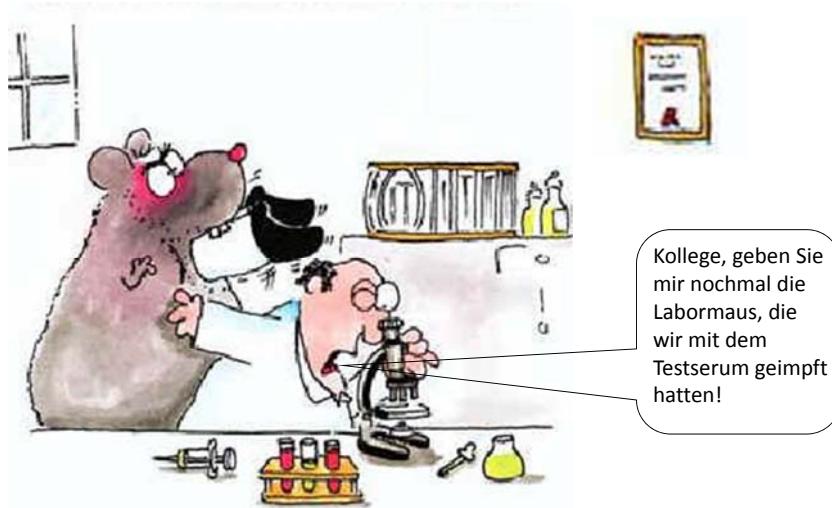


Hypothesenprüfungen. t-Tests



Schätzungen und Hypothesenprüfungen

Schätzungen

Wie gross ist eine Grösse?

Punktschätzungen

ein Wert ist gegeben und nichts über die Sicherheit

Parameter der Stichprobe

Parameter der Population

$$\begin{array}{l} \bar{X} \rightarrow \mu \quad (n \rightarrow \infty) \\ S \rightarrow \sigma \quad (n \rightarrow \infty) \end{array}$$

Intervall-schätzungen

95 % Konfidenzintervall für den Erwartungswert:

$$\bar{x} \pm 2 s_{\bar{x}}$$

ein Intervall ist mit einem Konfidenzniveau gegeben

(95 %) Referenzintervall:

$$\bar{x} \pm 2 s$$

Hypothesenprüfungen

Beantwortung einer Entscheidungsfrage

ja oder nicht mit einem Signifikanzniveau

2

Typische Entscheidungsfragen in der Medizin

- Ist die Therapie erfolgreich?
(Gibt es eine Änderung in der erwarteten Richtung?)
Hat eine Behandlung eine Wirkung?
Verkleinert ein Fiebermittel die Körpertemperatur?



- Gibt es einen Unterschied zwischen zwei Therapiemethoden?
- Gibt es eine Beziehung zwischen zwei Grössen?

3

Gibt es eine Wirkung einer Behandlung?

Verändert ein Fiebermittel die Körpertemperatur?



$$(a) \quad T_2 \neq T_1 \Leftrightarrow T_2 - T_1 \neq 0, \quad \text{z.B.: } T_2 - T_1 = -1.5 \text{ C}, \\ T_2 - T_1 = -2.1 \text{ C}, \\ T_2 - T_1 = +0.4 \text{ C}$$

...

$$(b) \quad T_2 \approx T_1 \Leftrightarrow T_2 - T_1 \approx 0$$

4

Die Nullhypothese

Es gibt keine Wirkung der Behandlung.

Die Wirkung der Behandlung ist Null (Nullhypothese, H_0).

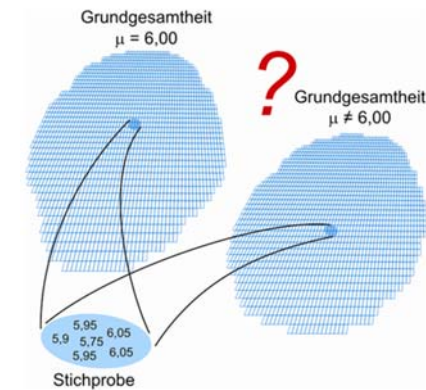
Das Fiebermittel verändert die Körpertemperatur nicht.

Wir müssen die Temperaturen in einer Gruppe (Stichprobe) messen.

Wenn die Nullhypothese richtig ist, müssen die Temperatur-differenzen um 0 streuen.
Alle Abweichungen von Null sind zufällig.

5

Die möglichen Grundgesamtheiten der Stichprobenentnahme



Wenn die Nullhypothese richtig ist, müssen die Daten der Stichprobe um den theoretischen Wert streuen. Alle Abweichungen von dem theoretischen Wert sind zufällig.

Pr.Buch Abb. 17

6

Die Alternativhypothese

Es gibt eine Wirkung der Behandlung.

Die Wirkung der Behandlung ist nicht Null (Alternativhypothese, H_1).

Das Fiebermittel verändert die Körpertemperatur.

Man unterscheidet als **Gegensatzpaar**
Nullhypothese und Alternativhypothese.

Entweder H_0 oder H_1 ist richtig.

Nehmen wir an, dass H_0 richtig ist!

Wenn Ergebnisse mit dieser Voraussetzung nicht passen: ablehnen wir H_0 richtig
 H_1 ist



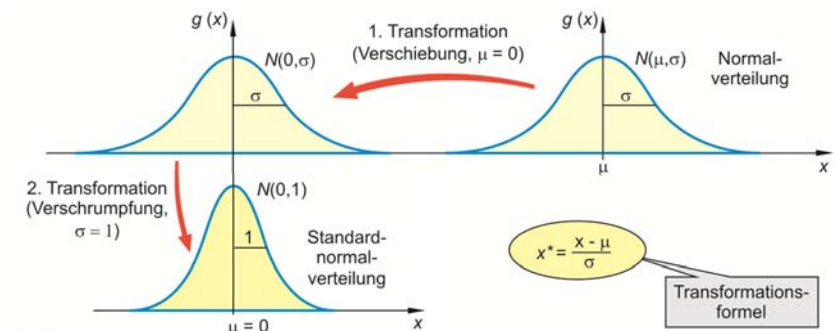
7

Transformation einer Normalverteilung mit allgemeiner Lage und Breite in eine Standardnormalverteilung

Mit welcher Verteilung sollen wir unsere Stichprobe vergleichen?

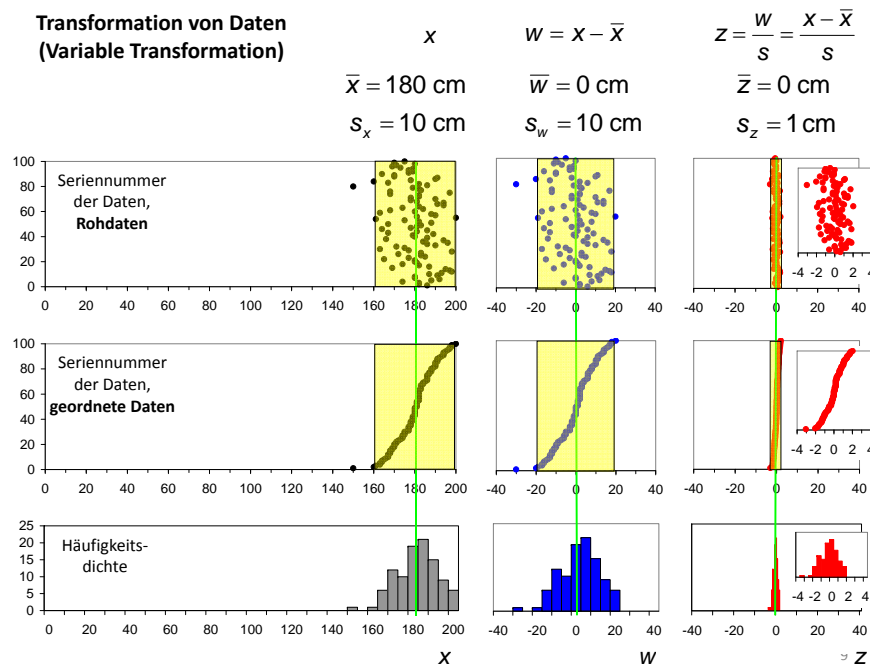
Die Standardnormalverteilung hat eine ausgezeichnete Rolle zwischen der Normalverteilungen.

Alle Normalverteilungen können in Standardnormalverteilung transformiert werden.



Pr.Buch Abb. 18

Transformation von Daten (Variable Transformation)



Einstichproben t-Test

Variable	x	$w = x - \mu$	$z = \frac{x - \mu}{\sigma}$
Verteilung	$N(\mu, \sigma)$	$N(0, \sigma)$	$N(0, 1)$

Wenn die originale Variable x zu einer **Normalverteilung** mit Parameter μ und σ gehört, dann gehört die transformierte Variable z zu der Standardnormalverteilung.

Wenn H_0 richtig ist, kennen wir den Wert von μ , aber σ nicht.

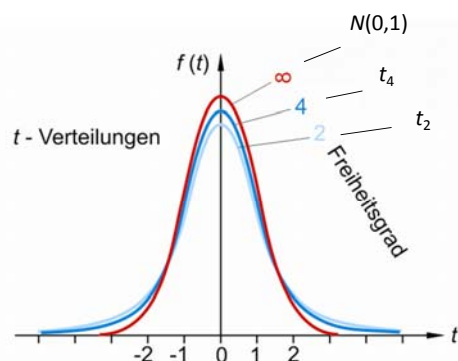
Die durchgeführte Transformation:

$$t = \frac{\bar{x} - \mu}{s_{\bar{x}}}$$

10

Variable x $z = \frac{x - \mu}{\sigma}$ $t = \frac{\bar{x} - \mu}{s_{\bar{x}}}$

Verteilung $N(\mu, \sigma)$ $N(0, 1)$ t_{n-1}



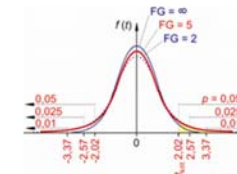
11

1. STATISTISCHE TABELLEN

t-Verteilungsfamilie

t-VERTEILUNG

Freiheitsgrad (FG)	p (Irrtumswahrscheinlichkeit, einseitiger Test)						
	0,4	0,25	0,1	0,05	0,025	0,01	0,005
	p (Irrtumswahrscheinlichkeit, zweiseitiger Test)						
	0,8	0,5	0,2	0,1	0,05	0,02	0,01
1	0,325	1,000	3,078	6,314	12,70	31,82	63,65
2	0,289	0,816	1,886	2,920	4,303	6,965	9,925
3	0,277	0,765	1,638	2,353	3,182	4,541	5,841
4	0,271	0,741	1,533	2,132	2,776	3,747	4,604
5	0,267	0,727	1,476	2,015	2,571	3,365	4,032
6	0,265	0,718	1,440	1,943	2,447	3,143	3,707
7	0,263	0,711	1,415	1,895	2,365	2,998	3,499



„Glockenkurven“

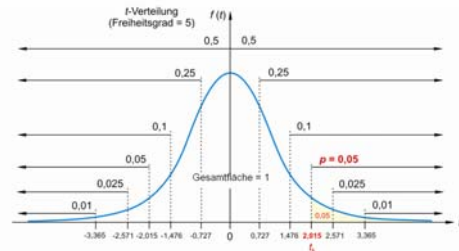
Je grösser ist der Freiheitsgrad, desto schmaler ist die Kurve.

$$t_{\infty} \equiv N(0, 1)$$

Kann der (aus der Stichprobe kalkulierte) t -Wert der t -Verteilung (mit entsprechendem Freiheitsgrad) gehören?

Alle Werte können zu der t -Verteilung gehören.
Aber: Wenn der t -Wert gross ist, dann ist die Wahrscheinlichkeit klein.

Deswegen benützen wir nicht die gesamte t -Verteilung, sondern eine abgestutzte t -Verteilung!



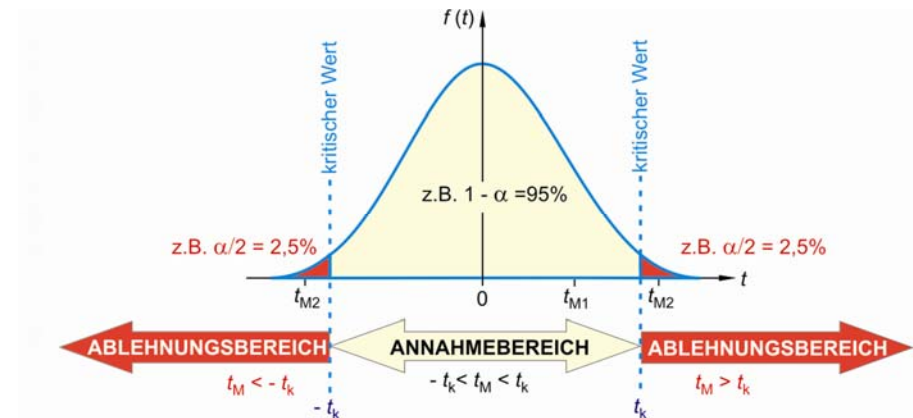
akzeptierbare Irrtumswahrscheinlichkeit in der Medizin: gleich 5 %

kleiner oder

13

Zweiseitiger t -Test

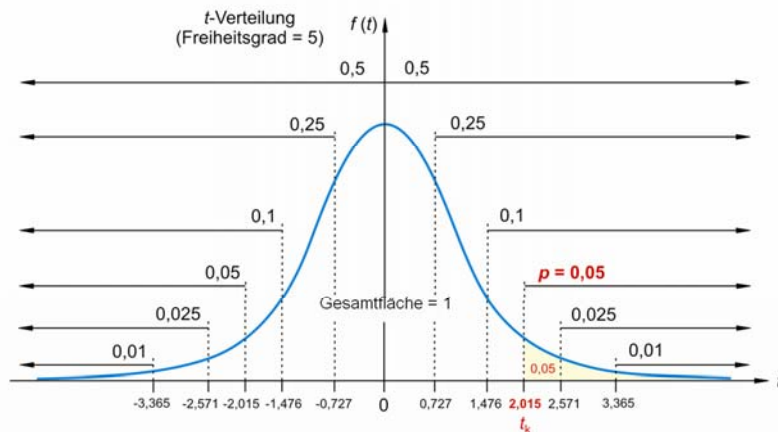
Verändert ein Fiebermittel die Körpertemperatur?



Pr.Buch Abb. 22

14

t -Verteilungskurve mit Freiheitsgrad 5. Die kritischen Werte und Wahrscheinlichkeiten des einseitigen t -tests



Pr.Buch Abb. 25

15

H_0 abgelehnt, obwohl richtig

ENTSCHEIDUNG AUF DIE WAHRE SITUATION	Nullhypothese ist angenommen		Nullhypothese ist abgelehnt	
	Nullhypothese ist angenommen		Nullhypothese ist abgelehnt	
Nullhypothese ist wahr	<p>RICHTIGE ENTSCHEIDUNG</p> <p>t-Verteilung</p> <p>$1-\alpha$</p> <p>t_k kritischer Wert</p>		<p>FEHL-ENTSCHEIDUNG (Fehler 1. Art)</p> <p>t-Verteilung</p> <p>$1-\alpha$</p> <p>t_k kritischer Wert</p> <p>Fehler 1. Art (bekannt)</p>	
Nullhypothese ist falsch	<p>FEHL-ENTSCHEIDUNG (Fehler 2. Art)</p> <p>Fehler 2. Art (unbekannt)</p> <p>$1-\beta$</p> <p>t_k kritischer Wert</p> <p>unbekannte Verteilung</p>		<p>RICHTIGE ENTSCHEIDUNG</p> <p>t-Verteilung</p> <p>$1-\beta$</p> <p>t_k kritischer Wert</p> <p>unbekannte Verteilung</p>	

H_0 angenommen, obwohl falsch

Pr.Buch Abb. 24

16

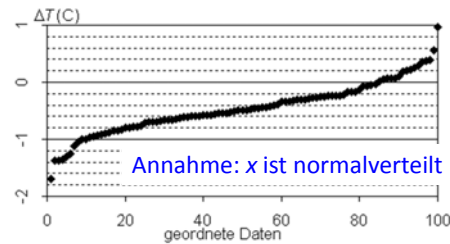
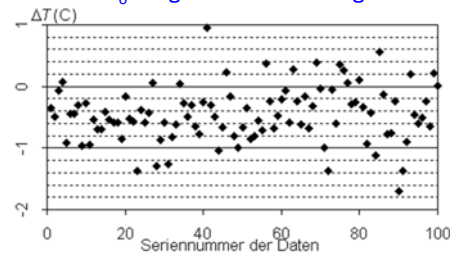
x = Temperaturdifferenzen

-0.35 -0.52 0.96 -0.07 -0.34
 -0.50 -0.57 -0.31 -0.58 -0.94
 -0.07 -1.37 -0.49 0.28 -0.43
 0.07 -0.39 -1.05 -0.24 -1.12
 -0.92 -0.58 -0.66 -0.62 0.56
 -0.44 -0.43 0.23 -0.16 -0.13
 -0.45 0.06 -0.16 -0.69 -0.78
 -0.30 -1.30 -0.80 -0.31 -0.77
 -0.97 -0.88 -1.00 0.38 -0.24
 -0.28 -0.58 -0.67 -0.04 -1.69
 -0.95 -1.26 -0.34 -1.00 -1.37
 -0.55 -0.82 -0.86 -1.36 -0.90
 -0.70 -0.61 -0.80 -0.05 0.19
 -0.70 0.05 -0.55 -0.60 -0.46
 -0.41 -0.27 -0.71 0.36 -0.61
 -0.54 -0.50 0.38 0.26 -0.50
 -0.59 -0.31 -0.24 0.06 -0.24
 -0.59 -0.65 -0.68 -0.29 -0.65
 -0.85 -0.77 -0.47 -0.25 0.21
 -0.17 -0.26 -0.22 0.10 0.01

Beispiel: Einstichproben t-Test

Verändert ein Fiebermittel die Körpertemperatur?

H_0 : es gibt keine Wirkung



Annahme: x ist normalverteilt

17

Kalkulation:

$$t_{n-1} = \frac{\bar{X} - \mu}{S_{\bar{X}}}, \quad \mu = 0$$

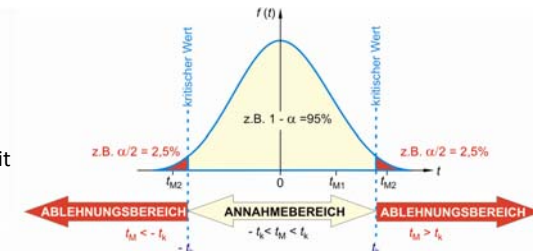
$$t_{n-1} = \frac{\bar{X}}{S_{\bar{X}}} = \frac{\bar{X}}{S} \sqrt{n}$$

zweiseitiger Test

$$|t| > t_{\text{krit}} \rightarrow$$

wir ablehnen die Nullhypothese mit einem Signifikanzniveau von 5%

Anzahl der Daten	n	100
Durchschnitt	avg	-0.457
Standardabweichung	stdev	0.454
Standardfehler	sem	0.045
t-Wert	t	-10.056
Freiheitsgrad	df	99
max. zulässige Irrtumswahrscheinlichkeit	α	0.05
kritischer t-Wert	t_{krit}	1.984



18

$$|t| > t_{\text{krit}} \rightarrow$$

Das Fiebermittel signifikant verändert (verkleinert) die Körpertemperatur ($p \leq 0.05$).

Im Klammer steht die Irrtumswahrscheinlichkeit. Es gibt die Wahrscheinlichkeit, dass die Nullhypothese richtig ist. In diesem Fall unsere Klassifikation ist falsch (Fehler 1. Art).

Die Nullhypothese abgelehnt wird, obwohl sie richtig ist.

Freiheitsgrad (FG)	p (Irrtumswahrscheinlichkeit, einseitiger Test)						
	0.4	0.25	0.1	0.05	0.025	0.01	0.005
1	0.325	1,000	3,078	6,314	12,70	31,82	63,65
	0,200	0,631	1,000	1,250	1,650	2,000	2,576
60	0,255	0,679	1,296	1,671	2,000	2,390	2,66
	0,254	0,677	1,289	1,658	1,980	2,358	2,617
∞	0,250	0,674	1,282	1,645	1,960	2,326	2,576

weitere Bemerkungen:

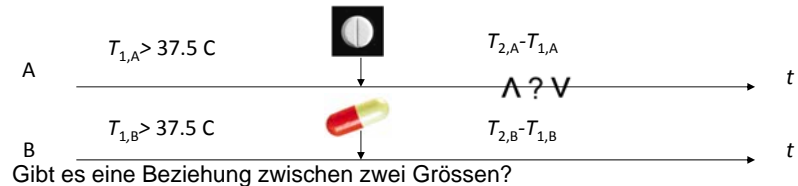
$$|t| = 10.056 > 2.66$$

$p \leq 0.01$ (zweiseitiger Test)

19

Typische Entscheidungsfragen in der Medizin

- Ist die Therapie erfolgreich?
(Gibt es eine Änderung in der erwarteten Richtung?)
Hat eine Behandlung eine Wirkung?
Verkleinert ein Fiebermittel die Körpertemperatur?
- Gibt es einen Unterschied zwischen zwei Therapiemethoden?



- Gibt es eine Beziehung zwischen zwei Größen?

20

Einstichproben t-Test

$$t_{n-1} = \frac{\bar{X} - \mu_0}{S_{\bar{X}}} = \frac{\bar{X} - \mu_0}{S} \sqrt{n}, \text{ wo}$$

$$s = \sqrt{\frac{Q}{n-1}}$$

Zweistichproben t-Test

$$t_{n_1+n_2-2} = \frac{\bar{X}_1 - \bar{X}_2}{S^*} \cdot \sqrt{\frac{n_1 n_2}{n_1 + n_2}}, \text{ wo}$$

$$s^* = \sqrt{\frac{Q_1 + Q_2}{n_1 + n_2 - 2}}$$

Vergleichen wir die Formeln!

$$t_{n-1} = \frac{\bar{X}}{s \sqrt{\frac{1}{n}}}$$

Einstichproben

$$t_{n_1+n_2-1} = \frac{\bar{X}_1 - \bar{X}_2}{s^* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Zweistichproben

21

Wirkung?:

Effekt der 5 Kniebeugen auf die Pulsfrequenz

Einstichproben t-Test

n: Pulsfrequenz (1/30s), 1: vor, 2 nach, d: Differenz

H_0 : keine Wirkung

				Zweistichproben t-Test bei abhängigen Stichproben (Paarvergleichstest)		
	n1	n2	dn		Variable 1	Variable 2
m	36	37	1			
m	33	35	2			
m	34	37	3			
m	33	37	4	Mittelwert	37.142857	42.285714
m	34	38	4	Varianz	43.516484	80.681319
m	37	41	4	Beobachtungen	14	14
m	37	41	4	Pearson Korrelation	0.9196855	
m	37	46	9	Hypothetische Differenz der Mittelwerte	0	
w	30	29	-1	Freiheitsgrade (df)	13	
w	45	49	4	t-Statistik	-4.9342498	
w	50	55	5	P(T<=t) einseitig	0.0001365	
w	35	40	5	Kritischer t-Wert bei einseitigem t-Test	1.7709334	
w	28	37	9	P(T<=t) zweiseitig	0.000273	
w	39	48	9	Kritischer t-Wert bei zweiseitigem t-Test	2.1603687	
w	49	63	14			

$|t| = 4.934 > t_{13, \text{krit}(0,05)} = 2.160 \Rightarrow H_0$ ist falsch ($p < 0.05$)

($p = 0.000273$)

22

Zweistichproben t-Test

Gibt es einen Unterschied zwischen zwei (Therapie)methoden?

m: männlich, w: weiblich

				Zweistichproben t-Test unter der Annahme gleicher Varianzen		
	n1	n2	dn		Variable 1	Variable 2
m	36	37	1			
m	33	35	2			
m	34	37	3	Mittelwert	3.8571429	6.4285714
m	33	37	4	Varianz	6.4761905	22.619048
m	34	38	4	Beobachtungen	7	7
m	37	41	4	Gepoolte Varianz	14.547619	
m	37	46	9	Hypothetische Differenz der Mittelwerte	0	
w	30	29	-1	Freiheitsgrade (df)	12	
w	45	49	4	t-Statistik	-1.261283	
w	50	55	5	P(T<=t) einseitig	0.1155878	
w	35	40	5	Kritischer t-Wert bei einseitigem t-Test	1.7822875	
w	28	37	9	P(T<=t) zweiseitig	0.2311755	
w	39	48	9	Kritischer t-Wert bei zweiseitigem t-Test	2.1788128	
w	49	63	14			

$|t| = 1.261 < t_{12, \text{krit}(0,05)} = 2.179$

$\Rightarrow H_0$ ist richtig

23

				Zweistichproben t-Test bei abhängigen Stichproben (Paarvergleichstest)		
	n1	n2	dn		Variable 1	Variable 2
m	36	37	1			
m	33	35	2			
m	34	37	3	Mittelwert	37.142857	42.285714
m	33	37	4	Varianz	43.516484	80.681319
m	34	38	4	Beobachtungen	14	14
m	37	41	4	Pearson Korrelation	0.9196855	
m	37	46	9	Hypothetische Differenz der Mittelwerte	0	
w	30	29	-1	Freiheitsgrade (df)	13	
w	45	49	4	t-Statistik	-4.9342498	
w	50	55	5	P(T<=t) einseitig	0.0001365	
w	35	40	5	Kritischer t-Wert bei einseitigem t-Test	1.7709334	
w	28	37	9	P(T<=t) zweiseitig	0.000273	
w	39	48	9	Kritischer t-Wert bei zweiseitigem t-Test	2.1603687	
w	49	63	14			

H_0 : keine Wirkung

H_0 ist falsch

H_0 : keine Differenz zwischen der Wirkungen

H_0 ist richtig

24

Freiheits- grad (FG)	p (Irrtumswahrscheinlichkeit, einseitiger Test)						
	0,4	0,25	0,1	0,05	0,025	0,01	0,005
	p (Irrtumswahrscheinlichkeit, zweiseitiger Test)						
	0,8	0,5	0,2	0,1	0,05	0,02	0,01
1	0,325	1,000	3,078	6,314	12,70	31,82	63,65
2	0,289	0,816	1,886	2,920	4,303	6,965	9,925
3	0,277	0,765	1,638	2,353	3,182	4,541	5,841
4	0,271	0,741	1,533	2,132	2,776	3,747	4,604
5	0,267	0,727	1,476	2,015	2,571	3,365	4,032
6	0,265	0,718	1,440	1,943	2,447	3,143	3,707
7	0,263	0,711	1,415	1,895	2,365	2,998	3,499
8	0,262	0,706	1,397	1,860	2,306	2,896	3,355
9	0,261	0,703	1,383	1,833	2,262	2,821	3,250
10	0,260	0,700	1,372	1,812	2,228	2,764	3,169
11	0,260	0,697	1,363	1,796	2,201	2,718	3,106
12	0,259	0,695	1,356	1,782	2,179	2,681	3,055
13	0,259	0,694	1,350	1,771	2,160	2,650	3,012
14	0,258	0,692	1,345	1,761	2,145	2,624	2,977
15	0,258	0,691	1,341	1,753	2,131	2,602	2,947

25

Einstichproben t-Test

$$|t| = 4.934 > t_{13, \text{krit}(0,05)} = 2.160 \Rightarrow H_0 \text{ ist falsch } (p < 0.05)$$

$$|t| = 4.934 > t_{13, \text{krit}(0,01)} = 3.012 \Rightarrow H_0 \text{ ist falsch } (p < 0.01)$$

$$|t| = 4.934 \geq t_{13, \text{krit}(0.00027)} = 4.934 \Rightarrow H_0 \text{ ist falsch } (p \leq 0.00027)$$

5 Kniebeugen verursachen Veränderung der Pulsfrequenzen mit einem Signifikanzniveau von 5 % (sogar: 1 %, ..., 0.027%)

Zweistichproben t-Test

$$|t| = 1.261 < t_{12, \text{krit}(0,05)} = 2.179 \quad H_0 \text{ ist richtig}$$

Es gibt keine signifikante Differenz zwischen der Pulsfrequenzen in der Männer- und Frauengruppen nach 5 Kniebeugen.

26