Signifikanztest

Die Situation eines Alternativtests, sich zwischen zwei einfachen Hypothesen entscheiden zu müssen, kommt in der Praxis selten vor, weil die Welt um uns dafür zu kompliziert ist.
Sehr häufig stellt sich einem jedoch das folgende **Problem:** Aufgrund irgendwelcher Erfahrungen und Überlegungen hegt man *eine* Vermutung, die nun durch einen Test, den sogenannten Signifikanztest, entweder bestätigt oder widerlegt werden soll.
Wird die Nullhypothese abgelehnt, spricht man von einem **signifikanten** Unterschied zwischen der in der Nullhypothese angenommenen und der tatsächlich vorliegenden Verteilung**. ABER:** Bei Ablehnung von H_0 ist damit aber keinesfalls „bewiesen“, dass H_0 falsch ist. Es wird aber mit einer Irrtumswahrscheinlichkeit α festgestellt, dass der beobachtete Wert von X mit der Nullhypothese unverträglich ist. Ein Verfahren, was auf die Widerlegung der Nullhypothese abzielt, heißt **Signifikanztest.**
Der Ansatz der Testtheorie (nach Neymann und Pearson) ähnelt in seiner Ausführung dem Ablauf eines indirekten Beweises der Aussagenlogik
Um eine Hypothese nicht zu verwerfen, untersucht man, ob die gegenteilige Annahme (= nicht gewünschte Hypothese = Nullhypothese H_0) mit dem Stichprobenergebnis unverträglich ist.
Man untersucht also, ob das Versuchsergebnis unter der Annahme der Nullhypothese mit einer sehr geringen Wahrscheinlichkeit eintritt. Als Nullhypothese H_0 wählt man immer die Hypothese, die man verwerfen will.
Die Wahrscheinlichkeit für den Fehler 1.Art – meistens 5% oder 1% - gibt man deshalb vor, weil die Nullhypothese (d.h. die Annahme einer falschen „Wunschhypothese“) als schwerwiegender Fehler angesehen werden kann. Durch die Vorgabe des Signifikanzniveaus α hat man eine Kontrollmöglichkeit mit einer quantitativen Abschätzung des Risikos einer entsprechenden Fehlentscheidung.
Je kleiner man α wählt , umso vorsichtiger ist man bei der Ablehnung der Nullhypothese.
Wenn selbst bei kleinem Wert α eine Ablehnung von H_0 erfolgt, spricht man von hoher Signifikanz und umgekehrt.
Bei der Beurteilung von Produkten geht man in der Praxis im Allgemeinen von einer statistischen Sicherheit (1 - α) von 95% ($\alpha \leq 0{,}05$) aus, es sei denn die Entscheidung Ist äußerst „kapitalintensiv“ etwa in Form hoher Investitionen,; da verwendet man eine Sicherheit von 99% ($\alpha \leq 0{,}01$)

Allgemein gilt:

Ein Signifikanztest läuft (fast) immer in folgenden Schritten ab:

1. Wie lautet die Nullhypothese H_0 (häufig die Verneinung der Vermutung)?
2. Wie groß ist der Stichprobenumfang des Tests, und welche Irrtumswahrscheinlichkeit (Signifikanzniveau α) ist vorgegeben?
3. Welche Größe X wird zur Prüfung verwendet (Sachsen nur Binomialverteilung), und wie lautet der Ablehnungsbereich $\overline{A}$?
4. Wie wird aufgrund des Stichprobenergebnisses entschieden?

Beispiele für Signifikanztests

1.Zweiseitiger Signifikanztest

Aufgabe 1: (Zweiseitiger Test (p = 0,5)

In der Berliner Münze soll die Vermutung getestet werden, ob eine neue Prägemaschine Münzen mit unausgeglichener Gewichtsverteilung herstellt, so genannte unfaire Münzen. Zu diesem Zweck wird eine der produzierten Münzen 100-mal geworfen und die Anzahl der Kopfwürfe werden gezählt. Weicht das Zählergebnis wenigstens um 10 vom Erwartungswert ab, so wird die Münze als unfair eingestuft. Welches Signifikanzniveau ergibt sich?

Lösung: n = 100; p: Wahrscheinlichkeit für Kopf; X: Anzahl der Kopfwürfe bei 100 Würfen
H_0: Die Münze ist fair: $p_0 = 0,5$
H_1: Die Münze ist unfair: $p_1 \neq 0,5 \rightarrow \overline{A}\{0,.....40\} \cup \{60,......100\} \rightarrow A\{41,.....59\}$
α = Fehler 1.Art $= P_{H_0}(\text{Entscheidung für } H_1)$
$= P(X \leq 40) + P(X \geq 60) = 1 - P(41 \leq X \leq 59) = 0,0568$

Zusatz: a) Welches Signifikanzniveau ergibt sich im obigen Beispiel für n = 80?
b) Wie groß ist der β-Fehler bei p_1= 0,4 bzw. p_1^* = 0,7
Wie kann durch Veränderung der im obigen Beispiel verwendeten Entscheidungsrege der Fehler 1.Art auf maximal 1% gedrückt werden?

Aufgabe 2: (Zweiseitiger Test p $\neq 0,5$)

Der Abgeordnete Egon Olsen hat bei der letzten Wahl 40% der Stimmen erhalten. Er möchte nun wissen, ob sich dieser Stimmenanteil inzwischen verändert hat; also lässt er 100 Personen aus seinem Wahlkreis befragen,
Sollten dabei erheblich weniger oder erheblich mehr als 40 Personen für ihn votieren, so wird er annehmen, dass sich der Stimmanteil verändert hat. Er möchte das Risiko, dass er aus dem Ergebnis der Umfrage irrtümlich aus einen veränderten Stimmanteil schließt, auf maximal 20% begrenzen. Welche Entscheidungsregel sollte er bei Auswertung des Umfrageergebnisses befolgen?

Lösung: n = 100; X: Anzahl der befragten Personen, die für Herrn Olsen stimmen würden
p: Anteil der Wahlberechtigten, die für Herrn Olsen stimmen würden
H_0: Stimmanteil unverändert: $p_0 = 0,4$; H_1: Stimmanteil verändert: $p_1 \neq 0,4$
Der kritische Bereich (Ablehnungsbereich von H_0) setzt sich aus zwei Intervallen zusammen: $[0;k_1]$ und $[k_2;100]$. Wir wählen die kritischen Zahlen g_L und g_R derart, dass gilt:

1) $\mathbf{P(X \leq k_1) \leq \frac{\alpha}{2}}$: $P(X \leq k_1) \leq 0,1 \rightarrow B(100; 0,4; X \leq k_1) \leq 0,1 \rightarrow$ gilt für $k_1 \leq 33$

2) $\mathbf{P(X \geq k_2) \leq \frac{\alpha}{2}}$: $P(X \geq k_2) \leq 0,1 \rightarrow 1 - B(100; 0,4; X \leq k_2-1) \leq 0,1 \rightarrow$

$0,9 \leq B(100, 0,4, X \leq k_2-1) \rightarrow$

Anmerkungen:

1. Wenn $\overline{A}$ gegeben ist, kann α bestimmt werden und umgekehrt.

2. Beim zweiseitigen Test d.h. wenn $\overline{A} = \{0,...,k_1\} \cup \{k_2,......n\}$ wird die Fehlerwahrscheinlichkeit α gleichmäßig aus beide Bereiche verteilt $\left(\frac{\alpha}{2}\right)$!!

Ein zweiseitiger Test liegt immer dann vor, wenn die Vermutung keine Richtung besitzt d.h. H_0 die Form H_0: $p = p_0$ ($p \neq p_0$) bzw. H_0: $p_1 < p_0 < p_2$ besitzt

3. Bei einem einseitigen Test gilt:
$H_0 : p \leq p_0$ ($p<p_0$) wird abgelehnt, wenn die Prüfgröße X sehr große Werte annimmt, d.h.
$\overline{A} = \{k,.....n\}$ → **rechtsseitiger Test**
$H_0 : p \geq p_0$ ($p > p_0$) wird abgelehnt, wenn die Prüfgröße X sehr kleine Werte annimmt, d.h.
$\overline{A} = \{0,.....k\}$ → **linksseitiger Test** gilt.

4. Ob einseitig oder zweiseitig getestet wird, muss man sich aufgrund der Problemstellung genau überlegen, d.h. ob irgendwelche Abweichungen nach oben und nach unten oder nur nach einer Seite von Interesse sind.
(Beispielsweise wird beim Ankauf einer Ware mit einer bestimmten Fehlerwahrscheinlichkeit kaum interessieren, ob der Anteil der fehlerhaften Ware in der Lieferung kleiner als vereinbart ist.)

2. Einseitiger Signifikanztest

Aufgabe1:(Bestimmung der Irrtumswahrscheinlichkeit)
Ein Pharma - Hersteller hat ein neues Medikament gegen Schlafstörung entwickelt.
Das beste bereits auf dem Markt eingeführte Medikament mit vergleichbar geringen Nebenwirkungen zeigt 50% der Anwendungsfälle eine ausreichende Wirkung.
Erste Anwendungen lassen die Forscher die Hypothese aufstellen, dass das neue Medikament in einem noch größeren Anteil der Anwendungsfälle ausreichend wirkt.
Dies soll in einer Studie an 50 Patienten überprüft werden. Die Forscher sind vorsichtig und legen fest, dass die Hypothese nur angenommen wird, wenn das Medikament bei mehr als 30 Patienten ausreichend wirkt.
Mit welcher Wahrscheinlichkeit wird dem Medikament eine bessere Wirkung als dem alten Medikament zugesprochen, wenn dieser Sachverhalt in Wirklichkeit gar nicht zutrifft?

Aufgabe 2:
(Vorgabe des Signifikanzniveaus, einseitiger Test)
Der Pharma – Hersteller möchte nun in einer 50 Patienten umfassenden Studie testen, ob sein neues Schlafmittel wirklich – wie die seine Forscher vermuten – besser ist als die besten marktgängigen Produkte, die in 50% der Fälle helfen.
Er möchte dieser Vermutung Glauben schenken, wenn das Medikament bei mehr als k Patienten wirkt.
Er ist recht vorsichtig und verlangt daher, dass die kritische Zahl k so bestimmt werden soll, dass der Test ein 1% - Signifikanzniveau besitzt, d.h. die Wahrscheinlichkeit dafür, dass das neue Medikament zu Unrecht als den alten Medikamenten überlegen eingestuft wird, darf maximal 1% betragen.

Übung 1:
Die Behauptung H_1, dass mehr als 20% aller ABC – Schützen Linkshänder sind, soll anhand einer Stichprobe von 80 Kindern getestet werden. Findet man mehr als 20 Linkshänder so wird H_1 als zutreffend eingestuft.
a) Wie groß ist das Signifikanzniveau der Tests (Irrtumswahrscheinlichkeit 1.Art)
b) Mit welcher Wahrscheinlichkeit wird verworfen, wenn der wahre Anteil von Linkshändern unter den ABC – Schützen 30% beträgt?