t-Test für unabhängige Stichproben

Inhaltsverzeichnis

t-Test für unabhängige Stichproben

Lernhinweise

Durch die Anwendung des t-Tests für unabhängige Stichproben können Sie entscheiden, ob zwei unabhängige Stichproben aus zwei Populationen stammen, die hinsichtlich der arithmetischen Mittelwerte identisch sind.

Benötigte Vorkenntnisse

Sie können diesen Lernschritt effizient bearbeiten, wenn Sie die in Ihrem Curriculum vorgesehene Vorbereitungsliteratur bearbeitet haben und mit dem Grundprinzip der entscheidungsstatistischen Verfahren vertraut sind.

Lernziele

Durch die Bearbeitung dieses Lernschritts können Sie folgende Fähigkeiten erwerben:

- Sie können entscheiden, in welchen Situationen der t-Test für unabhängige Stichproben angezeigt ist.
- Sie können den t-Test für unabhängige Stichproben "manuell" anwenden und entscheiden, ob zwei Stichproben aus Populationen stammen, die hinsichtlich des arithmetischen Mittelwertes identisch sind.

Hinweise zur Bearbeitung des Lernschritts

- Empfohlene Bearbeitung: Beginnen Sie mit "Einführung".
- Klicken Sie auf den Button "next", er führt Sie auf den empfohlenen Pfad.
- Wenn Sie auf eine Rubrik klicken, gelangen Sie automatisch zur entsprechenden ersten Seite. In der Regel gelangen Sie zum Inhaltsverzeichnis der angewählten Rubrik
- Für die Bearbeitung sollten Sie ca. eine Stunde Lernzeit einplanen.

Hinweise zur Bearbeitung des Fallbeispiels

- Das Ziel der manuellen Bearbeitung ist, dass Sie das Grundprinzip handelnd nachvollziehen können.
- Das Fallbeispiel ist so konstruiert, dass die Aufgaben problemlos mit einem Taschenrechner (mit Wurzelfunktion) gelöst werden kann. Sie brauchen kein Statistikprogramm wie z. B. SPSS.
- Bei der exemplarischen Problemlösung werden Vertiefungsfragen angeboten. Wenn Sie diese nicht lösen wollen, dann gehen Sie zur nächsten Seite.
- Wenn Sie die Vertiefungsaufgaben lösen, können Sie die Lösungen und deren Begründungen separat abrufen.
- Es kann sein, dass sich Ihre Resultate bei den Dezimalstellen aufgrund von Rundungsfehlern (zwei Dezimalstellen) leicht von den Musterlösungen unterscheiden. Das darf Sie nicht verwirren. Das Fallbeispiel ist so konzipiert, dass die Entscheidungen zur Signifikanz eindeutig ausfallen

Der vorliegende Lernschritt zum **t-Test für unabhängige Stichproben** bildet zusammen mit den Lernschritten zum **F-Test** und zum **t-Test für abhängige Stichproben** eine Einheit. Das Fallbeispiel berücksichtigt alle drei Testverfahren.

Einführung

Auszug aus dem Protokoll der Schulpflege Gerlechingen vom 15. Juli 2002

Schulpflegepräsident:

Unerfreuliches ist in unserer schönen Gemeinde passiert. Seit so vielen Jahren führen wir den Schülerwettkampf mit unseren Mittelstufenklassen durch, noch nie kam so etwas vor, dass ein so herrlicher Anlass von einem Dopingverdacht überschattet wurde. Sogar der Lokalanzeiger von Gerlechingen berichtete in seiner Juli-Ausgabe 2002 darüber. Frau Fischer aus der Schulpflege hat angeboten, sich der Sache anzunehmen.

Zeitungsartikel

Lokalanzeiger von Gerlechingen

Jahrgang 120 Auflage 20 000 Erscheint monatlich Juli 2002
Preis Fr. 1.25 Inserateannahmeschluss jeweils am 15. des Vormonats

Doping am Schuelerwett–kampf?

Seit ueber achtzig Jahren wird in unserer Gemeinde ein Schuelerwettkampf der Mittel–stufenklassen durchgefuehrt. In diesem Jahr wurde der Anlass durch den Verdacht auf einen Doping–Fall ueberschattet.

Seit über achtzig Jahren berichten wir auch von diesem herrlichen Anlass der Mittelstufenklassen unserer Gemeinde. Das Motto dieses Jahres hiess: "Auch fit im Kopf". Die 105 Kinder wurden durch Losentscheid in zwei Gruppen aufgeteilt. Die eine Gruppe nannte sich "Auerhahn" und die andere "Bibihenne". Neben den sieben sportlichen Aufgaben, hatten die Kinder einen Test zur Merkfähigkeit zu lösen. Die Gruppe "Bibihenne" oder auch Gruppe B genannt, gewann in dieser Disziplin. Ein erfreulicher Erfolg. Doch der Schein trügt, die Enttäuschung war riesengross. Der Hauswart fand im Aufenthaltsraum der Gruppe "Bibihenne" einige Getränkedosen mit der Aufschrift "Stimubrain". Ein Getränk, das die Denk- und Merkfähigkeit von Kindern steigern soll. Es soll auch wirksam gegen Schulmüdigkeit sein.
Waren die Kinder der Gruppe "Bibihenne" gedopt?
Die Lehrerinnen und Lehrer wandten sich an die Schulpflege und baten sie um Hilfe bei der Klärung des Falls.

Frau Fischer:

Ich habe mich entschlossen, der ganzen Sache mit statistischen Mitteln nachzugehen.
Zum jetzigen Zeitpunkt kann ich nur Folgendes sagen: Die Resultate beim Merkfähigkeitstest sind in den beiden Gruppen A und B ähnlich verteilt (homogene Varianzen: statistisches Prüfverfahren F-Test). Es ist nicht so, dass es in einer Gruppe nur extrem schlechte oder nur extrem gute Schüler gibt. Auffällig ist, dass in der Gruppe A niemand 10 und in der Gruppe B niemand einen Punkt gemacht hat. Wir können dennoch davon ausgehen, dass sich die Resultate jeweils etwa gleich um die durchschnittliche Gruppenleistung verteilen.
Im nächsten Schritt werde ich untersuchen, ob die Gruppe B durchschnittlich tatsächlich bessere Leistungen im Merkfähigkeitstest erbringt, oder ob der Unterschied zwischen den Gruppenmittelwerten auch zufällig entstanden sein könnte.

Schulpflegepräsidentin:
Ich danke Frau Fischer für die Bemühungen.

Theorie (1-7)

Inhaltsübersicht

1. Inhaltliche Fragestellung

In zwei unabhängigen Stichproben von unterschiedlichem Umfang wurde ein intervall-skaliertes Merkmal X erhoben. Dabei zeigt sich, dass die Mittelwerte der beiden Verteilungen des Merkmals nicht identisch sind.

Dieses Element (Animation, Video etc.) kann in der PDF version nicht dargestellt werden und ist nur in der online Version sichtbar. [link]

Es soll die Frage geklärt werden, ob der Unterschied zwischen den beiden Mittelwerten so gravierend ist, dass bei der Interpretation des Unterschieds der Zufall ausgeschlossen werden kann.
Anders ausgedrückt: Darf der Unterschied in den Stichprobenmittelwerten auf die hinter den Stichproben stehenden Populationen verallgemeinert werden?

2. Statistisch bearbeitbare Fragestellung

Auch hier müssen wir unsere Frage so formulieren, dass sie statistisch bearbeitbar ist.

Dieses Element (Animation, Video etc.) kann in der PDF version nicht dargestellt werden und ist nur in der online Version sichtbar. [link]

Die statistisch bearbeitbare Frage lautet: Stammen die beiden Stichproben aus Populationen, in denen das Merkmal X mit identischen Mittelwerten verteilt ist?

3. Prüfgrösse

Unser Interesse gilt dem Vergleich der beiden Stichprobenmittelwerte

$\overline{X}_A$

und

$\overline{X}_B$

.

Vergleiche können auf zwei Arten formalisiert werden:

1. Als Differenz von zwei Kennwerten. Die Kennwerte sind gleich, wenn die Differenz Null ergibt.
2. Als Quotient von zwei Kennwerten. Die Kennwerte sind gleich, wenn der Quotient gleich 1 ist.

Prüfgrösse:

$$\left(\overline{X}_A - \overline{X}_B\right)$$

Wie ist nun diese Prüfgrösse verteilt, wenn die beiden Stichproben tatsächlich aus Populationen mit identischen Mittelwerten stammen?

4. Arbeitshypothese und Alternativhypothese

Die Statistiker nennen uns die Verteilung der Prüfgrösse für den Fall, dass die Stichproben aus Populationen stammen, in denen das Merkmal X mit identischen Mittelwerten verteilt ist, d.h. wenn die sogenannte Arbeitshypothese H_0 gültig ist.

Arbeitshypothese H_0:

$$\mu_A = \mu_B$$

Zu jeder Arbeitshypothese kann eine Alternativhypothese formuliert werden, die gerichtet oder ungerichtet und spezifisch oder unspezifisch sein kann. Wir wählen als Beispiel die

Alternativhypothese H_1:

$$\mu_A \neq \mu_B$$

Diese Alternativhypothese ist **ungerichtet** und **unspezifisch**. Sie postuliert nur, dass die beiden Mittelwerte ungleich sind (

$$\mu_A \neq \mu_B$$

). Sie macht keine Angabe, welcher Mittelwert grösser sei (ungerichtet) und um welchen Betrag sich die beiden Mittelwerte unterscheiden (unspezifisch). Eine solche Alternativhypothese formulieren wir immer, wenn wir uns nur für den absoluten Betrag des Unterschieds zwischen den Stichprobenmittelwerten

$\overline{X}_A$

und

$\overline{X}_B$

interessieren, d.h. wenn das Vorzeichen des Unterschieds nicht von Bedeutung ist.

5. Prüfverteilung und Übertretungswahrscheinlichkeit

Bei Gültigkeit von H_0 ist die Verteilung der Prüfgrösse, d.h. die Prüfverteilung bekannt. Je nach der Grösse der beiden Stichproben müssen zwei Fälle unterschieden werden.

Prüfverteilung

- Für **grosse Stichproben**, d.h. in diesem Fall, wenn $(n_A + n_B) > 50$, ist die Prüfgrösse

  $$\left(\overline{X}_A - \overline{X}_B\right)$$

  mit den Parametern

  $$\mu_{(\overline{X}_A - \overline{X}_B)}$$

  und

  $$\sigma_{(\overline{X}_A - \overline{X}_B)}$$

  normalverteilt.
- Für **kleine Stichproben**, d.h. wenn $(n_A + n_B)$ # 50, ist die Prüfgrösse

  $$\left(\overline{X}_A - \overline{X}_B\right)$$

  mit den Parametern

  $$\mu_{(\overline{X}_A - \overline{X}_B)}$$

  und

  $$\sigma_{(\overline{X}_A - \overline{X}_B)}$$

  t-verteilt. In diesem Fall muss vorausgesetzt werden, dass das Merkmal in den hinter den Stichproben stehenden Populationen normalverteilt ist.

Wir wollen uns dies schrittweise veranschaulichen:

Dieses Element (Animation, Video etc.) kann in der PDF version nicht dargestellt werden und ist nur in der online Version sichtbar. [link]

Somit kennen wir die Form der Prüfverteilung. Was uns noch fehlt, sind die Parameter

$$\mu_{(\overline{X}_A - \overline{X}_B)}$$

und

$\sigma_{(\overline{x}_A - \overline{x}_B)}$

der Prüfverteilung und – bei kleinen Stichproben - der Freiheitsgrad df, der die spezifische t-Verteilung festlegt. Wir folgen den Überlegungen der Statistiker:

Dieses Element (Animation, Video etc.) kann in der PDF version nicht dargestellt werden und ist nur in der online Version sichtbar. [link]

Mit dem F-Test entscheiden wir als erstes, ob homogene oder inhomogene Varianzen vorliegen; dann schätzen wir die Standardabweichung der Prüfverteilung mit der entsprechenden Schätzformel. Die Schätzformeln sind in verschiedenen Lehrbüchern unterschiedlich dargestellt. Sie führen aber alle zum selben Resultat.
Achtung: Die Schätzformeln basieren auf sog. Punktschätzungen der Populationsvarianzen aus den Stichprobenvarianzen. Da diese Punktschätzungen ein normalverteiltes Merkmal in den Populationen voraussetzen, wird hier implizit auch für grosse Stichproben die Voraussetzung eines normalverteilten Merkmals in den Populationen eingeführt.
Damit sind die Form und die Parameter der Prüfverteilung bekannt. Dies bedeutet, dass wir wissen, wie die Prüfgrösse verteilt ist, wenn die Arbeitshypothese H_0 gültig ist, d.h. wenn die Stichproben aus Populationen stammen, in denen das Merkmal mit identischem Mittelwert verteilt ist.

Überschreitungswahrscheinlichkeit
Ordnen wir nun den konkreten Ausprägungsgrad unserer Prüfgrösse

$(\overline{x}_A - \overline{x}_B)$

in die Prüfverteilung ein, so lässt sich abschätzen, mit welcher Wahrscheinlichkeit dieser Ausprägungsgrad zufällig erreicht oder überschritten wird.
Wir haben eine ungerichtete Alternativhypothese H_1 formuliert, weil wir uns für die Zufälligkeit des absoluten Unterschieds in den Stichprobenmittelwerten interessieren. Dies bedeutet, dass das Vorzeichen der Prüfgrösse in diesem Fall keine Bedeutung hat und damit die zweiseitige Überschreitungswahrscheinlichkeit bestimmt werden muss.
Zur Bestimmung der Überschreitungswahrscheinlichkeit stehen uns nur die Tabellen zur z- und t-Verteilung zur Verfügung. Um diese zu nutzen, muss die Prüfgrösse in die entsprechende Verteilung transformiert werden.

Dieses Element (Animation, Video etc.) kann in der PDF version nicht dargestellt werden und ist nur in der online Version sichtbar. [link]

Mit der Überschreitungswahrscheinlichkeit p kennen wir die Wahrscheinlichkeit, mit der der absolute Unterschied zwischen den beiden Stichprobenmittelwerten

$\overline{x}_A$

und

$\bar{x}_B$

zufällig entstanden sein kann.

Ist p sehr klein, so ist es sehr unwahrscheinlich, dass die Stichproben aus Populationen mit identischen Mittelwerten stammen. Wir verwerfen H_0 zugunsten von H_1.

Ist p hingegen gross, so ist nicht auszuschliessen, dass die beiden Stichproben aus Populationen mit identischen Mittelwerten stammen. Wir behalten H_0 bei.

Was nun ,sehr klein' und ,gross' bedeutet, beurteilen wir wie immer anhand der Signifikanzniveaus.

6. Beurteilung der Übertretungswahrscheinlichkeit; Signifikanzniveau

Die ermittelte Überschreitungswahrscheinlichkeit beurteilen wir an den üblichen Signifikanzniveaus von 5%, 1% und 0,1 %.

Wir wollen uns dies am Beispiel grosser Stichproben und für den Fall ungerichteter und gerichteter Alternativhypothesen veranschaulichen. Wählen Sie erst eine gerichtete oder eine ungerichtete Alternativhypothese und dann ein Beispiel.

Dieses Element (Animation, Video etc.) kann in der PDF version nicht dargestellt werden und ist nur in der online Version sichtbar. [link]

Ein signifikantes Resultat bedeutet, dass bei der Interpretation des Unterschieds in den Stichprobenmittelwerten

$$\left(\bar{x}_A - \bar{x}_B\right)$$

der Zufall - mit der ermittelten Irrtumswahrscheinlichkeit p - ausgeschlossen werden darf.

Bei ungerichteter Alternativhypothese H1 interpretieren wir nur bezüglich des absoluten Betrags des Unterschieds, bei gerichteter Alternativhypothese H1 auch bezüglich der Richtung des Unterschieds, d.h. auch bezüglich seines Vorzeichens.

7. Zusammenfassung zum Lernschritt

Voraussetzungen

Die Ausprägungsgrade des intervall- oder proportional skalierten Merkmals müssen in den hinter den Stichproben stehenden Populationen näherungsweise normalverteilt sein.

Prüfgrösse

Arbeitshypothese und Alternativhypothese

H_0:

$$\mu_A = \mu_B$$

Beispiele für H_1:

$\mu_A \neq \mu_B$

(ungerichtet, unspezifisch)

μ_A

>

μ_B

(gerichtet, unspezifisch

Prüfverteilung

Grosse Stichproben, d.h. ($n_A + n_B$) > 50: Normalverteilung

Kleine Stichproben, d.h. ($n_A + n_B$) # 50: t-Verteilung mit dem Freiheitsgrad df

Parameter der Prüfverteilung:

- $\mu_{(\bar{x}_A - \bar{x}_B)}$

  = 0
- Homogene Varianzen, d.h.

  s_A^2

  und

  s_B^2

  unterscheiden sich nicht signifikant:

  $$\hat{\sigma}_{(x_A - x_B)} = \sqrt{\frac{s_A{}^2 * n_A + s_B{}^2 * n_B}{(n_A + n_B - 2)} * \left(\frac{1}{n_A} + \frac{1}{n_B}\right)}$$

  ; df =$n_A + n_B$ - 2
- Inhomogene Varianzen, d.h.

  s_A^2

  und

  s_B^2

  unterscheiden sich signifikant:

  $$\hat{\sigma}_{(x_A - x_B)} = \sqrt{\frac{s_A 2}{(n_A - 1)} + \frac{s_B{}^2}{(n_B - 1)}}$$

, df =(n_A + n_B - 2) / 2

Zur manuellen Bestimmung der Überschreitungswahrscheinlichkeit muss die Prüfgrösse in die z- resp. t-Verteilung transformiert werden.

$$z_{(x_A - x_B)} = \frac{(\bar{x}_A - \bar{x}_B) - \mu_{(x_A - x_B)}}{\hat{\sigma}_{(x_A - x_B)}}$$

resp.

$$t_{(x_A - x_B)} = \frac{(\bar{x}_A - \bar{x}_B) - \mu_{(x_A - x_B)}}{\hat{\sigma}_{(x_A - x_B)}}$$

Bei ungerichteter Alternativhypothese H_1: Zweiseitige Überschreitungswahrscheinlichkeit. Bei gerichteter Alternativhypothese H_1: Einseitige Überschreitungswahrscheinlichkeit.

Beurteilung der Überschreitungswahrscheinlichkeit

p > 5%:

H_0 wird beibehalten. Bei unspezifischen Alternativhypothesen kann keine Irrtumswahrscheinlichkeit angegeben werden.

p # 5%:

H_0 wird mit der Irrtumswahrscheinlichkeit p zugunsten von H_1 abgelehnt.

Fallbeispiel (1-9)

Inhaltsübersicht

1. Ausgangslage, Problemstellung t-Test

Im vorausgegangenen Arbeitsschritt untersuchten wir, ob die Stichproben A und B aus Populationen stammen, die hinsichtlich der Varianz identisch sind.

Diese Abklärung wurde durchgeführt, weil beim Vergleich der Populationsmittelwerte die Varianz der Prüfverteilung geschätzt werden muss. Die Schätzformeln berücksichtigen, ob homogene oder inhomogene Varianzen vorliegen. Anhand des F-Tests haben wir das geprüft und kamen zum Schluss, dass homogene Varianzen vorliegen.

Wollen Sie die wichtigsten Angaben zur Problemstellung sehen?

1. Am Sporttag eines Schulhauses mit 105 Mittelstufenkindern werden zwei Schülergruppen A und B gebildet.
2. Die Kinder werden per Losentscheid den Gruppen zugeteilt, damit sind die Stichproben voneinander unabhängig.
3. Die Kinder haben einen Merkfähigkeitstest zu lösen. Pro richtige Antwort erhalten die Kinder einen Punkt. Pro Kind gibt es maximal 10 Punkte im Merkfähigkeitstest.
4. Anhand der Gruppenmittelwerte soll entschieden werden, welche Gruppe den Merkfähigkeitstest besser löste.
5. Beim Aufräumen entdeckt man bei der Gruppe B einige Dosen "Stimubrain". Es besteht nun der Verdacht, dass die Kinder der Gruppe B gedopt waren.
6. Frau Fischer, ein Mitglied der Schulpflege, ist dabei, mit statistischen Mitteln herauszufinden, ob an dem Verdacht etwas Wahres ist.

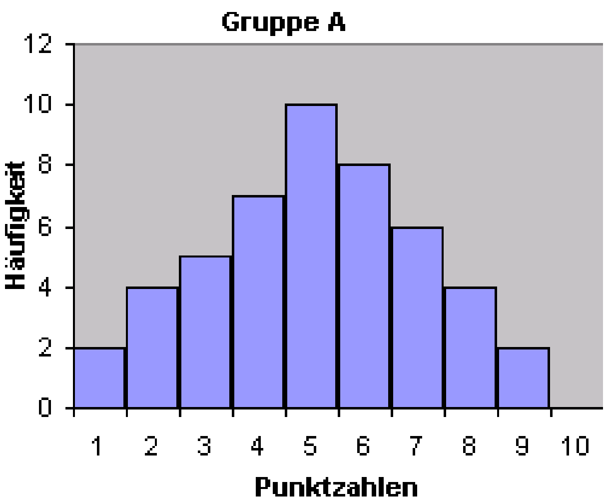
Wollen sie die statistisch relevanten Zahlen sehen: Häufigkeitsverteilungen und Kennzahlen

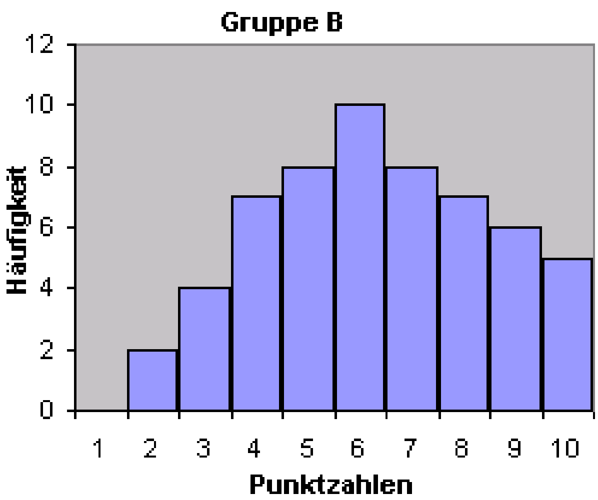
Gruppe A

Punkte	f
0	0
1	2
2	4
3	5
4	7
5	10
6	8
7	6
8	4
9	2
10	0
Total	**48**

Gruppe B

Punkte	f
0	0
1	0
2	2
3	4
4	7
5	8
6	10
7	8
8	7
9	6
10	5
	57





Kennzahlen	Gruppe A	Gruppe B
Anzahl Kinder	48	57
Gesamtpunktzahl	243	360
arith. Mittelwert	5.06	6.32
Standardabweichung	2.04	2.19

Im nächsten Schritt formulieren wir die inhaltliche Fragestellung.

2. Inhaltliche Fragestellung

Unterscheiden sich die beiden Schülergruppen A und B in ihrer durchschnittlichen Leistung so stark, dass davon ausgegangen werden muss, dass das Getränk "Stimubrain" eine leistungssteigernde Wirkung beim Lösen des Merkfähigkeitstests entfaltet?
Oder
Lösen Kinder mit "Stimubrain" Merkfähigkeitstests besser als Kinder ohne "Stimubrain"?
Im nächsten Schritt wird die Frage so formuliert, dass wir sie mit statistischen Verfahren bearbeiten können.

3. Statistisch bearbeitbare Fragestellung

Stammt die Gruppe B aus einer Population, die hinsichtlich der durchschnittlichen Leistung im Lösen des Merkfähigkeitstests bessere Werte erzielt als die Population aus der die Gruppe A stammt?
Im nächsten Schritt bestimmen wir die Prüfgrösse.

4. Prüfgrösse

Beim Vergleich der Stichprobenmittelwerte interessiert uns die Differenz zwischen

$\overline{x}_A$

und

$\overline{x}_B$

.
Die Prüfgrösse lautet somit:

$$\left(\overline{x}_A - \overline{x}_B\right)$$

Der konkrete Ausprägungsgrad der Prüfgrösse beträgt -1.26
Nun stellt sich die Frage, wie diese Prüfgrösse verteilt ist, d. h. es ist die Prüfverteilung zu bestimmen.
Hierfür müssen wir im nächsten Arbeitsschritt die Arbeitshypothese H_0 und die Alternativhypothese H_1 formulieren.

5. Arbeitshypothese und Alternativhypothese

Arbeitshypothese: Die beiden Stichproben A und B stammen hinsichtlich des Mittelwertes aus identischen Populationen (Nullhypothese H_0).

$$H_0 : \mu_A = \mu_B$$

Alternativhypothese: Die beidem Stichproben A und B stammen hinsichtlich des Mittelwertes aus unterschiedlichen Populationen, wobei der Populationsmittelwert der Gruppe A kleiner ist als derjenige der Gruppe B, bzw. derjenige der Gruppe B grösser ist als derjenige der Gruppe A (Alternativhypothese H_1).

$$H_1 : \mu_A < \mu_B$$

Diese Alternativhypothese ist **gerichtet** und **unspezifisch**.
Sie ist **gerichtet**, weil sie Angaben über die unterschiedliche Grösse der Mittelwerte macht. Sie postuliert, dass

μ_A

kleiner als

μ_B

ist.
Sie ist unspezifisch, weil sie keine Angaben darüber macht, um welchen Betrag

μ_A

kleiner ist als

μ_B

.
Jetzt können wir die Prüfverteilung bestimmen.

6. Prüfverteilung

Wenn es gemäss H_0 zutrifft, dass sich die beiden Populationen, aus denen die Stichproben stammen, hinsichtlich des Mittelwertes nicht unterscheiden, dann ist die Verteilung der Prüfgrösse bekannt.
In unserem Falle, d. h. bei grossen Stichproben, ist die Prüfgrösse

$\left(\overline{X}_A - \overline{X}_B\right)$

mit den Parametern

$\mu_{(\overline{X}_A - \overline{X}_B)}$

normalverteilt.
Wenn Ihnen klar ist, weshalb die Prüfgrösse

$\left(\overline{X}_A - \overline{X}_B\right)$

normalverteilt und nicht t-verteilt ist, dann gehen Sie weiter, sonst auf

Erklärung

Bei grossen Stichproben $(n_A + n_B) > 50$, bei uns 48 + 57 = 105, ist die Prüfgrösse mit den Parametern

$\mu_{(\overline{X}_A - \overline{X}_B)}$

und

> $\sigma_{(\overline{x}_A-\overline{x}_B)}$
>
> normalverteilt.
> Bei kleinen Stichproben ($n_A + n_B$) # 50, ist die Prüfgrösse
>
> $\left(\overline{x}_A - \overline{x}_B\right)$
>
> mit den Parametern
>
> $\mu_{(\overline{x}_A-\overline{x}_B)}$
>
> und
>
> $\sigma_{(\overline{x}_A-\overline{x}_B)}$
>
> t-verteilt.

Nun kennen wir die Form der Prüfverteilung. Was uns noch fehlt, sind die Parameter

$\mu_{(\overline{x}_A-\overline{x}_B)}$

und

$\sigma_{(\overline{x}_A-\overline{x}_B)}$

. Da wir von homogenen Varianzen ausgehen können, schätzen wir diese Parameter wie folgt:

$\mu_{(\overline{x}_A-\overline{x}_B)}$

= 0

Schätzformel für homogene Varianzen

$$\hat{\sigma}_{(\overline{x}_A-\overline{x}_B)} = \sqrt{\frac{4.15*48+4.79*57}{103}*0.04} = 0.42$$

Schätzformel

$$\hat{\sigma}_{(\overline{x}_A-\overline{x}_B)} = \sqrt{\frac{s_A{}^2*n_A+s_B{}^2*n_B}{(n_A+n_B-2)}*\left(\frac{1}{n_A}+\frac{1}{n_B}\right)}$$

Mit diesen Schätzwerten für die Parameter der Prüfverteilung können wir unseren konkret beobachteten Ausprägungsgrad der Prüfgrösse

$$\left(\overline{x}_A - \overline{x}_B\right)$$

= -1,26 in die z-Verteilung transformieren:
Formel für die z-Transformation

$$z_{(\overline{x}_A - \overline{x}_B)} = \frac{-1.26}{0.42} = -3.00$$

Formel für z-Transformation

mit:

$$\mu_{(\overline{x}_A - \overline{x}_B)}$$

= 0

$$z_{(\overline{x}_A - \overline{x}_B)} = \frac{(\overline{x}_A - \overline{x}_B) - \mu_{(\overline{x}_A - \overline{x}_B)}}{\hat{\sigma}_{(\overline{x}_A - \overline{x}_B)}}$$

$$z_{(\overline{x}_A - \overline{x}_B)} = \frac{(\overline{x}_A - \overline{x}_B)}{\hat{\sigma}_{(\overline{x}_A - \overline{x}_B)}}$$

Im nächsten Schritt bestimmen wir anhand der z-Tabelle die einseitige Überschreitungswahrscheinlichkeit. Wir bestimmen die einseitige Überschreitungswahrscheinlichkeit, weil wir eine gerichtete Alternativhypothese

$$H_0: \mu_A = \mu_B$$

postuliert haben.

7. Prüfung der Arbeitshypothese und Interpretation

Ermittlung der Überschreitungswahrscheinlichkeit
Die ermittelte Überschreitungswahrscheinlichkeit beurteilen wir anhand der üblichen Signifikanzniveaus von 5%, 1% und 0,1%.
Anhand der einseitigen z-Tabelle ermitteln wir die Wahrscheinlichkeit, mit welcher der z-Wert - 3.00 erreicht respektive unterschritten wird.
Ergebnis: Unsere Arbeitshypothese

$$H_0: \mu_A = \mu_B$$

wird zugunsten der **gerichteten** und **unspezifischen** Alternativhypothese

$$H_1: \mu_A < \mu_B$$

mit einer Irrtumswahrscheinlichkeit von p # 0,1% abgelehnt.

Interpretation

Die Ablehnung von H_0 bedeutet: Die beiden Gruppen A und B stammen aus Populationen mit unterschiedlichem Mittelwert.

Die Annahme von H_1 bedeutet: Die Gruppe B stammt mit einer Irrtumswahrscheinlichkeit von p # 1% aus einer Population, die einen grösseren Mittelwert hat als die Gruppe A.

8. Inhaltliche Interpretation, Anwendung auf Fallgeschichte

In der Schulbehörde und im Lehrerkollegium brechen aufgrund der Resultate Verwirrung aus. Frau Fischer wird mit verschiedenen Fragen konfrontiert. In der Hitze der Diskussion kann sie die Fragen und Ausrufe vorerst nur mit "ja" oder "nein" beantworten. Am Schluss gibt sie eine Interpretation und kündigt weitere Schritte an.

Dieses Element (Animation, Video etc.) kann in der PDF version nicht dargestellt werden und ist nur in der online Version sichtbar. [link]

Frau Fischer erklärt, wie das Resultat zu deuten ist

Wir wissen nur etwas: "Die Gruppe B löste den Merkfähigkeitstest durchschnittlich besser".

Sonst können wir keine Aussagen machen. Bei sportlichen Anlässen sind eben die einen Gruppen besser als die anderen. Nicht jede Gewinnermannschaft ist gedopt.

Den Grund, weshalb die Gruppe B besser ist, kann ich nicht sagen. Zum jetzigen Zeitpunkt kann ich nicht ausschliessen, dass die Gruppe B gedopt war, ich kann auch nicht sagen, dass sie es nicht war. Die durchschnittlich bessere Leistung der Gruppe B kann durch verschiedene Gründe bzw. Einflüsse zustande-kommen. Ich zähle Ihnen zwei Gründe auf:

- Die Gruppe B hat den Merkfähigkeitstest durchschnittlich besser gelöst, weil sie das besser kann.
- Die Gruppe B hat den Merkfähigkeitstest durchschnittlich besser gelöst, weil sie gedopt war.

Über das weitere Vorgehen werde ich bald informieren. Vorerst habe ich noch einige Abklärungen zu treffen. Zum jetzigen Zeitpunkt kann ich lediglich sagen, dass ich in meinem nächsten Arbeitsschritt den **t-Test für abhängige Stichproben** anwenden werde.

Auf der nächsten Seite wird die Berechnung mit SPSS vorgestellt.

9. SPSS-Ausdruck zum t-Test für unabhängige Stichproben

Zu den einzelnen Arbeitsschritten können Sie sich die entsprechenden SPSS-Ausdrucken mit den dazu gehörenden SPSS-Befehlen abrufen.

Häufigkeitsverteilung und Graphik der Gruppe A

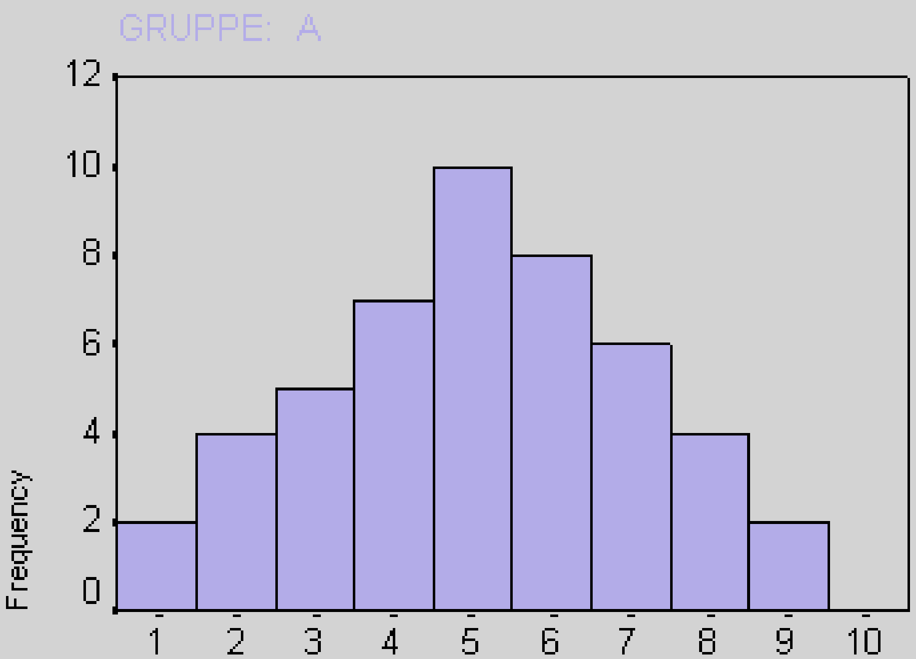
SPSS-Befehle
GET FILE "C:\DATEN\wettkampf.sav" SPLIT FILE SEPARATE BY gruppe. FREQUENCES VARIABLES = punktzah/HISTOGRAM.

PUNKTZAHL[a]

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	2	4.2	4.2	4.2
	2	4	8.3	8.3	12.5
	3	5	10.4	10.4	22.9
	4	7	14.6	14.6	37.5
	5	10	20.8	20.8	58.3
	6	8	16.7	16.7	75.0
	7	6	12.5	12.5	87.5
	8	4	8.3	8.3	95.8
	9	2	4.2	4.2	100.0
	Total	48	100.0	100.0	

a. GRUPPE = A

PUNKTZAHL



Häufigkeitsverteilung und Graphik der Gruppe B

SPSS-Befehle

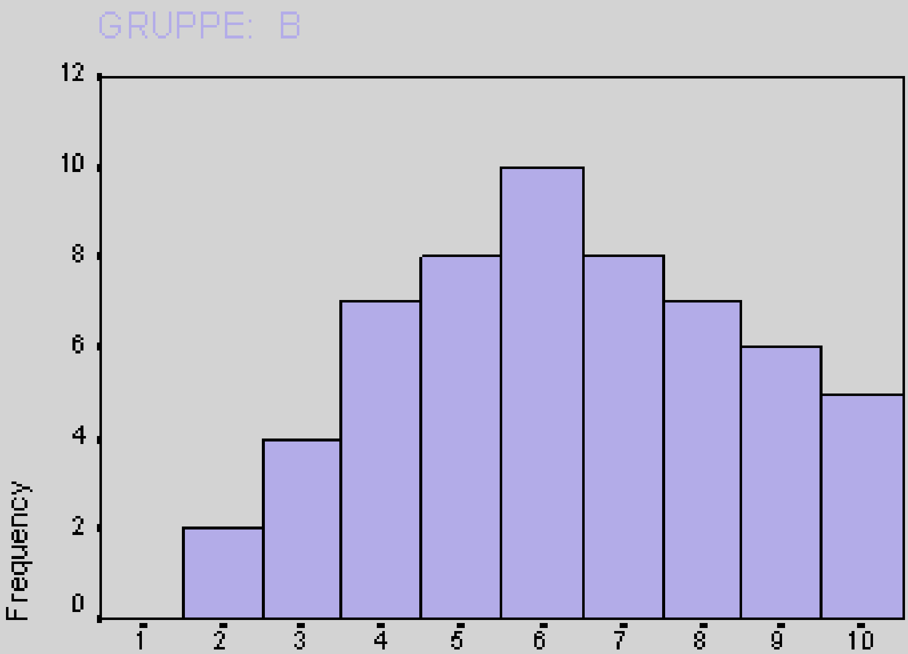
GET FILE "C:\DATEN\wettkampf.sav". SPLIT FILE SEPARATE BY gruppe. FREQUENCES VARIABLES = punktzah/HISTOGRAM.

PUNKTZAHL[a]

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2	2	3.5	3.5	3.5
	3	4	7.0	7.0	3.5
	4	7	12.3	3.5	22.8
	5	8	14.0	14.0	36.8
	6	2	17.5	17.5	54.4
	7	8	14.0	3.5	68.4
	8	7	12.3	12.3	80.7
	9	6	10.5	10.5	91.2
	10	5	8.8	8.8	100.0
	Total	57	100.0	100.0	

a. GRUPPE = B

PUNKTZAHL



Statistisch relevante Masszahlen und Signifikanzberechnung der Gruppen A und B

SPSS-Befehle

GET FILE "C:\DATEN\wettkampf.sav" T-TEST/GROUPS=GRUPPE ("A","B")/ VARIABLES=punktzah.

Group Statistics

GRUPPE		N	Mean	Std. Deviation	Std. Error Mean
PUNKTZAH	A	48	5.06	2.04	.29
	B	57	6.32	2.19	.29

Independent Samples Test

		Levene's Test for Equality of Variances		t-Test for Equality of Means			
		F	Sig	t	df	Sig. (2-tailed)	Mean Difference
PUNKTZAH	Equal variances assumed	.735	.393	-3.017	103	.003	-1.25
	Equal variances not assumed			-3.036	102.0	.003	-1.25

Erklärung zur SPSS-Ausgabe

Die SPSS-Ausgabe zeigt die Verteilungskennwerte für die beiden Stichproben A und B und die Resultate des t-Tests. Dieser gliedert sich in zwei Teile:

1. Prüfung auf Homogenität der Varianzen: Anstelle des besprochenen F-Tests, der speziell bei manuellen Auswertungen zum Einsatz kommt, prüft SPSS die Homogenität der Varianzen anhand des Tests nach Levene. Auch dieser Test verwendet eine F-verteilte Prüfgrösse, ihr Ausprägungsgrad und die Überschreitungswahrscheinlichkeit werden direkt ausgegeben. Bei einer Überschreitungswahrscheinlichkeit von 39,3% können wir von homogenen Varianzen ausgehen.
2. Bezüglich des t-Tests werden in zwei Zeilen die Testresultate bei homogenen und inhomogenen Varianzen ausgegeben. Anhand des Resultates des Tests nach Levene muss entscheiden werden, welche Auswertung gültig ist. In unserem Fall liegen homogene Varianzen vor, es gelten die Testresultate der Zeile "Equal variances assumed". SPSS nennt den Ausprägungsgrad der Prüfgrösse ($t = -3{,}017$), den Freiheitsgrad des Schätzwertes für die Varianz der Prüfverteilung (df) und die zweiseitige Überschreitungswahrscheinlichkeit des Ausprägungsgrades der Prüfgrösse (Sig. 2-tailed: $p = 0.003$). Da als Prüfverteilungen nur die symmetrischen z- oder t-

> Verteilungen in Frage kommen, kann die einseitige Überschreitungswahrscheinlichkeit durch Halbierung der zweiseitigen Überschreitungswahrscheinlichkeit gewonnen werden. In unserem Fall beträgt die einseitige Überschreitungswahrscheinlichkeit 0,15%, die Differenz zwischen den Stichprobenmittelwerten ist damit auf dem 1%-Niveau signifikant.

Nun sind Sie am Ende des Lernschritts t-Test für unabhängige Stichproben. Im nächsten Lernschritt wird anhand des t-Tests für abhängige Stichproben das Fallbeispiel definitv gelöst.