2 Test auf den Mittelwert μ

Frage: Wird ein Sollwert signifikant unterschritten oder übertroffen?

Voraussetzung:
Die Stichprobe stammt aus einer $N(\mu, \sigma)$-normalverteilten Population.

1. Fall: σ ist bekannt

Einstichprobentest Normalverteilung

Beispiel:
Einhalten von Füllgewichten (Mittag/Schüller p. 272, B 16.3)

In einer Fabrik wird Zucker in 2kg-Säcke abgefüllt.
Aus Voruntersuchungen ist bekannt, dass die Zufallsvariable X des Inhalts normalverteilt mit der Standardabweichung $\sigma = 0.01\ kg$ ist. Es soll geprüft werden, ob das Sollgewicht $\mu_0 = 2$ systematisch unterschritten wird und die Maschine neu eingestellt werden muss.

1, Nullhypothese H0: $\mu_0 = 2$
2, Alternativhypothese HA: $\mu < \mu_0 = 2$

3. Testgrösse $z = \frac{\bar{x}-2}{0.01} \cdot \sqrt{10} = (1.996 - 2) \cdot \sqrt{10} \approx -1.265$

4. Zufallsexperiment:
Eine Zufallsstichprobe vom Umfang $n = 10$ ergibt $\bar{x} = 1.996$.

Kritischer Wert bei einseitigem Test und Signifikanzniveau 5% ist $z_{0.05} = -z_{0.95} \approx -1.65$

Schlussfolgerung:
Da die Testgrösse den kritischen Wert nicht unterschreitet, kann die Nullhypothese nicht abgelehnt werden. Die Differenz zum Sollgewicht ist zufällig und statistisch nicht signifikant.

Der sogenannte p-Wert ist das Niveau α' für das $z_{\alpha'} = -1.265$ gilt.
Mit einer Tabelle erhält man $\Phi(1.265) \approx 0.897$ und daraus $\alpha' \approx 1 - 0.897 \approx 0.103$ mit dem Fazit, dass die Nullhypothese nicht abgelehnt werden kann.

Bemerkung
Mit der Tendenz in wissenschaftlichen Publikationen eher Testergebnisse zu publizieren, bei denen die Nullhypothese abgelehnt werden kann, entsteht eine Publikationsverzerrung. Dies kann in der Medizin etwa dazu führen, dass die Wirksamkeit von Medikamenten oder Therapien ¨überschätzt» wird.

2, Fall: σist unbekannt

σ kann durch den erwartungstreuen Schätzer

$$\hat{\sigma}^2 = \frac{1}{n-1} \cdot \left(\sum_i X_i^2 - n \cdot \bar{X}^2 \right)$$

geschätzt werden.

Es können die folgenden drei Fälle auftreten:

2.1
t-Test für eine Stichprobe
2.2
t-Test für zwei verbundene (abhängige Stichproben)
2.3
t-Test für 2 unabhängige Stichproben

2.1 t-Test für eine Stichprobe

Fragetyp: Wird ein Sollwert signifikant überschritten oder unterschritten?

In diesem Fall ist die folgende Testzufallsvariable T Student t-verteilt mit dem Freiheitsgrad $n-1$ d,h,

$$T = \frac{\bar{x} - \mu}{\hat{\sigma}} \cdot \sqrt{n} \sim t(n-1)$$

Beispiel: Oktoberfest
Am Oktoberfest wird jährlich von einem Verein mit Stichproben überprüft, ob die 1 Liter Masskrüge aus 12 Festzelten eine genügende Füllhöhe Bier enthalten. Toleriert wird eine maximale Abweichung von 10 mm unter den Eichstrich.

Die Ergebnisse sind in der folgenden Tabelle auf der folgenden Seite angegeben.
1.
Alternativhypothese H_A: der Mittelwert der Füllhöhe ist zu klein: $\mu < 0.9$ Die Lady hat

2.
Nullhypothese H_0: Die Füllhöhe ist 0.9 Liter (oder grösser)

3.
Planen eines Zufallsexperiments

Messungen der mittleren Füllhöhe in 12 Festzelten

4.
Es wird für die Zufallsvariable X (gemessene Füllhöhe).
Vorgegebenes **Signifikanzniveau 5%**

5.
Das Zufallsexperiment wird durchgeführt und der Wert der Zufallsvariable bestimmt.

Abfüllmenge von Bier auf dem Oktoberfest 2013

Vorgabe: 1 Litermass enthält mindestens 0.9 Liter
Indikatorvariable ist binomialverteilt

Festzelt	Mittelwerte (in Liter)	erfüllt
1	0.87	0
2	0.86	0
3	0.88	0
4	0.86	0
5	0.92	1
6	0.90	1
7	0.8	0
8	0.94	1
9	0.87	0
10	0.81	0
11	0.84	0
12	0.84	0

Summe	10.39
Mittelwert X_quer	0.86583333
emp. Standardabweichung	0.04122187

n	12
Mittelwert X_quer	0.86583333
Erwartungswert müh 0	0.9
Standardisierung	-0.0341667
Testvariable T	**-2.871214**

Als Wert der Prüfgrösse erhält man $t = -2.871$
Bei einer Signifikanz von $\alpha = 0.05$ ist bei einem Freiheitsgrad $df = n - 1$
der t-Wert $-t_{11;0.95} = -1.796$
Damit ist die Nullhypothese deutlich abgelehnt, denn der p-Wert für eine so grosse Abweichung beträgt nur 0.0076.
Die Füllhöhen liegen also bei statistischer Betrachtung deutlich
unter der Toleranzgrenze 0.9 Liter

6.
Fehlermöglichkeiten:

Fehler 1. Art α:

Bei diesem Beispiel wollte man den Fehler möglichst klein halten, die Festzelte fälschlicherweise zu verdächtigen.
Ist die Voraussetzung der Normalverteilung nicht erfüllt, so kann der **Wilcoxon-Vorzeichen-Rang-Test** durchgeführt werden.

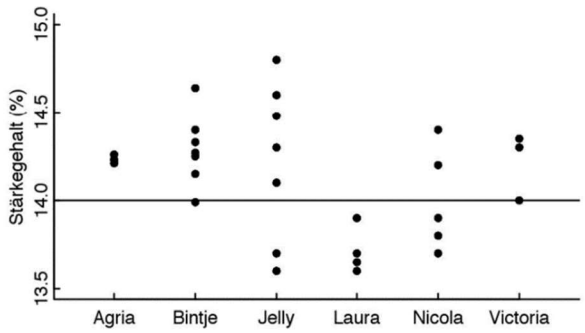
Übungsaufgabe:
Es ist bei Kartoffelsorten zu prüfen, ob der mittlere Stärkegehalt grösser als 14% beträgt. Damit ist nämlich eine Sorte für die Herstellung von Chips geeignet.

1 Hypothese H_A: Der mittlere Stärkegehalt grösser % als 14%
2 Hypothese H_0: Der mittlere Stärkegehalt ist 14%

3 Planen eines Zufallsexperiments

Es werden die sechs in der Tabelle und der Grafik aufgeführten Sorten geprüft.

4.
Zufallsvariable X ist der Stärkegehalt.
Vorgegebenes **Signifikanzniveau 5%**
Die verschiedenen Ergebnisse in einem Acker sind in der Abbildung links durch einen Punkt dargestellt. In der Tabelle rechts sind der mittlere Stärkegehalt, die Standardabweichung und der Stichprobenumfang angegeben.



Sorte	$\bar{x}$	s	n
Agria	14.23	0.025	3
Bintje	14.29	0.203	7
Jelly	14.23	0.454	7
Laura	13.71	0.131	4
Nicola	14.00	0.296	5
Victoria	14.22	0.189	3

Die Frage der Eignung wird für die Sorte Jelly durchgeführt, für die die einzelnen 7 Stichprobenergebnisse bekannt sind:

Ergebnis:
Diese Sorte «Bintje» Sorte eignet sich für die Herstellung von Chips, denn ihr mittlerer Stärkegehalt beträgt 14.29 %. Er liegt signifikant über 14%.
Für den t-Wert erhält man nämlich den Wert $t = 3.78$ $(n = 7, df = 6)$.
Der zugehörige p-wert 0.5% liegt deutlich unter 5%.

	Jelly	Bintje
1	13.6	
2	13.7	
3	14.1	
4	14.3	
5	14.5	
6	14.6	
7	14.8	
$\bar{x}$	14.229	14.29
s	0.454	0.203
t	1.333	3.780
	11.5%	0.5%

Für die Sorte «Jelly» ergibt sich zwar ein mittlerer Stärkegehalt von 14.23%. Ihre Eignung kann aber nicht nachgewiesen werden, denn der Stärkegehalt liegt nicht signifikant über 14%. Für den t-Wert erhält man nämlich den Wert $t = 1.33$ $(n = 7, df = 6)$. Der zugehörige p-Wert 11.5% liegt deutlich über 5%.

2.2 t-Test für zwei abhängige (gepaarte, verbundene) Stichproben

Fragestellung:
Unterscheiden sich die beiden gepaarten Stichproben? Gepaart bedeutet z.B,
- dass man eine Messung zu verschiedenen Zeitpunkten durchführt,
- dass man aus einer Population zufällig möglichst ähnliche Paare bildet und bei ihnen ein anderes Merkmal untersucht.
- dass man bei einem Paar die Wirkung von zwei Behandlungen untersucht

Die Fragestellung kann auf den Fall einer Stichprobe zurückgeführt werden, indem man die Differenzen bildet.

Dazu ein Beispiel aus Stahel: Angewandte Statistik:
Eine Rfenfirma will für Winterreifen zwei Profile entwickeln. Sie möchte deren Bremswirkung vergleichen. Dazu werden zehn Testfahrzeuge bei gleicher Geschwindigkeit abgebremst, einmal mit Profil A und einmal mit Profil B. Die gemessenen Bremswege sind in der folgenden Tabelle eingetragen:

Es sind die beiden Hypothesen bei einem Signifikanziveau 5% zu testen:

HA: $d \neq 0$
Die beiden Bremswege unterscheiden sich signifikant
Ihre Differenzen sind ungleich 0.

H0: $d = 0$

Testgrösse: $T = \frac{\bar{d} - d_0}{s_{\bar{d}}}$ mit $s_{\bar{d}} = \frac{s_d}{\sqrt{n}}$

Der kritische t-Wert für den Freiheitsgrad $df = 9$ bei einem Signifikanzniveau 5% (beidseitig) liegt bei 2.262
(Tabelle t-Verteilung oder Excel).

i	Profil A	Profil B	Diff
1	44.5	44.9	0.4
2	55.0	54.8	-0.2
3	52.5	55.6	3.1
4	50.2	55.2	5.0
5	45.3	55.6	10.3
6	46.1	47.7	1.6
7	52.1	53.0	0.9
8	50.5	49.1	-1.4
9	50.6	52.3	1.7
10	49.2	50.7	1.5
Summe			22.9
emp. Mittelwert	$\bar{x}$	**2.29**	
Stand.abw	s_x	**3.315**	
Standardfehler		**1.048**	
df		**9**	
Test-Statistik		**2.184**	
Tabellenwert für df = 9		**2.26**	
Die Nullhypothese wird also nicht verworfen			

Ergebnis:
Damit kann die Nullhypothesen nicht verworfen werden. Die beiden Reifen unterscheiden sich also bezüglich der Bremswirkung nicht signifikant.

Da σ der Population nicht bekannt ist, also aus den Daten geschätzt werden muss, ist die t-Verteilung massgebend.

Zum Vergleich:
Der kritische Wert 1.96 bei Normalverteilung führt zu einer Ablehnung der Nullhypothese.

Beispiel: Vergleich zweier Methoden
Vor der Untersuchung wurden Paare von Schülerinnen mit gleichem IQ gebildet. Damit sollte die Störvariable IQ kontrolliert werden. Je einer Schülerin des Probandenpaares wurde dann zufällig eine Methode (zu. Methode 1) zugewiesen, der anderen Schülerin die Methode 2. Es wurde davon ausgegangen, dass die Methode 1 besser abschneiden würde als Methode 2

Es sind die beiden Hypothesen bei einem Signifikanziveau 5% zu testen:

H_A: $d > 0$
Methode 1 ist besser als Methode 2

H_0: $d = 0$
Die beiden Methoden unterscheiden sich nicht

Abhängige Stichproben

Paar	Methode 1	Methode 2	Differenzen D
1	15	5	10
2	45	36	9
3	16	18	-2
4	41	25	16
5	7	10	-3
6	48	40	8
7	46	43	3
8	37	30	7
9	40	35	5
10	35	29	6
Summe	**330**	**271**	59
	33	**27.1**	
	14.76	**12.65**	

5.9 Mittelwert der Differenzen

31.66 Varianz der Differenzen

1.78 Standardfehler der Differenzen

3.32 Prüfgrösse

Der kritische Wert beim t-Test mit $df = 9$ Freiheitsgraden und einem Signifikanzniveau von 5% (einseitig) ist 1.83. Die Prüfgrösse 3.32 ist bedeutend grösser.

Die nach
Methode 1 (Mittelwert **33.0**, Standardabweichung **14.76**)
Unterrichteten schneiden signifikant besser ab als die nach
Methode 2 (Mittelwert **27.1**, Standardabweichung **12.65**)

Beispiel für eine andere Fragestellung:
Hat der Einsatz eines neuen Medikaments eine bessere Wirkung als das bisherige?

2.3 t-Test für 2 unabhängige Stichproben mit verschiedenen Varianzen

Fragestellung:
Sind die arithmetischen Mittel μ_1 und μ_2 eines Merkmals in zwei unabhängigen Grundgesamtheiten signifikant verschieden.

Als Testgrösse dient die Differenz der Stichprobenmittelwerte $d = \bar{x}_1 - \bar{x}_2$

Es sind die beiden Hypothesen bei einem Signifikanziveau 5% zu testen:

H_A: $d \neq 0$ (oder $d > 0$ oder $d < 0$)
Die beiden Mittelwerte unterscheiden sich.

H_0: $d = 0$
Die beiden Mittelwerte unterscheiden sich nicht

Führt man einen t-Test für 2 unabhängige Stichproben durch, so sind zwei Fälle zu unterscheiden:

Varianzhomogenität
Die beiden Stichproben stammen aus Grundgesamtheiten mit (annähernd) gleicher Varianz. Diese Bedingung ist wohl in den meisten Fällen nicht erfüllt. Die Varianzhomogenität kann auch grafisch mit einem Boxplot der beiden Gruppen oder mit dem sogenannten **Levenetest** überprüft werden.
Näheres dazu diesem Fall findet man bei UZH Methodenberatung Statistik.

oder
Varianzheterogenität.
Die beiden Stichproben stammen aus Grundgesamtheiten mit verschiedenen Varianzen.

Der sogenannte Welchtest ermöglicht eine Lösung in beiden Fällen. Bei der Berechnung der Testgrösse ist der Freiheitsgrad nicht unbedingt ganzzahlig. Die Testgrösse ist allerdings nur annähernd t-verteilt.

Welchtest

Zunächst wird der Standardfehler der Mittelwertsdifferenz geschätzt.

$$\hat{\sigma}_{\bar{x}_1-\bar{x}_2} = S_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{S_{X_1}^2}{n_1} + \frac{S_{X_2}^2}{n_2}}$$

Im Unterschied zum t-Test wird keine gepoolte Varianz berechnet, sondern die Varianzen werden beim Freiheitsgrad berücksichtigt:

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{s_1^4}{n_1^2(n_1-1)} + \frac{s_2^4}{n_2^2(n_2-1)}}$$

Der Freiheitsgrad ist damit in der Regel eine Dezimalzahl.

Testgrösse:

$$t = \frac{\bar{x}_1 - \bar{x}_2 - \omega_0}{s_{\bar{x}_1-\bar{x}_2}} = \frac{\bar{x}_1 - \bar{x}_2 - \omega_0}{\sqrt{\frac{s_{X_1}^2}{n_1} + \frac{s_{X_2}^2}{n_2}}} \sim t(df)$$

In den meisten Fällen $\omega_0 = 0$.
Die Testgrösse ist vor allem für kleine n nur annähernd t-verteilt. Mit steigendem Freiheitsgrad konvergiert sie aber schnell gegen die Normalverteilung.

Ein weiterer Vorteil des Welchtests besteht darin, dass die Stichproben nicht aus einer normalverteilten Population stammen müssen.

Es folgt ein Rechenbeispiel zum Welch-Test

Daten		1. Stichprobe		2. Stichprobe
Stichprobenumfang	1)	15	2)	17
Mittelwert der Stichproben	3)	70.3	4)	57.4
Varianz	5)	17.64	6)	9.61
Differenz der Mittelwerte	8)	12.9	3) und 4)	
Standardfehler der Mittelwertsdifferenz	9)	1.3196	1), 5) und 2), 6)	
Testgrösse aus den Daten		9.78	8), 9)	
(in den meisten Fällen ist ω = 0)				
		$\frac{s_1^2}{n_1}$	$\frac{s_2^2}{n_2}$	
	1), 5) und 2), 6)	1.176	0.565	
Hilfsgrössen				
Zähler Freiheitsgrad		3.032		
		0.084	0.033	
Nenner Freiheistgrad		0.117		
Freiheitsgrad		25.86		

$$\hat{\sigma}_{\bar{x}_1-\bar{x}_2} = S_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{S_{X_1}^2}{n_1} + \frac{S_{X_2}^2}{n_2}}$$

$$t = \frac{\bar{x}_1 - \bar{x}_2 - \omega_0}{\sqrt{\frac{S_{X_1}^2}{n_1} + \frac{S_{X_2}^2}{n_2}}} \sim t(df)$$

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{s_1^4}{n_1^2(n_1-1)} + \frac{s_2^4}{n_2^2(n_2-1)}}$$

signifikanzniveau		5%		Excel
kritisches t aus Tabelle	beidseitig	2.06	7), 12)	T.INV(0.975; 25)
Testgrösse aus den Daten		9.78		
Da die Testgrösse 9.78 grösser als der kritische Wert ist, wird man die Nullhypothese verwerfen				

Aufgabe (BMS gibb):
Gemäss der EZB (Europäische Zentralbank ist die Masse einer 1-Euromünze 7.5 Gramm.
Im Rahmen eines Tests wurden mehrere Packungen mit je 250 Münzen geprüft.
In der folgenden Tabelle sind die Daten von zwei Stichproben aufgeführt.
Besteht ein signifikanter Unterschied zwischen den Mittelwerten in den beiden Stichproben?
Signifikanzniveau: 5%

Daten			1. Stichprobe		2. Stichprobe
Stichprobenumfang		1)	250	2)	250
Mittelwert der Stichproben		3)	7.52	4)	7.52
Varianz		5)	0.00111556	6)	0.0010890
Differenz der Mittelwerte		8)	-0.0078	3) und 4)	
Standardfehler der Mittelwertsdifferenz		9)	0.00296955215	1), 5) und 2), 6)	
Testgrösse aus den Daten			-2.6267	8), 9)	
(in den meisten Fällen ist ω = 0)					
			$\frac{s_1^2}{n_1}$	$\frac{s_2^2}{n_2}$	
		1), 5) und 2), 6)	0.00000446	0.00000436	

$$\hat{\sigma}_{\bar{x}_1-\bar{x}_2} = S_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{S_{\bar{x}_1}^2}{n_1} + \frac{S_{\bar{x}_2}^2}{n_2}}$$

$$t = \frac{\bar{x}_1 - \bar{x}_2 - \omega_0}{\sqrt{\frac{s_{\bar{x}_1}^2}{n_1} + \frac{s_{\bar{x}_2}^2}{n_2}}} \sim t(df)$$

Hilfsgrössen					
Zähler Freiheitsgrad			0.000000000077761		
			0.000000000000080	0.000	
Nenner Freiheitsgrad			0.000000000000156		
Freiheitsgrad			497.928		
signifikanzniveau			5%		Excel
kritisches t aus Tabelle		beidseitig	1.965	7), 12)	T.INV(0.975; 18)
Testgrösse aus den Daten			-2.63		
Da der Betrag der Testgrösse (2.63) grösser als der kritische Wert(1.965) ist, wird man die Nullhypothese verwerfen.					
Die Diifferenz der beiden Stichproben ist auffällig gross.					

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{s_1^4}{n_1^2(n_1-1)} + \frac{s_2^4}{n_2^2(n_2-1)}}$$

Beispiel:
Zwei Gruppen mit 20 bzw. 11 Teilnehmerinnen nahmen an einem Gedächtnistest teil. Es mussten möglichst viele vorgegebene Wörter gemerkt werden. Die erzielten Leistungen sind in der Tabelle dargestellt:
Frage:
Unterscheiden sich die beiden Erwartungswerte?

HA: Die beiden Erwartungswerte unterscheiden sich; $\mu_1 \neq \mu_2$

H0: $\mu_1 = \mu_2$

	Gruppe 1	Gruppe 2
1	22	35
2	27	26
3	28	34
4	30	24
5	23	27
6	25	25
7	26	28
8	29	24
9	32	25
10	25	30
11	23	34
12	29	
13	29	
14	28	
15	30	
16	21	
17	26	
18	16	
19	23	
20	25	
Umfang	20	11
Mittelwert	25.85	28.36
Varianz	14.45	17.85

Auswertung mit Excel:

Daten			1. Stichprobe		2. Stichprobe
Stichprobenumfang		1)	20	2)	11
Mittelwert der Stichproben		3)	25.85	4)	28.36
Varianz		5)	14.45	6)	17.85
Differenz der Mittelwerte		8)	-2.5136	3) und 4)	
Standardfehler der Mittelwertsdifferenz		9)	0.359	1), 5) und 2), 6)	
Testgrösse aus den Daten			-6.9926	8), 9)	
(in den meisten Fällen ist ω = 0)					
			$\frac{s_1^2}{n_1}$	$\frac{s_2^2}{n_2}$	
		1), 5) und 2), 6	0.7225	1.6231	
Hilfsgrössen					
Zähler Freiheitsgrad			5.5020		
			0.0275	0.26346	
Nenner Freiheitsgrad			0.2909		
Freiheitsgrad			18.912		
signifikanzniveau			5%		Excel
kritisches t aus Tabelle		beidseitig	2.101	7), 12)	T.INV(0.975; 18)
Testgrösse aus den Daten			-6.99		
Da der Betrag der Testgrösse (6.99) grösser als der kritische Wert2.101) ist, wird man die Nullhypothese verwerfen.					
Die Gedächtnisleistung in den beiden Gruppen ist verschieden.					

$$\hat{\sigma}_{\bar{x}_1-\bar{x}_2} = S_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{S_{x_1}^2}{n_1} + \frac{S_{x_2}^2}{n_2}}$$

$$t = \frac{\bar{x}_1 - \bar{x}_2 - \omega_0}{\sqrt{\frac{s_{x_1}^2}{n_1} + \frac{s_{x_2}^2}{n_2}}} \sim t(df)$$

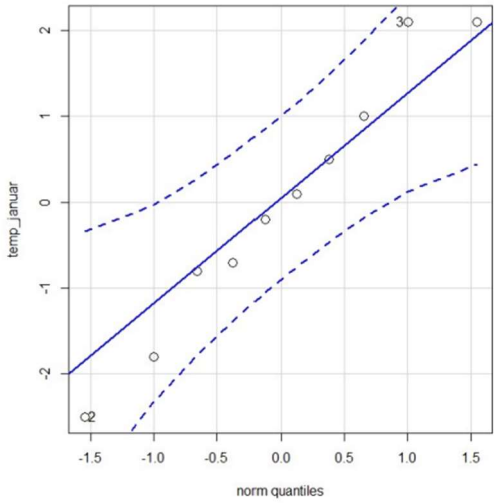
$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{s_1^4}{n_1^2(n_1-1)} + \frac{s_2^4}{n_2^2(n_2-1)}}$$

Voraussetzungen beim t-Test:

Der t-Test setzt eine Normalverteilung der Daten in den untersuchten Populationen (in der Grundgesamtheit) voraus.

Eine Prüfung kann mit dem Quantil-Quantil-Diagramm erfolgen **(Q-Q-Plot).**
Auf der x-Achse sind die Quantile der Normalverteilung und auf der y-Achse die Daten der Stichprobe dargestellt. Bei einer perfekten Normalverteilung würden die Daten auf der blau dargestellten Winkelhalbierenden liegen.
Die Daten sollten zufällig um diese Gerade streuen und keine Muster aufweisen.
Mindestens 95% der Punkte sollten wie im folgenden Beispiel innerhalb der gestrichelten Kurven liegen:

Beispiel (HAfL):
Januartemperaturen von 1991bis 2000 in Grad
0.1 -2.5 2.1 2.1 -0.7 -0.2 -1.8 1.0 0.5 -0.8



Die Prüfung ob die Daten normalverteilt sind, dann auch rechnerisch mit dem **Shapiro-Wilk-test** erfolgen

Falls die Daten nicht normalverteilt sind, kann ein parameterfreier Test, wie etwa der → **Wilcoxon-Rangtest** angewendet werden.

Vergleich mehrerer Mittelwerte

Das Problem, die Gleichheit mehrerer Mittelwerte zu testen wird im Abschnitt Varianzanalyse behandelt. Falls die Annahme der Normalverteilung verletzt ist, kann der Test mit dem Test von Kruskal und Wallis erfolgen.

Die bisher erwähnten Tests sind eine Auswahl. Eine Uebersicht stellt etwa die folgende Zusammenfassung der «Methodenberatung Statistik der Universität Zürich» dar.

Weitere Quellen:
Stahel: Angewandte Statistik
Statistik verstehen youtube
Mittag Schüller: Statistik