

# Chapter 7

## Unpaired Continuous Data (Unpaired T-Test, Mann-Whitney)

### 1 General Purpose

For the study of two outcomes often two parallel groups of similar age, gender and other characteristics are applied, and the studies are called parallel-group studies, and the two groups are called independent of one another. This study gives examples of parallel-group analyses.

### 2 Schematic Overview of Type of Data File

Outcome	Parallel-group (1,2)
.	1
.	1
.	1
.	1
.	1
.	1
.	1
.	1
.	1
.	1
.	2
.	2
.	2
.	2
.	2
.	2
.	2
.	2

### 3 Primary Scientific Question

Are unpaired t-tests or Mann-Whitney tests appropriate for testing, whether the outcome of the first group is significantly different from that of the second group?

### 4 Data Example

The underneath study assesses in 20 patients whether some sleeping pill (parallel-group 1) is more efficacious than a placebo (parallel-group 2). The hours of sleep is the outcome value.

Outcome parallel-group 1	Outcome parallel-group 2
6,0	5,1
7,1	8,0
8,1	3,8
7,5	4,7
6,4	5,2
7,9	5,4
6,8	4,3
6,6	6,0
7,3	3,7
5,6	6,2

Outcome = hours of sleep after treatment

### 5 Unpaired T-Test

Two age- and gender- matched parallel-groups are compared with one another. For the calculation of SDs see the Chap. 1.

group 1:

6,0, 7,1, 8,1, 7,5, 6,4, 7,9, 6,8, 6,6, 7,3, 5,6

group 2:

5,1, 8,0, 3,8, 4,4, 5,2, 5,4, 4,3, 6,0, 3,7, 6,2

Mean group 1 = 6.93 SD = 0.806 SE =  $SD/\sqrt{10} = 0.255$

Mean Group 2 = 5.21 SD = 1.299 SE =  $SD/\sqrt{10} = 0.411$

- A. not significant
- B.  $0.05 < p < 0.10$
- C.  $p < 0.05$
- D.  $p < 0.01$

Is there a significant difference between the two groups, which level of significance is correct?

Mean    standard deviation (SD)

6.93    0.806

5.21 - 1.299 pepe

$$1.72 \quad \text{pooled SE} = \sqrt{\left( \frac{0.806^2}{10} + \frac{1.299^2}{10} \right)} = 0.483.$$

The t-value =  $(6.93 - 5.21) / 0.483 = 3.56$ .

$20 - 2 = 18$  degrees of freedom, because we have 20 patients and 2 groups.

According to the t-table of page 16 the p-value is  $< 0.01$ , and we can conclude that a very significant difference exists between the two groups. The values of group 1 are higher than those of group 2. The answer D is correct.

## 6 T-Table

The t-table has a left-end column giving degrees of freedom ( $\approx$  sample sizes), and two top rows with p-values (areas under the curve = p - values), one-tail meaning that only one end of the curve, two-tail meaning that both ends are assessed simultaneously. The t-table is, furthermore, full of t-values, that, with  $\infty$  degrees of freedom, are equal to z-values (Chap.36). The t-values are to be understood as mean results of studies, but not expressed in mmol/l, kilograms, but in so-called SEM-units (Standard error of the mean units), that are obtained by dividing your mean result by its own standard error. With many degrees of freedom (large samples) the curve will be a little bit narrower, and more in agreement with nature.

df	One-Tail = .4 Two-Tail = .8	.25 .5	.1 .2	.05 .1	.025 .05	.01 .02	.005 .01	.0025 .005	.001 .002	.0005 .001
1	0.325	1.000	3.078	6.314	12.706	31.821	63.657	127.32	318.31	636.62
2	0.289	0.816	1.886	2.920	4.303	6.965	9.925	14.089	22.327	31.598
3	0.277	0.765	1.638	2.353	3.182	4.541	5.841	7.453	10.214	12.924
4	0.271	0.741	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	0.265	0.718	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	0.263	0.711	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8	0.262	0.706	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	0.261	0.703	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	0.260	0.700	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	0.260	0.697	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	0.259	0.695	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	0.259	0.694	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	0.258	0.692	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15	0.258	0.691	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16	0.258	0.690	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17	0.257	0.689	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	0.257	0.688	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19	0.257	0.688	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	0.257	0.687	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21	0.257	0.686	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	0.256	0.686	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	0.256	0.685	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.767
24	0.256	0.685	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	0.256	0.684	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26	0.256	0.684	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27	0.256	0.684	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28	0.256	0.683	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29	0.256	0.683	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30	0.256	0.683	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646
40	0.255	0.681	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551
60	0.254	0.679	1.296	1.671	2.000	2.390	2.660	2.915	3.232	3.460
120	0.254	0.677	1.289	1.658	1.980	2.358	2.617	2.860	3.160	3.373
∞	0.253	0.674	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291

A t-value of 3.56 with 18 degrees of freedom indicates, that we will need the 18th row of t-values. The upper row of the table gives the area under the curve of the Gaussian-like t-distribution. The t-value 3.56 is left from 4.297, right from 3.250. Now look right up at the upper row: we are right from 0.01. The p-value equals  $<0.01$ . The hours of sleep during the sleeping pill are significantly better than those during placebo.

## 7 Mann-Whitney test

Like the Wilcoxon test, being the non-parametric alternative for the paired t-test, the Mann-Whitney test is the non-parametric alternative for the unpaired t-test. Also this test is applicable for all kinds of data, and, therefore, particularly, to be recommended for investigators with little affection to medical statistics.

Calculate the p-value of the difference between two groups of 10 patients with the help of this test.

group 1:

6.0, 7.1, 8.1, 7.5, 6.4, 7.9, 6.8, 6.6, 7.3, 5.6,

group 2:

5.1, 8.0, 3.8, 4.4, 5.2, 5.4, 4.3, 6.0, 3.7, 6.2

- A. not significant
- B.  $0.05 < p < 0.10$
- C.  $p < 0.05$
- D.  $p < 0.01$

Is there a significant difference between the two groups? What significance level is correct?

All values are ranked together in ascending order of magnitude. The values from group 1 are printed thin, those from group 2 are printed fat. Add a rank number to each value. If there are identical values, for example, the rank numbers 9 and 10, then replace those rank numbers with average rank numbers, 9.5 and 9.5. Subsequently, all fat printed rank numbers are added up, and so are the thin printed rank numbers. We will find the values 142.5 for fat print, and 67.5 for thin print.

According to the underneath Mann-Whitney table. the difference should be larger than 71 in order for the significance level of difference to be  $<0.05$ . We find a difference of 75, which means that there is a p-value  $<0.05$  and that the difference between the two groups is, thus, significant.

3.7	1
3.8	2
4.3	3
4.4	4
5.1	5
5.2	6
5.4	7
<b>5.6</b>	<b>8</b>
<b>6.0</b>	<b>9.5</b>
6.0	9.5
6.2	11
<b>6.4</b>	<b>12</b>
<b>6.6</b>	<b>13</b>



### 9 Mann-Whitney Table $P < 0.05$

$P < 0.05$  levels

$P < 0.05$ levels	
$n_1 \rightarrow$	
$n_2 \downarrow$	2 3 4 5 6 7 8 9 10 11 12 13 14 15
5	15
6	10 16 23
7	10 17 24 32
8	11 17 25 34 43
9	6 11 18 26 35 45 56
10	6 12 19 27 37 47 58 71
11	6 12 20 28 38 49 61 74 87
12	7 13 21 30 40 51 63 76 90 106
13	7 14 22 31 41 53 65 79 93 109 125
14	7 14 22 32 43 54 67 81 96 112 129 147
15	8 15 23 33 44 56 70 84 99 115 133 151 171
16	8 15 24 34 46 58 72 86 102 119 137 155
17	8 16 25 36 47 60 74 89 105 122 140
18	8 16 26 37 49 62 76 92 108 125
19	3 9 17 27 38 50 64 78 94 111
20	3 9 18 28 39 52 66 81 97
21	3 9 18 29 40 53 68 83
22	3 10 19 29 42 55 70
23	3 10 19 30 43 57
24	3 10 20 31 44
25	3 11 20 32
26	3 11 21
27	4 11
28	4

### 10 Conclusion

For the study of two outcomes two parallel groups of similar age, gender and other characteristics are often applied, and the studies are called parallel-group studies, and the two groups are called independent of one another. Unpaired tests, like the unpaired t-test and Mann-Whitney test are appropriate for analysis.

### 11 Note

The theories of null-hypotheses and frequency distributions and additional examples of unpaired t-tests and Mann-Whitney tests are reviewed in Statistics applied to clinical studies 5th edition, Chaps. 1 and 2, entitled “Hypotheses data stratification” and “The analysis of efficacy data”, Springer Heidelberg Germany, 2012, from the same authors.