^2}{10}\right) = 5.57 \Rightarrow$$

$$s_z = \sqrt{5.57} = 2.36$$

Computations for Example (2)

X	Y	Z	Z^2
13	12	1	1
15	13	2	4
18	15	3	9
14	15	-1	1
12	14	-2	4
13	13	0	0
15	13	2	4
16	14	2	4
18	14	4	16
19	13	6	36
	ΑΘΡΟΙΣΜΑ	17	79

$\overline{z} = 1.7$ and $s_z = \sqrt{5.57} = 2.36$

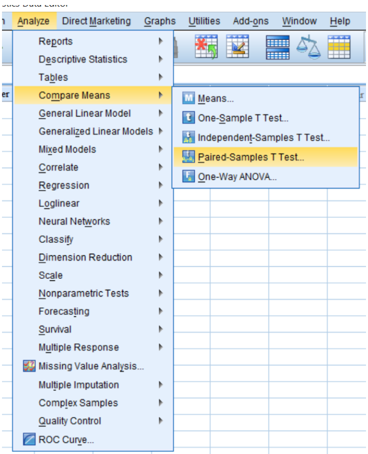
- Test statistic $t = \frac{\overline{z}}{s_z/\sqrt{n}} = \frac{1.7}{2.36/\sqrt{10}} = 2.278$
- Rejection region: $(t_{\alpha;n-1}, \infty) = (t_{0.05;9}, \infty) = (1.833, \infty)$
- Clearly, test statistic lies in rejection region, so H_0 is rejected.

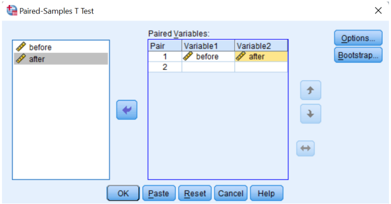
Dependent samples test in SPSS: Variables & Data

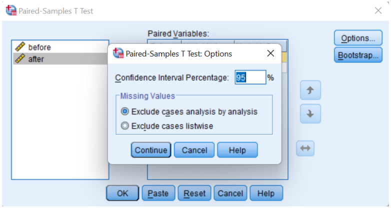
	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	pressure1	Numeric	8	2	Pressure (before)	None	None	8	Right	Scale	Input
2	pressure2	Numeric	8	2	Pressure (After)	None	None	8	Right	Scale	Input
3											
4											

	pressure1	pressure2	var
1	13	12	
2	15	13	
3	18	15	
4	14	15	
5	12	14	
6	13	13	
7	15	13	
8	16	14	
9	18	14	
10	19	13	
11			

Dependent samples test in SPSS







Dependent samples test in SPSS: Ouput

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Before medicine administration	15,30	10	2,406	,761
	After medicine administration	13,60	10	,966	,306

Paired Samples Correlations

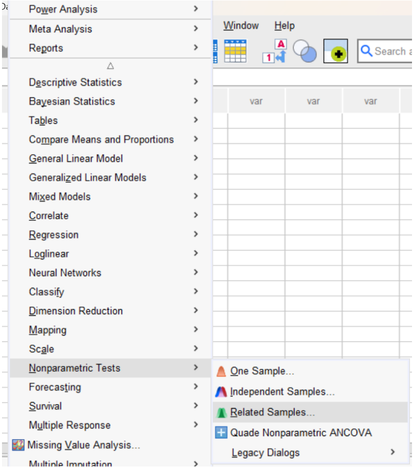
		N	Correlation	Sig.
Pair 1	Before medicine administration & After medicine administration	10	,249	,489

Paired Samples Test

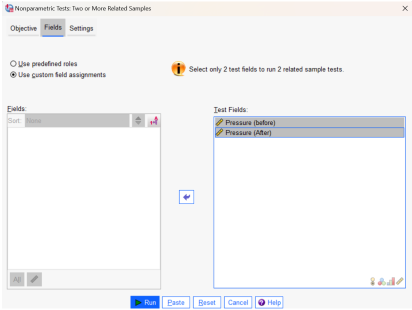
		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference: Lower	95% Confidence Interval of the Difference: Upper			
Pair 1	Before medicine administration - After medicine administration	1,700	2,359	,746	,012	3,388	2,279	9	,049

The *p*-value reported here is for the 2–tailed test. For our purposes, we need to compare $\alpha = 0.05$ with the correct *p*-value for the one-tailed test, which is 0.049/2.

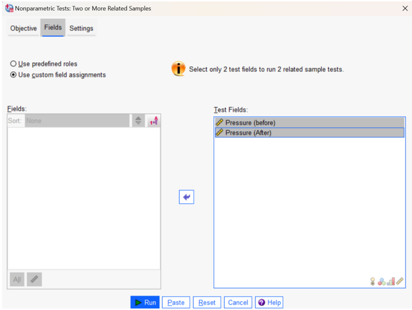
Nonparametric alternative in SPSS: Wilkoxon's signed rank test

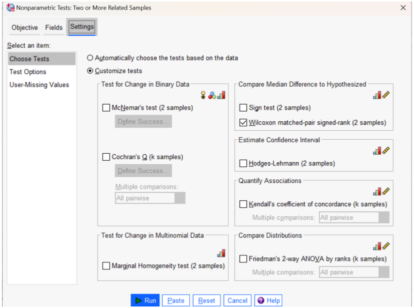


- An essential assumption for the dependent samples t-test is that the values of both scale variables are distributed according to the normal distribution in (i.e., in both dependent samples).
- Should normality tests fail in either of these variables, a non–parametric alternative provided in SPSS is Wilkoxon's signed rank test.



Wilkoxon signed rank test





Nonparametric Tests

Hypothesis Test Summary

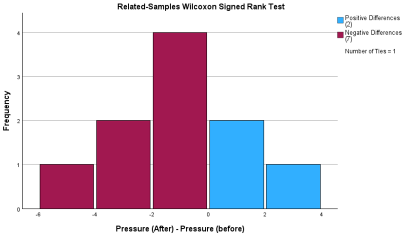
	Null Hypothesis	Test	Sig.[a,b]	Decision
1	The median of differences between Pressure (before) and Pressure (After) equals 0.	Related-Samples Wilcoxon Signed Rank Test	,048	Reject the null hypothesis.

a. The significance level is ,050.

b. Asymptotic significance is displayed.

Related-Samples Wilcoxon Signed Rank Test Summary

Total N	10
Test Statistic	6,000
Standard Error	8,359
Standardized Test Statistic	-1,974
Asymptotic Sig.(2-sided test)	,048



Proportion tests

One sample proportion test

Hypotheses	Test statistic	Reject H_0, whenever
$\begin{cases} H_0 : p = p_0 \\ H_1 : p \neq p_0 \end{cases}$	$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$	$\lvert z \rvert > z_{\alpha/2}$
$\begin{cases} H_0 : p = p_0 \\ H_1 : p > p_0 \end{cases}$	$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$	$z > z_\alpha$
$\begin{cases} H_0 : p = p_0 \\ H_1 : p < p_0 \end{cases}$	$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$	$z < -z_\alpha$

Independent samples proportions comparison test

Hypotheses	Test statistic	Reject H_0, whenever
$\begin{cases} H_0 : p_1 - p_2 = 0 \\ H_1 : p_1 - p_2 \neq 0 \end{cases}$	$z = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$	$\lvert z \rvert > z_{\alpha/2}$
$\begin{cases} H_0 : p_1 - p_2 = 0 \\ H_1 : p_1 - p_2 > 0 \end{cases}$	$z = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$	$z > z_\alpha$
$\begin{cases} H_0 : p_1 - p_2 = 0 \\ H_1 : p_1 - p_2 < 0 \end{cases}$	$z = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$	$z < -z_\alpha$

Variance tests

One sample variance test

Hypotheses	Test statistic	Reject H_0, whenever
$\begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_1 : \sigma^2 \neq \sigma_0^2 \end{cases}$	$X^2 = \frac{(n-1)s^2}{\sigma_0^2}$	$X^2 > \chi^2_{\alpha/2;n-1}$ or $X^2 < \chi^2_{1-\alpha/2;n-1}$
$\begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_1 : \sigma^2 > \sigma_0^2 \end{cases}$	$X^2 = \frac{(n-1)s^2}{\sigma_0^2}$	$X^2 > \chi^2_{\alpha/2;n-1}$
$\begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_1 : \sigma^2 < \sigma_0^2 \end{cases}$	$X^2 = \frac{(n-1)s^2}{\sigma_0^2}$	$X^2 < \chi^2_{1-\alpha/2;n-1}$

Independent samples variance test

Hypotheses	Test statistic	Reject H_0, whenever
$\begin{cases} H_0 : \frac{\sigma_2^2}{\sigma_1^2} = 1 \\ H_1 : \frac{\sigma_2^2}{\sigma_1^2} \neq 1 \end{cases}$	$F = \frac{s_1^2}{s_2^2}$	$F > F_{\alpha/2;(n_1-1,n_2-1)}$ or $F < F_{1-\alpha/2;(n_1-1,n_2-1)}$
$\begin{cases} H_0 : \frac{\sigma_2^2}{\sigma_1^2} = 1 \\ H_1 : \frac{\sigma_2^2}{\sigma_1^2} > 1 \end{cases}$	$F = \frac{s_1^2}{s_2^2}$	$F < F_{1-\alpha;(n_1-1,n_2-1)}$
$\begin{cases} H_0 : \frac{\sigma_2^2}{\sigma_1^2} = 1 \\ H_1 : \frac{\sigma_2^2}{\sigma_1^2} < 1 \end{cases}$	$F = \frac{s_1^2}{s_2^2}$	$F > F_{\alpha;(n_1-1,n_2-1)}$

Analysis of Variance

Analysis of Variance

- If several competing treatments are being used in the sampling process, the problem involves one factor with more than two levels.
- In the $k > 2$ sample problem, it will be assumed that there are k samples populations.
- One very common procedure used to deal with testing population means is called the analysis of variance, or ANOVA.
- Test the hypotheses:

$$\begin{cases} H_0 : \mu_1 = \mu_2 = \cdots = \mu_k, \\ H_1 : \text{ at least two of the means are not equal} \end{cases}$$

Assumptions

The ANOVA test makes the following assumptions about the data:

- Independence of the observations. Each subject should belong to only one group. There is no relationship between the observations in each group. Having repeated measures for the same participants is not allowed.
- No significant outliers in any cell of the design
- Normality. The data for each design cell should be approximately normally distributed.
- Homogeneity of variances. The variance of the outcome variable should be equal in every cell of the design.

Notation

- Samples $j = 1, \ldots, k$
- Observations in sample-j ($j = 1, \ldots, k$):

$$x_{1j}, x_{2j}, \ldots, x_{n_j, j},$$

  where n_j denotes the j–sample size.
- $\overline{x} = \frac{\sum_{j=1}^{k} \sum_{i=1}^{n_j} x_{ij}}{\sum_{j=1}^{k} n_j}$ is the grand mean.
- $\overline{x}_j = \frac{\sum_{i=1}^{n_j} x_{ij}}{n_j}$ is the j–sample mean.

Sources of variability within data

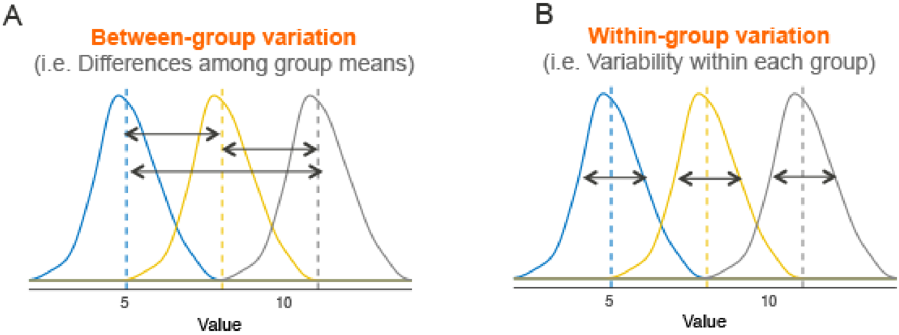
- Analysis-of-Variance approach: A procedure whereby the total variation is subdivided into components. Part of the goal of the analysis of variance is to determine if differences among the k sample means are what we would expect due to random variation alone or, rather, due to variation beyond merely random effects.
- It can be shown that

$$SST = SSB + SSW$$

$$\underbrace{\sum_{j=1}^{k}\sum_{i=1}^{n_j}(x_{ij}-\overline{x})^2}_{\text{Total Sum of Squares}} = \underbrace{\sum_{j=1}^{k}\sum_{i=1}^{n_j}(x_{ij}-\overline{x}_j)^2}_{\text{Sum of Squares Within Groups}} + \underbrace{\sum_{j=1}^{k} n_j(\overline{x}_j-\overline{x})^2}_{\text{Sum of Squares Between Groups}}$$

- The quantities $(x_{ij}-\overline{x}_j)$ forming SSW encode variability *within each group*.
- The quantities $n_j(\overline{x}_j-\overline{x})$ forming SSW encode variability *between groups*.

Between and within group variation



- Basic idea: if the average variation between groups is large enough compared to the average variation within groups, then you could conclude that at least one group mean is not equal to the others. Hence, $H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$ would be rejected.
- Thus, it's possible to evaluate whether the differences between the group means are significant by comparing the two variance estimates.

One-way AnoVa Table

- Test:

$$\begin{cases} H_0 : \mu_1 = \mu_2 = \cdots = \mu_k, \\ H_1 : \text{ at least two of the } \mu_i\text{'s are not equal} \end{cases}$$

- Test statistic:

$$F = \frac{SSB/(k-1)}{SSW/(n-k)} = \frac{MSB}{MSW} \sim F_{(k-1,n-k)}$$

- Rejection region $R = (F_{\alpha;(k-1,n-k)}, \infty)$ for a level of significance α.
- A high ratio implies that the variation among group means are greatly different from each other compared to the variation of the individual observations in each group.

One-way ANOVA Table

Source of Variation	**Sum of Squares**	**Degrees of Freedom**	**Mean Square**	**Computed f**
Between Groups	SSB	$k-1$	$MSB = \frac{SSB}{k-1}$	$f = \frac{MSB}{MSW}$
Within Groups	SSW	$n-k$	$MSW = \frac{SSW}{n-k}$	
Total	SST	$n-1$		

Data for AnoVa application

Application: Differences in steady-state haemoglobin levels ($g/declitre$) between patients with different types of sickle cell disease.

- Do patients in these groups have identical mean levels of haemoglobin?

Type	n_j	Mean ($\overline{x}_j$)	s.d. (s_j)
1: Hb SS	16	8.7125	0.8445
2: Hb S/β	10	10.6300	1.2841
3: Hb SC	15	12.3000	0.9419

Group 1	Group 2	Group 3
7,2	8,1	10,7
7,7	9,2	11,3
8	10	11,5
8,1	10,4	11,6
8,3	10,6	11,7
8,4	10,9	11,8
8,4	11,1	12
8,5	11,9	12,1
8,6	12	12,3
8,7	12,1	12,6
9,1		12,6
9,1		13,3
9,1		13,3
9,8		13,8
10,1		13,9
10,3		

AnoVa computations

	Group 1	Gr. 1 SQUARES		Group 2	Gr. 2 SQUARES		Group 3	Gr. 3 SQUARES
	7,2	51,84		8,1	65,61		10,7	114,49
	7,7	59,29		9,2	84,64		11,3	127,69
	8	64		10	100		11,5	132,25
	8,1	65,61		10,4	108,16		11,6	134,56
	8,3	68,89		10,6	112,36		11,7	136,89
	8,4	70,56		10,9	118,81		11,8	139,24
	8,4	70,56		11,1	123,21		12	144
	8,5	72,25		11,9	141,61		12,1	146,41
	8,6	73,96		12	144		12,3	151,29
	8,7	75,69		12,1	146,41		12,6	158,76
	9,1	82,81	SUM	106,3	1144,81		12,6	158,76
	9,1	82,81	bar{x}_2	10,63			13,3	176,89
	9,1	82,81					13,3	176,89
	9,8	96,04					13,8	190,44
	10,1	102,01					13,9	193,21
	10,3	106,09				SUM	184,5	2281,77
SUM	139,4	1225,22				bar{x}_3	12,3	
bar{x}_1	8,7125							

AnoVa computations (SST)

$$n = \sum_{j=1}^{3} n_j = 16 + 10 + 15 = 41,$$

$$\sum_{j=1}^{3}\sum_{i=1}^{n_j} x_{ij} = 139.4 + 106.3 + 184.5 = 430.2 \Rightarrow \overline{x} = \frac{\sum_{j=1}^{3}\sum_{i=1}^{n_j} x_{ij}}{n} = \frac{430.2}{41} = 10.49$$

$$\sum_{j=1}^{3}\sum_{i=1}^{n_j} x_{ij}^2 = 1225.22 + 1144.81 + 22181.77 = 4651.8$$

$$\textbf{Total: } SST = \sum_{j=1}^{3}\sum_{i=1}^{n_j} (x_{ij} - \overline{x})^2 = \sum_{j=1}^{3}\sum_{i=1}^{n_j} x_{ij}^2 - \frac{\left(\sum_{j=1}^{3}\sum_{i=1}^{n_j} x_{ij}\right)^2}{n} =$$

$$= 4651.8 - \frac{430.2^2}{41} = 137.85$$

$$\text{d.f.} = n - 1 = 40$$

AnoVa computations (SSB)

Type	n_j	Mean ($\overline{x}_j$)	s.d. (s_j)
1: Hb SS	16	8.7125	0.8445
2: Hb S/β	10	10.6300	1.2841
3: Hb SC	15	12.3000	0.9419

$$n = \sum_{j=1}^{3} n_j = 16 + 10 + 15 = 41,$$

$$\sum_{j=1}^{3} \sum_{i=1}^{n_j} x_{ij} = 430.2 \Rightarrow \overline{x} = \frac{430.2}{41} = 10.49$$

$$\sum_{j=1}^{3} \sum_{i=1}^{n_j} x_{ij}^2 = 4651.8$$

Between: $$SSB = \sum_{j=1}^{3} n_j \left(\overline{x}_j - \overline{x}\right)^2 = 16(8.71 - 10.49)^2 + 10(10.63 - 10.49)^2 +$$

$$+ 15(12.3 - 10.49)^2 = 99.89$$

d.f. $= k - 1 = 2$

Easier calculation: $$SSB = \sum_{j=1}^{3} n_j \overline{x}_j^2 - \frac{\left(\sum_{j=1}^{3} \sum_{i=1}^{n_j} x_{ij}\right)^2}{n} = 16 \cdot 8.7125^2 + 10 \cdot 10.63^2 + 15 \cdot 12.3^2 - \frac{430.2^2}{41}$$

$$= 99.89$$

AnoVa computations (SSW)

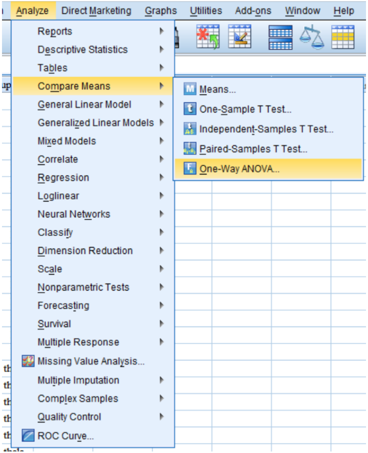
Within: $SSW = \sum_{j=1}^{3}\sum_{i=1}^{n_j}(x_{ij} - \overline{x}_j)^2 = \sum_{j=1}^{3}(n_j - 1)s_j^2$

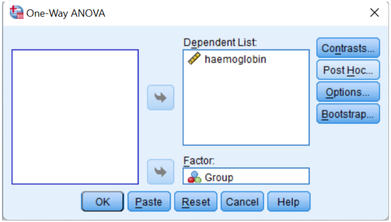
$= 15 \cdot 0.8445^2 + 9 \cdot 1.2841^2 + 14 \cdot 0.9419^2 = 37.96$

d.f. $= n - k = 41 - 3 = 38$

Type	n_j	**Mean** $(\overline{x}_j)$	**s.d.** (s_j)
1: Hb SS	16	8.7125	0.8445
2: Hb S/β	10	10.6300	1.2841
3: Hb SC	15	12.3000	0.9419

One-way AnoVa in SPSS



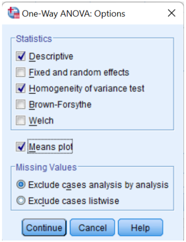


ANOVA

Haemoglobin (g/declitre)

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	99,889	2	49,945	49,999	,000
Within Groups	37,959	38	,999		
Total	137,848	40			

AnoVa Group Descriptives in SPSS



Descriptives

Haemoglobin (g/declitre)

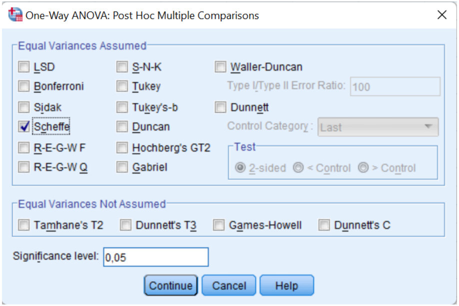
	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Hb S	16	8,712	,8445	,2111	8,263	9,162	7,2	10,3
Hb S/β - thalassaemia	10	10,630	1,2841	,4061	9,711	11,549	8,1	12,1
Hb SC	15	12,300	,9419	,2432	11,778	12,822	10,7	13,9
Total	41	10,493	1,8564	,2899	9,907	11,079	7,2	13,9

Test of Homogeneity of Variances

Haemoglobin (g/declitre)

Levene Statistic	df1	df2	Sig.
,902	2	38	,414

Post-Hoc in SPSS



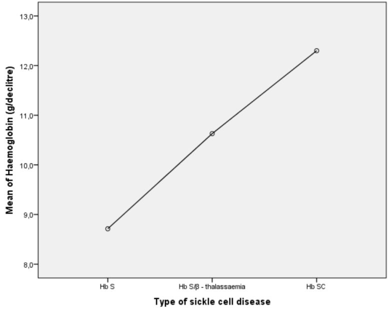
Multiple Comparisons

Dependent Variable: Haemoglobin (g/declitre)
Scheffe

(I) Type of sickle cell disease	(J) Type of sickle cell disease	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Hb S	Hb S/β - thalassaemia	-1,9175*	,4029	,000	-2,944	-,891
	Hb SC	-3,5875*	,3592	,000	-4,503	-2,672
Hb S/β - thalassaemia	Hb S	1,9175*	,4029	,000	,891	2,944
	Hb SC	-1,6700*	,4080	,001	-2,709	-,631
Hb SC	Hb S	3,5875*	,3592	,000	2,672	4,503
	Hb S/β - thalassaemia	1,6700*	,4080	,001	,631	2,709

*. The mean difference is significant at the 0.05 level.

Homogeneous subsets



Haemoglobin (g/decltre)

Scheffe[a,b]

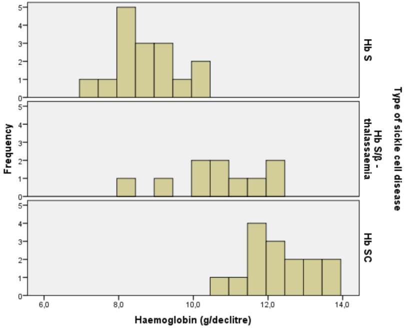
Type of sickle cell disease	N	Subset for alpha = 0.05		
		1	2	3
Hb S	16	8,712		
Hb S/β - thalassaemia	10		10,630	
Hb SC	15			12,300
Sig.		1,000	1,000	1,000

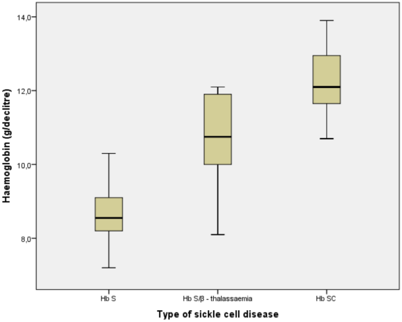
Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 13,091.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

ANOVA Visualization

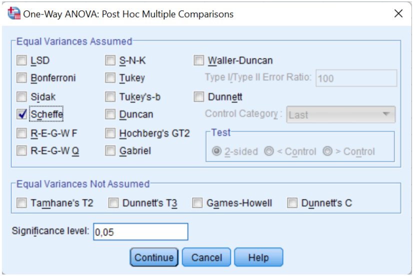




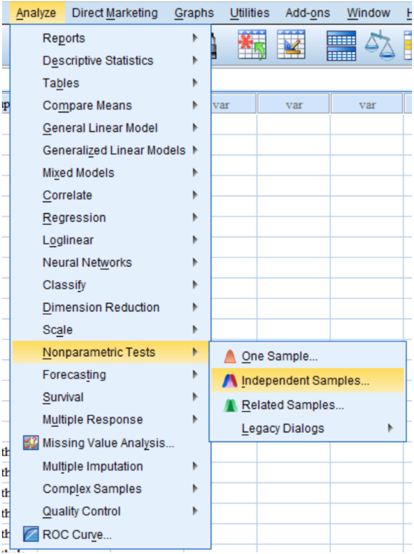
Relaxing the ANOVA assumptions

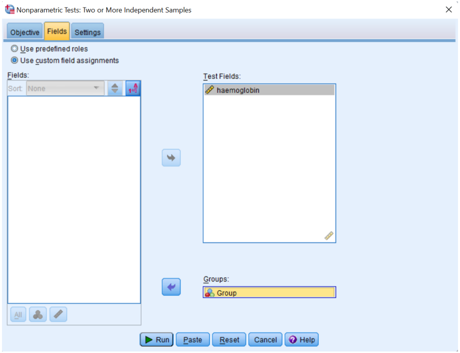
In case the homogeneity of variance or normality assumptions are violated:

- The Independent Samples Kruskal-Wallis or the Welch one-way test is an alternative to the standard one-way ANOVA in the situation where the homogeneity of variance can't be assumed (i.e., Levene test is significant).
- In this case, the Games-Howell post hoc test or pairwise t-tests (with no assumption of equal variances) can be used to compare all possible combinations of group differences.

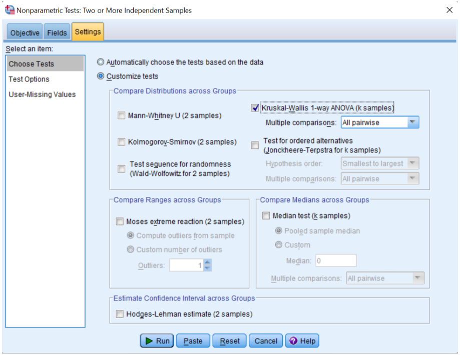


Non–parametric alternatives in SPSS





Non–parametric alternatives in SPSS



Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Haemoglobin (g/declitre) is the same across categories of Type of sickle cell disease.	Independent-Samples Kruskal-Wallis Test	,000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.