

Einführung in statistische Tests

veröffentlicht im Internet unter aufgabomat.de

Gliederung:

1	Einführung	1
2	Ausgangspunkt: Intervallschätzung für den Erwartungswert einer normalverteilten Variablen	1
3	Übergang zum Parametertest für den Erwartungswert einer normalverteilten Zufallsvariable	2
4	Zusammenfassung	5

1 Einführung

Die Anzahl unterschiedlicher statistischer Tests ist so groß, dass es undenkbar ist, sie auch nur annähernd alle im Rahmen einer Lehrveranstaltung vorzustellen. Um eine aktuell zu behandelnde Fragestellung bearbeiten zu können, wird man sich in der Praxis daher häufig eigenständig einen statistischen Test suchen und sich mit ihm vertraut machen müssen. Dabei ist es äußerst hilfreich, wenn nicht sogar Voraussetzung, das grundlegende Konzept hinter statistischen Tests verstanden zu haben.

2 Ausgangspunkt: Intervallschätzung für den Erwartungswert einer normalverteilten Variablen

Ziel des vorliegenden Skripts ist es, das Konzept statistischer Tests ausgehend von demjenigen statistischen Verfahren zu entwickeln, das als die Intervallschätzung für den Erwartungswert einer normalverteilten Variable¹ bezeichnet wird. Dabei wird auf Grundlage einer Stichprobe von Werten einer normalverteilten Zufallsvariable X ein so genanntes Konfidenzintervall für deren Erwartungswert $E(X) = \mu$ ermittelt. Obere und untere Begrenzung des Intervalls werden bestimmt als

$$\mu = \bar{X} \pm t_{1-\alpha/2} \frac{S}{\sqrt{N}} \quad (1)$$

mit N : Anzahl der Stichprobenwerte

$\bar{X}$: empirischer Mittelwert der Stichprobenwerte

S : empirische Standardabweichung der Stichprobenwerte

$t_{1-\alpha/2}$: $(1 - \alpha/2)$ -Quantil der t -Verteilung mit dem Freiheitsgrad $f = N - 1$.

Beispiel: Der Durchmesser maschinell hergestellter Schrauben soll 12,00 mm betragen. Zur Kontrolle werden fünf Schrauben ausgemessen. Es ergeben sich die Werte 11,95 mm, 11,97 mm, 11,98 mm, 11,99 mm und 12,03 mm.

Empirischer Mittelwert und empirische Standardabweichung in dieser Stichprobe betragen $\bar{X} = 11,98$ mm und $s = 0,03$ mm. Wird als Konfidenzniveau $\gamma = 0,95$ festgelegt, entspricht dem

¹ Siehe Skript „Intervallschätzung für den Erwartungswert“ unter aufgabomat.de.

$\alpha = 1 - \gamma = 0,05$ und $1 - \alpha/2 = 0,975$. Das 0,975-Quantil der t-Verteilung mit dem Freiheitsgrad $f = 4$ ist $t_{0,975} = 2,78$ (Tabelle 1). Es ergibt sich das Konfidenzintervall [11,94 mm; 12,02 mm].

f	p						f	p					
	0,841	0,900	0,950	0,975	0,990	0,995		0,841	0,900	0,950	0,975	0,990	0,995
2	1,32	1,89	2,92	4,30	6,96	9,92	19	1,03	1,33	1,73	2,09	2,54	2,86
3	1,20	1,64	2,35	3,18	4,54	5,84	20	1,03	1,33	1,72	2,09	2,53	2,85
4	1,14	1,53	2,13	2,78	3,75	4,60	21	1,02	1,32	1,72	2,08	2,52	2,83
5	1,11	1,48	2,02	2,57	3,36	4,03	22	1,02	1,32	1,72	2,07	2,51	2,82
6	1,09	1,44	1,94	2,45	3,14	3,71	23	1,02	1,32	1,71	2,07	2,50	2,81
7	1,08	1,41	1,89	2,36	3,00	3,50	24	1,02	1,32	1,71	2,06	2,49	2,80
8	1,07	1,40	1,86	2,31	2,90	3,36	25	1,02	1,32	1,71	2,06	2,49	2,79
9	1,06	1,38	1,83	2,26	2,82	3,25	26	1,02	1,31	1,71	2,06	2,48	2,78
10	1,05	1,37	1,81	2,23	2,76	3,17	27	1,02	1,31	1,70	2,05	2,47	2,77
11	1,05	1,36	1,80	2,20	2,72	3,11	28	1,02	1,31	1,70	2,05	2,47	2,76
12	1,04	1,36	1,78	2,18	2,68	3,05	29	1,02	1,31	1,70	2,05	2,46	2,76
13	1,04	1,35	1,77	2,16	2,65	3,01	30	1,02	1,31	1,70	2,04	2,46	2,75
14	1,04	1,35	1,76	2,14	2,62	2,98	40	1,01	1,30	1,68	2,02	2,42	2,70
15	1,03	1,34	1,75	2,13	2,60	2,95	50	1,01	1,30	1,68	2,01	2,40	2,68
16	1,03	1,34	1,75	2,12	2,58	2,92	70	1,01	1,29	1,67	1,99	2,38	2,65
17	1,03	1,33	1,74	2,11	2,57	2,90	100	1,01	1,29	1,66	1,98	2,36	2,63
18	1,03	1,33	1,73	2,10	2,55	2,88	∞	1,00	1,28	1,64	1,96	2,33	2,58

Tabelle 1: Quantile t_p der t-Verteilung.

3 Übergang zum Parametertest für den Erwartungswert einer normalverteilten Zufallsvariable

Die Ausgangssituation für den im Folgenden eingeführten statistischen Test ist:

- Der Erwartungswert $E(X) = \mu$ einer normalverteilten Zufallsvariable X soll einem vorgegebenen Wert μ_0 entsprechen.
- Es liegt eine Stichprobe von N Werten der Variable vor.
- Der empirische Mittelwert $\bar{x}$ in der Stichprobe stimmt nicht mit dem Sollwert μ_0 überein.

Der empirische Mittelwert ist ein Schätzwert für den Erwartungswert. Da $\bar{x} \neq \mu_0$ ist, liegt der Verdacht nahe, dass die Forderung $E(X) = \mu_0$ nicht eingehalten wird. Dieser Verdacht ($E(X) \neq \mu_0$) wird zur so genannten **Alternativhypothese** des Tests gemacht. Jeder statistische Test ist mit der Angabe von Null- und Alternativhypothese einzuleiten. Anstatt diese beiden Begriffe auszuschreiben, ist es auch üblich, die abkürzenden Bezeichnungen H_0 und H_1 , manchmal auch H_A , zu verwenden (mit H für Hypothese). Die **Nullhypothese** ist das logische Gegenteil der Alternativhypothese, sodass sich im vorliegenden Fall ein statistischer Test mit den folgenden Hypothesen ergibt:

$$H_0: E(X) = \mu_0$$

$$H_1: E(X) \neq \mu_0.$$

Beispiel: Herstellung von Schrauben

Der empirische Mittelwert $\bar{x} = 11,98$ mm des Schraubendurchmessers in der Stichprobe weicht vom Sollwert $\mu_0 = 12,00$ mm ab. Kann dennoch davon ausgegangen werden, dass der Erwartungswert des Durchmessers in der Grundgesamtheit aller hergestellten Schrauben dem Sollwert entspricht?

$$H_0: E(X) = 12,00 \text{ mm}$$

$$H_1: E(X) \neq 12,00 \text{ mm.}$$

Das 0,95-Konfidenzintervall ist [11,94 mm; 12,02 mm]. Der Sollwert $\mu_0 = 12,00 \text{ mm}$ liegt innerhalb dieses Intervalls. Ist dies möglicherweise ein Hinweis darauf, dass die Nullhypothese gilt und nicht die Alternativhypothese?

Die Intervallschätzung liefert die folgende Aussage über den Erwartungswert μ :

$$P\left(\bar{X} - t_{1-\alpha/2} \frac{S}{\sqrt{N}} \leq \mu \leq \bar{X} + t_{1-\alpha/2} \frac{S}{\sqrt{N}}\right) = 1 - \alpha.$$

$1 - \alpha = \gamma$ ist das Konfidenzniveau der Intervallschätzung. Seine Vorgabe bestimmt die Breite des Intervalls. Im Zusammenhang mit statistischen Tests wird stattdessen üblicherweise $\alpha = 1 - \gamma$ angegeben, das **Signifikanzniveau** oder, wie man auch sagt, die **Irrtumswahrscheinlichkeit** des Tests.

Um zwischen der Null- und der Alternativhypothese eine Entscheidung zu treffen, wird angesetzt, dass H_0 gilt, falls μ_0 ebenfalls in diesem Konfidenzintervall liegt, d. h. falls

$$\bar{X} - t_{1-\alpha/2} \frac{S}{\sqrt{N}} \leq \mu_0 \leq \bar{X} + t_{1-\alpha/2} \frac{S}{\sqrt{N}}$$

ist. Dieser mathematische Ausdruck wird nun umgeformt, ohne dass an der darin liegenden mathematischen Aussage etwas verändert wird. Zuerst wird $\bar{X}$ subtrahiert.

$$- t_{1-\alpha/2} \frac{S}{\sqrt{N}} \leq \mu_0 - \bar{X} \leq t_{1-\alpha/2} \frac{S}{\sqrt{N}}$$

Als nächstes wird mit -1 multipliziert. Dabei ist zu beachten, dass sich die Kleinerzeichen umkehren.

$$t_{1-\alpha/2} \frac{S}{\sqrt{N}} \geq \bar{X} - \mu_0 \geq - t_{1-\alpha/2} \frac{S}{\sqrt{N}} \quad | : \frac{S}{\sqrt{N}}$$

$$t_{1-\alpha/2} \geq \frac{\bar{X} - \mu_0}{S / \sqrt{N}} \geq - t_{1-\alpha/2}$$

Dieser Ausdruck wird jetzt anstatt von links nach rechts von rechts nach links gelesen:

$$- t_{1-\alpha/2} \leq \frac{\bar{X} - \mu_0}{S / \sqrt{N}} \leq t_{1-\alpha/2}.$$

Aufgrund der speziellen Symmetrie der Wahrscheinlichkeitsdichtefunktion der t-Verteilung (Abbildung 1) folgt außerdem $- t_{1-\alpha/2} = t_{\alpha/2}$. Damit ergibt sich

$$t_{\alpha/2} \leq \frac{\bar{X} - \mu_0}{S / \sqrt{N}} \leq t_{1-\alpha/2}. \quad (2)$$

Falls diese Bedingung erfüllt ist, wird gemäß dem oben formulierten Ansatz angenommen, dass die Nullhypothese des Tests gilt.

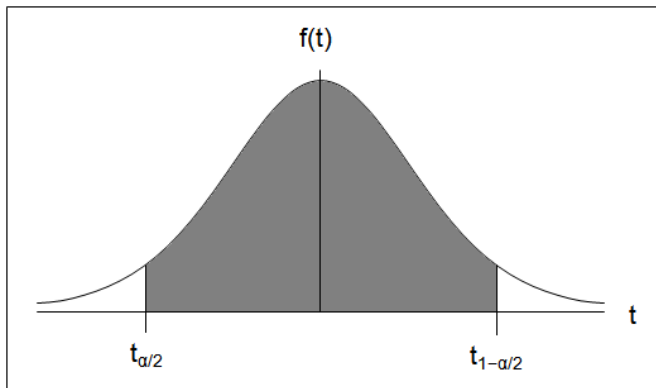


Abbildung 1: Wahrscheinlichkeitsdichtefunktion der t-Verteilung.

Damit ist der statistische Test bereits vollständig skizziert. Er wird als zweiseitiger **Parameter-test für den Erwartungswert einer normalverteilten Variable** oder als zweiseitiger **Einstichproben-t-Test** bezeichnet.

Für die Zufallsvariablen $\bar{X}$ und S ergibt sich aus der Stichprobe jeweils ein Wert $\bar{x}$ und s . Mit diesen lässt sich für die Zufallsvariable

$$T = \frac{\bar{X} - \mu}{S / \sqrt{N}}$$

genau ein Wert

$$t = \frac{\bar{x} - \mu_0}{s / \sqrt{N}} \quad (3)$$

berechnen. Die Zufallsvariable T wird als die **Teststatistik** bezeichnet. Sie ist charakteristisch für den betreffenden Test und wird so definiert, dass sie bei Gültigkeit der Nullhypothese des Tests bekannte Eigenschaften aufweist. Im vorliegenden Fall ist die Teststatistik bei Gültigkeit von H_0 t-verteilt mit dem Freiheitsgrad $f = N - 1$.

Ihr Wert t , der sich aus der vorliegenden Stichprobe ergibt, wird **Wert der Teststatistik** oder, kurz, **Prüfwert** genannt.

Das Intervall $[t_{\alpha/2}; t_{1-\alpha/2}]$ ist das so genannte **Annahmeintervall**. Allgemein sind die Begrenzungen des Annahmeintervalls Quantile derjenigen Wahrscheinlichkeitsverteilung, durch welche die Teststatistik bei Gültigkeit von H_0 beschrieben wird, hier der t-Verteilung.

Falls der Prüfwert im Annahmeintervall liegt, wird angenommen, dass die Nullhypothese H_0 gilt. Im Folgenden wird als Abschluss des Tests formuliert „ H_0 wird beibehalten.“. Falls der Prüfwert außerhalb des Annahmeintervalls liegt, wird angenommen, dass die Alternativhypothese H_1 gilt. Im Folgenden wird als Abschluss des Tests formuliert „ H_0 wird verworfen.“.

Beispiel: Herstellung von Schrauben

Signifikanzniveau: $\alpha = 0,05$

Der Wert der Teststatistik ergibt sich mit $N = 5$, $\bar{x} = 11,98$ mm und $s = 0,03$ mm gemäß Gleichung 3 zu $t = -1,49$.

Teststatistik ist die t-Verteilung mit dem Freiheitsgrad $f = 4$. Das $(1 - \alpha/2)$ -Quantil der t-Verteilung mit $f = 4$ ist $t_{0,975} = 2,78$. Das $\alpha/2$ -Quantil der t-Verteilung mit $f = 4$ ist $t_{0,025} = -t_{0,975} = -2,78$. Das Annahmeintervall ist folglich $[-2,78; 2,78]$.

Aus $t_{0,025} \leq t \leq t_{0,975}$ folgt: Die Nullhypothese wird beibehalten.

4 Zusammenfassung

Ein statistischer Test läuft grundsätzlich in den folgenden Schritten ab:

1. Nullhypothese H_0 und Alternativhypothese H_1 formulieren
2. Irrtumswahrscheinlichkeit bzw. Signifikanzniveau α festlegen
3. Wert der Teststatistik (Prüfwert) berechnen (Tabelle 2)

Test	Voraussetzungen	Prüfwert	Teststatistik bei Gültigkeit von H_0
Parametertest für den Erwartungswert	X ist normalverteilt.	$t = \frac{\bar{x} - \mu_0}{s / \sqrt{N}}$	t-verteilt mit $f = N - 1$
Parametertest für die Varianz	X ist normalverteilt.	$\chi^2 = (N - 1) \frac{s^2}{\sigma_0^2}$	Chi-Quadrat-verteilt mit $f = N - 1$
F-Test	X_1 und X_2 sind normalverteilt und unabhängig voneinander.	$F = \frac{s_1^2}{s_2^2}$	F-verteilt mit $f_1 = N_1 - 1, f_2 = N_2 - 1$
Zweistichproben-t-Test	X_1 und X_2 sind normalverteilt, unabhängig voneinander und besitzen dieselbe Varianz.	$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}}$	t-verteilt mit $f = N_1 + N_2 - 2$

Tabelle 2: Zusammenfassung von Informationen zu einigen statistischen Tests.

4. Annahmeintervall bestimmen

Bei einem zweiseitigen Test wird das Annahmeintervall durch das $\alpha/2$ -Quantil und das $(1 - \alpha/2)$ -Quantil derjenigen Wahrscheinlichkeitsverteilung begrenzt, durch welche die Teststatistik bei Gültigkeit von H_0 beschrieben wird. Bei einem linksseitigen Test ist das Annahmeintervall nur nach links durch das α -Quantil und bei einem rechtsseitigen Test nur nach rechts durch das $(1 - \alpha)$ -Quantil dieser Wahrscheinlichkeitsverteilung begrenzt.

5. Testentscheid treffen

Prüfwert innerhalb des Annahmeintervalls $\Rightarrow H_0$ wird beibehalten.

Prüfwert außerhalb des Annahmeintervalls $\Rightarrow H_0$ wird verworfen.