# Der Signifikanztest

Carlo Vöst, Oliva, Spanien
Illustrationen von Carlo Vöst

*Foto: Klaus Vedfelt/DigitalVision/Getty Images Plus*

Ein statistischer Test erlaubt es den Statistikern in vielen Bereichen, in denen die Wahrscheinlichkeit für das Auftreten eines bestimmten Ereignisses nicht bekannt ist, sich auf der Basis einer empirischen Stichprobe zwischen zwei konkurrierenden wissenschaftlichen Hypothesen zu entscheiden. Dabei treten gewisse Irrtumswahrscheinlichkeiten auf, die aufgrund der getroffenen Vorgaben von besonderem Interesse sind. In diesem Beitrag lernen die Jugendlichen die verschiedenen Signifikanztests in Theorie und Praxis kennen. An zahlreichen Aufgaben wenden sie ihr neues Wissen an und testen sich in einer Lernerfolgskontrolle.

# Der Signifikanztest

## Oberstufe (grundlegend)

Carlo Vöst, Oliva, Spanien
Illustrationen von Carlo Vöst





## Die Schüler lernen:

die verschiedenen Arten von Signifikanztests zunächst von ihrer theoretischen Seite, anhand von Beispielen illustriert, kennen. Dies sind der rechtsseitige, der linksseitige und der zweiseitige Signifikanztest. Ihre Schüler üben anschließend Anwendungen dieser Testverfahren anhand einer Reihe von Beispielen ein. Eine Lernerfolgskontrolle rundet den Beitrag ab.

## Überblick:

Legende der Abkürzungen:
**Ab** = Arbeitsblatt **LEK** = Lernerfolgskontrolle

| Thema | Material | Methode |
|---|---|---|
| Der Signifikanztest (Theorieteil) | M1 | Ab |
| Aufgaben | M2 | Ab |
| Klassenarbeit | M3 | LEK |

**Erklärung zu Differenzierungssymbolen**

| | | |
|---|---|---|
| einfaches Niveau | mittleres Niveau | schwieriges Niveau |
| | Dieses Symbol markiert Zusatzaufgaben. | |

**Kompetenzprofil:**

**Inhalt:** rechtsseitiger, linksseitiger, beidseitiger Signifikanztest, Übungsaufgaben, Lernerfolgskontrolle

**Medien:** GTR/CAS, Taschenrechner, stochastisches Tafelwerk

**Kompetenzen:** Mathematisch argumentieren und beweisen (K1), Probleme mathematisch lösen (K2), mathematisch modellieren (K3), mathematische Darstellungen verwenden (K4), mit symbolischen, formalen und technischen Elementen der Mathematik umgehen (K5), Kommunizieren (K6)

## Hinweise

Dieser Beitrag eignet sich besonders gut zum Einstieg in das Thema „Signifikanztests". Sie können ihn aber auch den Jugendlichen zum Selbststudium überlassen, damit sie sich eigenständig erste Kenntnisse erarbeiten. Der Beitrag bzw. die Aufgaben erfordern öfter auch den Einsatz eines CAS-Rechners, welcher gerade hier ein mächtiges Hilfsmittel ist, um sich langwierige Arbeit mit stochastischen Tafelwerken und Taschenrechner zu ersparen. Allerdings sind die Beispiele so gewählt, dass die Lernenden sie auch ohne CAS-Rechner (allerdings mit mehr Aufwand) bewältigen können.

### Lehrplanbezug

Am Beispiel des einseitigen Signifikanztests erhalten die Schüler einen Einblick in die beurteilende Statistik. Sie lernen einzuschätzen, wie sich Änderungen von Stichprobenlänge, Ablehnungsbereich oder Signifikanzniveau auf die Aussage des Tests auswirken.[1]



### Zu den einzelnen Materialien

Im Material **M 1** ist die Theorie zum Verstehen des Signifikanztests zusammengestellt. Damit Sie mit Ihrer Klasse diesen theoretischen Teil besser erarbeiten können, greift dieser immer wieder auf anschauliche Beispiele zurück.

Das Material **M 2** bietet Ihnen insgesamt zehn ausführlich formulierte Aufgaben, mit denen die Lernenden die Möglichkeit haben, ihr erworbenes Wissen einzuüben.

Material **M 3** ist eine mögliche Lernerfolgskontrolle zu diesem Themenkomplex mit Bewertungseinheiten und Bewertungsschlüssel. So können entweder Ihre Schülerinnen und Schüler den „Ernstfall" simulieren, oder Sie, als Lehrer, bekommen eine Anregung, wie Sie eine Klassenarbeit gestalten können.

Schließlich sind ausführliche Lösungen zu den Aufgaben und der Lernerfolgskontrolle angefügt. Gerade in Zeiten des Home-Schooling eignet sich der Beitrag damit auch zum Selbststudium.



[1] Lehrplan Bayern, http://www.gym8-lehrplan.bayern.de/contentserv/3.1.neu/g8.de/id_26192.html (aufgerufen am 5.02.2021)

## M 1 Der Signifikanztest (Theorieteil)

In der Wahrscheinlichkeitsrechnung ist meistens die Trefferwahrscheinlichkeit p bekannt (oder berechenbar) und man berechnet damit die Wahrscheinlichkeiten von Ereignissen. In der Statistik ist oft die „umgekehrte Aufgabe" zu lösen: Der Anteil p der „Treffer" in einer Gesamtheit ist unbekannt. Man entnimmt der Gesamtheit eine Stichprobe und folgert daraus p. Dieses Verfahren wird beispielsweise bei Hochrechnungen und Prognosen von Wahlen benutzt. Aufgrund der Umfrageergebnisse schätzt man p ab. Häufig hat man aber von vornherein über den Trefferanteil p gewisse Hypothesen (Vermutungen). Aufgrund des Ergebnisses einer Stichprobe entscheidet man sich dann, ob man die Hypothese verwerfen muss oder nicht. Wenn man testen will, ob man sich für eine ganz bestimmte Hypothese entscheiden will oder nicht, spricht man von einem Signifikanztest.

**Beispiel:**

Bei einer Urne mit 10 Kugeln ist bekannt, dass sich darin schwarze und weiße Kugeln befinden, die sich nur durch ihre Farbe unterscheiden.

Es soll nun durch ein 5-maliges Ziehen mit Zurücklegen (ZL) getestet werden, ob man sich für die Hypothese entscheiden will, dass sich in der Urne gleich viele schwarze und weiße Kugeln befinden, oder nicht.

Wenn p die relative Häufigkeit der weißen Kugeln angibt, dann muss man sich zwischen der Annahme oder Ablehnung der sogenannten Nullhypothese $H_0: p = 0{,}5$ (dass sich also in der Urne gleich viele schwarze und weiße Kugeln befinden) entscheiden.

**Hier gilt:** $p \in \{0;\ 0{,}1;\ 0{,}2;\ 0{,}3;\ 0{,}4;\ 0{,}5;\ 0{,}6;\ 0{,}7;\ 0{,}8;\ 0{,}9;\ 1\}$, weil prinzipiell 0–10 von 10 Kugeln weiß sein können.

Zufallsvariable X := „Anzahl der gezogenen weißen Kugeln" (diese Wahl ist beliebig)

**Mögliche Entscheidungsregel:**

$X \in \{2, 3\} \rightarrow H_0$ wird angenommen $\rightarrow$ man sagt dann, dass X im sog. Annahmebereich $A = \{2, 3\}$ für $H_0$ liegt.

$X \in (\{0;1\} \cup \{4;5\}) \rightarrow H_0$ wird abgelehnt $\rightarrow$ X liegt im sog. Ablehnungsbereich $\overline{A} = \{0;1\} \cup \{4;5\}$ für $H_0$.

Folgende Fälle sind möglich:

| | $H_0$ wird abgelehnt | $H_0$ wird angenommen |
|---|---|---|
| $H_0$ ist wahr | Fehlentscheidung | richtige Entscheidung |
| $H_0$ ist falsch | richtige Entscheidung | Fehlentscheidung |

Folgende Bezeichnungen sind nun in der Statistik üblich:

**Fehler erster Art:** die Nullhypothese wird zu Unrecht abgelehnt

**Risiko erster Art:** Wahrscheinlichkeit für den Fehler erster Art

**Fehler zweiter Art:** die Nullhypothese wird zu Unrecht angenommen

**Risiko zweiter Art:** Wahrscheinlichkeit für den Fehler zweiter Art

**Überblick:**

| | Entscheidung | Entscheidung ist | Fehler |
|---|---|---|---|
| $H_0$ ist wahr | Ablehnung | falsch | erster Art |
| | Annahme | richtig | ------------ |
| $H_0$ ist falsch | Ablehnung | richtig | ------------ |
| | Annahme | falsch | zweiter Art |

Berechnung des Risikos erster Art in unserem Beispiel:

$$P\left(X \in \left(\{0;1\} \cup \{4;5\}\right)\right) = 1 - P\left(X \in \{2;3\}\right) = 1 - \sum_{i=2}^{3} B(5;0{,}5;i) = 0{,}375\,,$$

das heißt: in 37,5 % aller Fälle nimmt man zu Unrecht an, dass sich von einer Sorte mehr Kugeln als von der anderen Sorte in der Urne befinden.

**Anmerkung:**

Grundsätzlich ist die Festlegung des Annahmebereichs beliebig. Es gehört nur zu jedem Wert von k (Grenze des Annahmebereichs) dann ein bestimmtes Risiko 1. Art für die irrtümliche Ablehnung einer an sich wahren Hypothese $H_0$ und ein bestimmtes Risiko 2. Art für die irrtümliche Annahme einer an sich falschen Hypothese.

Zu beachten ist, dass es in praktischen Beispielen oft sinnvoll ist, bei der Festlegung der

Entscheidungsregel für ein möglichst geringes Risiko 1. Art zu sorgen. Man gibt sich sogar oft ganz bewusst eine Zahl $\alpha$ mit $0<\alpha<1$ als obere Schranke für das Risiko 1. Art vor. $\alpha$ nennt man Signifikanzniveau. Dann ist der Ablehnungsbereich A bestimmt durch die Bedingung: $P(X\in\overline{A})\leq\alpha$, falls $H_0$ wahr ist.

Die Zahl $\beta=1-\alpha$ stellt dann eine untere Schranke für die Wahrscheinlichkeit, eine wahre Hypothese zu Recht nicht abzulehnen, dar, also ein Maß für die Sicherheit der Entscheidung ($\beta$ heißt Sicherheits-Niveau). Also gilt die Bedingung:

$P(X\in A)\geq\beta$, falls $H_0$ wahr ist.

Wenn die Durchführung der Stichprobe nach Anwendung der durch das Signifikanzniveau $\alpha$ festgelegten Entscheidungsregel zur Ablehnung der Nullhypothese führt, dann sagt man, dass das Testergebnis signifikant auf dem Niveau $\alpha$ ist (übliche Sprechweise).

Den Ablehnungsbereich $\overline{A}$ nennt man auch kritischen Bereich, weil es durch ihn zur Ablehnung einer eigentlich wahren Hypothese kommen könnte.

## A) Der rechtsseitige Signifikanztest

**Beispiel:**

Wir wollen Transistoren kaufen.

Der Hersteller behauptet, die Ausschussquote sei höchstens 10 %.

$\Rightarrow$ Hypothese $H_0: p=0{,}1$ für unsere Überlegung ergibt sich die Gegenhypothese: $p>0{,}1$ (rechtsseitig). Die Gegenhypothese wird oft auch mit $H_1$ bezeichnet; hier gilt also: $H_1: p>0{,}1$.

**Folgendes Verfahren wird vereinbart:**

Als Erstes wird die Art der Stichprobe festgelegt:

Aus den Transistoren werden (zum Beispiel) willkürlich 20 entnommen, ihre Funktionstüchtigkeit überprüft und zurückgelegt. Die Anzahl der defekten Transistoren (Testgröße) wird mit T (Zufallsgröße) bezeichnet.

Als Zweites wird eine (vernünftige) Entscheidungsregel formuliert:

Ist $T > 4 \rightarrow$ Entscheidung gegen $H_0$

Ist $T \leq 4 \rightarrow$ Entscheidung für $H_0$

Der Annahmebereich A für $H_0$ ist dann das Ereignis $A = \{0;\ 1;\ 2;\ 3;\ 4\}$,

Der Ablehnungsbereich $\overline{A}$ für $H_0$ ist dann das Ereignis $\overline{A} = \{5;\ 6; \ldots;\ 20\}$.

Dadurch sind zwei **Fehlentscheidungen** möglich:

| | **$H_0$ wird abgelehnt** | **$H_0$ wird angenommen** |
|---|---|---|
| $H_0$ ist wahr | Fehlentscheidung:<br>Fehler 1. Art:<br>$p \leq 0{,}1$ jedoch $T > 4$<br>Wahrscheinlichkeit hierfür<br>(Risiko 1. Art):<br>$P_{H_0}\left(\overline{A}\right) = \sum_{i=5}^{20} B(20; 0{,}1; i)$<br>$\approx 0{,}04317$<br>(ist ein Maximalwert für<br>p = 0,1; für kleinere p kleiner) | richtige Entscheidung |
| $H_0$ ist falsch | richtige Entscheidung | Fehlentscheidung: Fehler 2. Art:<br>$p > 0{,}1$ jedoch $T \leq 4$<br>Wahrscheinlichkeit hierfür<br>(Risiko 2. Art):<br>$P_{\overline{H_0}}(A)$<br>(Dieser Wert ist für jeden Wert p > 0,1<br>unterschiedlich; das größte Risiko ist in<br>der Nähe des Nullhypothesenwerts!) |

**Bemerkungen:**

Für das Risiko 2. Art gilt beispielhaft:

p = 0,4: $\sum_{i=0}^{4} B(20; 0{,}4; i) \approx 0{,}05095$

p = 0,2: $\sum_{i=0}^{4} B(20; 0{,}2; i) \approx 0{,}62965$

p = 0,15: $\sum_{i=0}^{4} B(20; 0{,}15; i) \approx 0{,}82985$

p = 0,11: $\sum_{i=0}^{4} B(20; 0{,}11; i) \approx 0{,}93898$

Sie sehen, dass das Risiko 2. Art immer größer wird, je näher der Wahrscheinlichkeitswert an dem Wert von p für die Nullhypothese liegt.

Oft ist die irrtümliche Verwerfung der Nullhypothese verhängnisvoll. Für das Risiko für einen Fehler 1. Art wird dann ein bestimmter Schwellenwert (Signifikanzniveau $\alpha$) vorgegeben, der nicht überschritten werden darf. Üblicherweise liegt dieser bei höchstens 5 %.

**Festlegung:**

Ist $\alpha \le 0{,}05 = 5\ \%$, so heißt der Test signifikant.

Ist $\alpha \le 0{,}01 = 1\ \%$, so heißt der Test hochsignifikant.

Ist $\alpha \le 0{,}001 = 0{,}1\ \%$, so heißt der Test höchstsignifikant.

### B) Der linksseitige Signifikanztest

**Beispiel:**

Laut Angabe des Herstellers keimen mindestens 90 % aller Samenkörner seines Saatguts. Ein Gärtner sät 100 Samenkörner aus und beobachtet deren Keimverhalten.

Wie muss der Ablehnungsbereich $\overline{A}$ für die Behauptung des Herstellers gewählt werden, wenn man sich höchstens mit einer Wahrscheinlichkeit von 5 % irren will?

$\rightarrow$ Hypothese $H_0: p = 0{,}9$; Gegenhypothese: $p < 0{,}9$ (linksseitig)

**Folgendes Verfahren wird vereinbart:**

Zuerst die Stichprobe: 100 Samenkörner (n = 100);

Testgröße T ist die Anzahl der keimenden Samenkörner.

Dann die Entscheidungsregel $P_{H_0}\left(\overline{A}\right) = \sum_{k=0}^{N} B(100; 0{,}9; k) \le 0{,}05 \quad \Rightarrow \quad N = 0,1,\ldots,84:$

Ist $T \ge 85 \rightarrow$ Entscheidung für $H_0$

Ist $T < 85 \rightarrow$ Entscheidung gegen $H_0$

Der Annahmebereich A für $H_0$ ist dann das Ereignis $A = \{85;\ 86; \ldots;\ 100\}$

Der Ablehnungsbereich $\overline{A}$ für $H_0$ ist dann das Ereignis $\overline{A} = \{0;\ 1; \ldots;\ 83;\ 84\}$

**Anmerkung:** Die Festlegung der Stichprobe ist im Prinzip willkürlich. Allerdings wird die Güte des Tests umso besser, je größer die Stichprobenlänge ist (zu groß ist aus praktischen Gründen auch wieder schlecht) und je näher der Grenzwert der Entscheidungsregel der relativen Häufigkeit der Nullhypothese entspricht (hier: 0,85).

Es sind hier wieder zwei Fehlentscheidungen möglich:

| | $H_0$ wird abgelehnt | $H_0$ wird angenommen |
|---|---|---|
| $H_0$ ist wahr | Fehlentscheidung: Fehler 1. Art: $p \geq 0{,}9$ jedoch $T < 85$ Die Wahrscheinlichkeit hierfür (also das Risiko 1. Art): $P_{H_0}(\overline{A}) = \sum_{k=0}^{84} B(100; 0{,}9; k) \approx 0{,}03989$ (ist ein Maximalwert für p = 0,9; für größere p kleiner.) | richtige Entscheidung |
| $H_0$ ist falsch | richtige Entscheidung | Fehlentscheidung: Fehler 2. Art: $p < 0{,}9$ jedoch $T > 84$ Die Wahrscheinlichkeit hierfür (also das Risiko 2. Art): $P_{\overline{H_0}}(A)$ (Der Wert hierfür ist für jeden Wert p < 0,9 unterschiedlich; das größte Risiko ist in der Nähe des Nullhypothesenwerts.) |

**Zum Risiko 2. Art:**

für $p = 0{,}8$: $\sum_{i=85}^{100} B(100; 0{,}8; i) \approx 0{,}12851$

für $p = 0{,}85$: $\sum_{i=85}^{100} B(100; 0{,}85; i) \approx 0{,}56832$

für $p = 0{,}88$: $\sum_{i=85}^{100} B(100; 0{,}88; i) \approx 0{,}85855$

für $p = 0{,}89$: $\sum_{i=85}^{100} B(100; 0{,}89; i) \approx 0{,}91977$

## C) Der zweiseitige Signifikanztest

**Beispiel:**

Jemand vermutet, dass der bei einem Spiel verwendete Würfel kein Laplace-Würfel (L-Würfel) ist. Um dieser Vermutung nachzugehen, würfelt er 100-mal und bestimmt die Anzahl der Sechser. Er wählt folgende Entscheidungsregel:

Sind es mehr als 10 Sechser, aber höchstens 24, so erkennt er den Würfel als L-Würfel an.

a) Mit welcher Wahrscheinlichkeit verwirft er einen L-Würfel irrtümlich als solchen?

b) Die Sechs fällt mit einer Wahrscheinlichkeit von 20 %. Mit welcher Wahrscheinlichkeit hält er bei obiger Entscheidungsregel einen solchen Würfel irrtümlich für einen L-Würfel?

Die Nullhypothese $H_0$ ist also hier: $p = \frac{1}{6}$; Gegenhypothese: $p \neq \frac{1}{6}$ (zweiseitig)

**Er wählt folgendes Verfahren:**

Stichprobe: 100 Würfe (n = 100);

Testgröße T ist die Anzahl gewürfelten Sechser.

Er legt sich für folgende (naheliegende) Entscheidungsregel fest:

Ist $11 \leq T \leq 24$, dann entscheidet er sich für $H_0$,

ist $T < 11 \vee T > 24$, dann entscheidet er sich gegen $H_0$.

Der Annahmebereich A für $H_0$ ist dann das Ereignis $A = \{11;\ 12; \ldots;\ 23;\ 24\}$

Entsprechend ist der Ablehnungsbereich $\overline{A}$ für $H_0$ dann das Ereignis

$\overline{A} = \{0;\ 1; \ldots;\ 10\} \cup \{25; \ldots;\ 100\}$.