Einfache statistische Testverfahren

Johannes Hain

Lehrstuhl für Mathematik VIII (Statistik)

Hypothesentesten: Allgemeine Situation

Im Folgenden wird die statistische Vorgehensweise zur Durchführung einfacher Mittelwertsvergleiche behandelt. Wir unterscheiden mehrere Fälle:

- Eine Stichprobe
- Zwei Stichproben:
  - Ungepaarte Stichproben
  - Gepaarte Stichproben

Einer elementaren Eigenschaft müssen alle diese Stichproben genügen:

Grundlegende Voraussetzung

Die Daten jeder Teilstichprobe müssen in sich unabhängig und identisch verteilt sein.

Der Einstichproben t-Test

Voraussetzungen

Gegeben ist eine Stichprobe $X_1, \ldots, X_n$ von n unabhängigen Beobachtungen einer $N(\mu, \sigma^2)$-verteilten Zufallsvariable mit unbekanntem μ und σ^2.

Die zu untersuchende Nullhypothese lautet

$$H_0 : \mu = \mu_0$$

mit einem hypothetischen Wert μ_0. Der Name des Tests lautet **Einstichproben t-Test**.

Beispiel:
Eine Herstellerfirma umweltfreundlicher Engergiesparlampen behauptet, dass die Haltbarkeit ihrer Lampen 10.000 Stunden beträgt. In einem Langzeitversuch werden von $n = 25$ Energiesparlampen die Stundenzahlen gemessen, wie lange es dauert, bis die Lampe durchbrennt.

Der Einstichproben t-Test

Grundlegender Gedanke:
Berechnet man den Mittelwert $\bar{X}_n$ der Stundenzahl der 25 Energiesparlampen, so sollte sich dieser bei Gültigkeit der H_0 nicht stark von μ_0 unterscheiden. Je großer also die Differenz von $\bar{X}_n$ und μ_0 ist, desto eher wird man H_0 anzweifeln. Wird die Differenz zu groß, muss die Nullhypothese verworfen werden.

Um eine Aussage über die Gültigkeit von H_0 machen zu können schaut man auf die Teststatistik

$$T := \sqrt{n} \cdot \frac{\bar{X}_n - \mu_0}{S_n}$$

auch t-**Statistik** genannt. Diese ist t-verteilt mit $(n-1)$ Freiheitsgraden.

Der Einstichproben t-Test

Der Einstichproben t-Test in SPSS

$\rightarrow$ *Analysieren*

$\rightarrow$ *Mittelwerte vergleichen*

$\rightarrow$ *T-Test bei einer Stichprobe*

$\rightarrow$ *Variable auswählen:* Gebe Variable, deren Mittelwert untersucht werden soll in Feld *Testvariable* ein

$\rightarrow$ *Testwert:* Trage Wert ein, mit dem der Mittelwert verglichen werden soll

Nichtparametrische Testverfahren als Alternative

Bei diesem und auch allen anderen t-Tests wird in irgend einer Weise die Normalverteilung vorausgesetzt. Was aber, wenn diese Annahme nicht gerechtfertigt ist?

Ist die Normalverteilungsannahme verletzt, behilft man sich mit sog. **nichtparametrischen Verfahren**. Für jeden t-Test gibt es ein alternatives nichtparametrisches Testverfahren. Die Nullhypothese ist dabei die gleiche wie beim t-Test, d.h. bei der Interpretation des Testergebnisses muss man nichts neues beachten.

Vorteil der nichtparametrischen Verfahren ist, dass für diese Tests keine Normalverteilung vorausgesetzt wird!

Nichtparametrischer Einstichprobentest

Voraussetzungen

Gegeben ist eine unabhängige und identisch verteilte Stichprobe $X_1, \ldots, X_n$ mit dem unbekanntem Median m.

Die zu untersuchende Nullhypothese lautet

$$H_0 : m = m_0$$

mit einem hypothetischen Wert m_0.

- Ein verteilungsfreier Test für den Einstichprobenfall liegt in SPSS nicht vor. Um dennoch H_0 zu überprüfen kann man sich aber mit einem Trick behelfen. Man definiere sich eine neue Variable, bei der in jedem Fall der Wert m_0 steht.
- Für dieses Szenario – den Zweistichprobenfall – liegt ein verteilungsfreier Test in SPSS vor, der sogenannte **Wilcoxon-Test**.
- Details zu diesem Test findet man weiter unten auf Folie 24.

Nichtparametrischer Einstichprobentest

Der nichtparametrische Einstichprobentest in SPSS

- → Definiere neue Variable Y mittels *Transformieren* → *Variable berechnen*
- → *Analysieren*
- → *Nichtparametrische Tests*
- → *Alte Dialogfelder*
- → *Zwei verbundene Stichproben...*
- → Übertrage die zu testende Variable und die neu erzeugte Variable in das Feld *Testpaare* – eine Variable in die Spalte *Variable1*, die andere Variable in die Spalte *Variable2*

→ Eine alternative Durchführung des Wilcoxon-Tests wird auf Folie 27 beschrieben.

Der Zweistichproben t-Test für ungepaarte Stichproben

Voraussetzungen

- Es liegen zwei Teilstichproben $X_1, \ldots, X_m \sim N(\mu_X, \sigma_X^2)$ und $Y_1, \ldots, Y_n \sim N(\mu_Y, \sigma_Y^2)$ vor.
- Es liegt **Varianzhomogenität** vor: $\sigma_X^2 = \sigma_Y^2$.
- Die beiden Teilstichproben sind unabhängig voneinander erhoben worden.

Beispiel:
Anhand zweier zufällig ausgewählter Patientengruppen, soll die Wirksamkeit eines fiebersenkenden Medikaments untersucht werden. Dazu wird den $m = 35$ Patienten der **Behandlungsgruppe** das Medikament verabreicht, die $n = 40$ Patienten der **Kontrollgruppe** bekommen kein Medikament. Nach einem festgelegten Zeitraum wird von allen Patienten die Körpertemperatur gemessen.

Der Zweistichproben t-Test für ungepaarte Stichproben

Die Nullhypothese beim **Zweistichproben t-Test für unabhängige Stichproben** lautet:

$$H_0 : \mu_X = \mu_Y.$$

In Worten bedeutet dies, dass die Mittelwerte der beiden Zufallsvariablen X und Y gleich sind. Bezogen auf das Beispiel also, dass die Körpertemperatur von Behandlungs- und Kontrollgruppe gleich hoch sind.

Grundlegender Gedanke:
Hat das fiebersenkende Medikament keine Wirkung auf die Patienten, so sollten die Mittelwerte der beiden Patientengruppen in etwa gleich sein. Je größer also die Differenz von $\bar{X}_m$ und $\bar{Y}_n$ ist, desto eher wird man H_0 anzweifeln. Wird die Differenz zu groß, muss die Nullhypothese verworfen werden.

Der Zweistichproben t-Test für ungepaarte Stichproben

Um eine Aussage über die Gültigkeit von H_0 machen zu können schaut man auf die Statistik

$$T := \frac{\bar{X}_m - \bar{Y}_n}{\sqrt{\frac{1}{m} + \frac{1}{n}} \cdot S_p},$$

die t-verteilt mit $(m + n - 2)$ Freiheitsgraden, wobei

$$S_p^2 := \frac{(m-1)S_{X,m}^2 + (n-1)S_{Y,n}^2}{m+n-2}.$$

Bei der Gültigkeit von H_0 sollte T nahe bei Null liegen. Je größer, desto eher wird H_0 in Zweifel gezogen.

Die Voraussetzung der Varianzhomogenität wird mit dem **Levene-Test** von SPSS automatisch überprüft. Die Nullhypothese lautet:

$$H_0: \quad \text{Die Varianzen in beiden Stichproben sind gleich}$$

Wird H_0 verworfen, kann der t-Test nicht mehr angewendet werden (**Behrens-Fisher-Problem**).

Testen bei Varianzheterogenität

Im Fall von Varianzheterogenität wird statt des t-Tests der **Welch-Test** (auch **Satterthwaite-Test**) durchgeführt, bei dem die Voraussetzung der Varianzhomogenität fallen gelassen werden kann (aber auch nur diese!). Die Teststatistik ist in diesem Fall nur approximativ t-verteilt.

In SPSS wird dieser Welch-Test automatisch immer auch gleichzeitig mit dem t-Test ausgegeben.

Der Zweistichproben t-Test für ungepaarte Stichproben

Der t-Test für ungepaarte Stichproben in SPSS

- → *Analysieren*
- → *Mittelwerte vergleichen*
- → *T-Test bei unabhängigen Stichproben*
- → *Testvariablen auswählen*: Die Testvariable muss vom stetigen Typ sein!
- → *Gruppierungsvariable auswählen:* Gebe Gruppierungsvariable (z.B. Geschlecht) ein und klicke auf Schaltfläche *Gruppen definieren* um die Werte für die jeweiligen beiden Gruppen zu bestimmen

Der Mann-Whitney U-Test

Voraussetzungen

Gegeben sind zwei jeweils identisch verteilte Teilstichproben $X_1, \ldots, X_m$ und $Y_1, \ldots, Y_n$, die unabhängig voneinander erhoben wurden.

Es soll nun die folgende Nullhypothese untersucht werden:

H_0 : Die beiden Stichproben entstammen der gleichen Grundgesamtheit

Dies wird mit dem **Mann-Whitney** U-**Test** oder auch nur U-**Test** überprüft.

Um das Vorgehen dieses Tests zu verstehen, muss man sich mit dem Begriff des **Rangs** einer Beobachtung vertraut machen:

Der Mann-Whitney U-Test

Eine häufige Vorgehensweise bei nichtparametrischen Verfahren ist die Bildung von **Rängen** aus der Stichprobe. Dies bedeutet, dass man bei einer realisierten Stichprobe $X_1, \dots, X_m$ die Werte **aufsteigend** nach der Größe ordnet, d.h. es gilt

$$X_{(1,m)} < X_{(2,m)} < \cdots < X_{(m-1,m)} < X_{(m,m)}.$$

Dann ist der **Rang** von $X_{(1,m)}$ gerade 1, usw., der Rang von $X_{(m,m)}$ ist also m.

Für den U-Test werden beide Gruppen zunächst zusammengefasst und jedem Wert wird ein Rang zugeordnet, d.h. die kleinste Beobachtung bekommt den Rang 1, die größte Beobachtung den Rang $m + n$ zugewiesen.

Hat man die Ränge bestimmt, wird von jeder Gruppe die jeweilige Rangsumme R_x und R_y berechnet, sowie die beiden folgenden Größen bestimmt:

$$U_x = mn + \frac{m(m+1)}{2} - R_x \quad \text{und} \quad U_y = mn + \frac{n(n+1)}{2} - R_y.$$

Grundlegender Gedanke:
Bei Gültigkeit von H_0, sollten die beiden Gruppen in der zu Beginn gebildeten Reihenfolge in etwa gleichmäßig verteilt sein, die Rangsummen R_x und R_y sollten also ungefähr die gleiche Größe haben.

Die Berechnung von U_x und U_y geschieht, um sicher zu stellen, dass nicht beispielsweise eine Gruppe sehr hohe und sehr niedrige Rangwerte besitzt, wohingegen sich für die andere Gruppe die mittleren Rangwerte ergeben hätten. In diesem Fall wären nämlich die beiden Rangsummen in etwa gleich, die Gruppen würden aber nicht einer gemeinsamen Grundgesamtheit zu Grunde liegen.

Der Mann-Whitney U-Test

Für die Teststatistik gilt dann:

$$U := \min\{U_x, U_y\}$$

U ist unter H_0 **approximativ** $N(\frac{mn}{2}, \frac{mn(m+n+1)}{12})$-verteilt, d.h. für **hinreichend große** m und n liefert der Test brauchbare Ergebnisse.

Faustregel für den U-Test

Damit die Ergebnisse des U-Tests genau genug sind, müssen gelten:

- $n \geq 4$ **und** $m \geq 4$
- $n + m \geq 20$

Der Mann-Whitney *U*-Test in SPSS

→ *Analysieren*

→ *Nichtparametrische Tests*

→ *Alte Dialogfelder*

→ *Zwei unabhängige Stichproben*

→ Bestimme die *Testvariable*

→ Bestimme die *Gruppenvariable* und gebe unter *Gruppen definieren* die relevanten Codierungen für die Gruppe ein

Der Mann-Whitney *U*-Test in SPSS

→ *Analysieren*

→ *Nichtparametrische Tests*

→ *Unabhängige Stichproben...*

→ Aktiviere das Feld *Felder*

→ Übertrage die unabhängigen Variablen das Feld *Testfelder*

→ Übertrage die abhängige Variable in das Feld *Gruppen* und bestätige mit *Ausführen*

Der t-Test für gepaarte Stichproben

Voraussetzungen

Gegeben sind zwei Stichproben $X_1, \dots, X_n$ und $Y_1, \dots, Y_n$, die durch die Bildung von Paaren $(X_i, Y_i), i = 1, \dots, n$ erhoben wurden.
Die Differenz der beiden Teilstichproben $D_i = X_i - Y_i, i = 1, \dots, n$ muss normalverteilt sein.

Da es sich hier um zwei verbundene Stichproben handelt (man spricht auch von einem **matched pairs-design**) muss die Annahme der Unabhängigkeit der beiden Stichproben fallen gelassen werden.

Beispiel:
Von $n = 35$ Patienten wird der Bluckdruck vor und nach der Einnahme eines blutdrucksenkenden Medikamentes gemessen. Es soll untersucht werden ob sich der Blutdruck gesenkt hat.

Der t-Test für gepaarte Stichproben

Die Nullhypothese lautet

$$H_0 : \mu_X = \mu_Y \quad \text{bzw.} \quad \mu_X - \mu_Y = 0,$$

also beispielsweise dass das Medikament keinen Einfluss hat. Dies ist der **Zweistichproben t-Test für gepaarte (verbundene) Stichproben**.

Grundlegender Gedanke:
Um zu untersuchen ob die Behandlung mit dem Medikament erfolgreich war wird von jedem Patient die Differenz $D_i := X_i - Y_i, i = 1, \ldots, n$ gebildet. Bei der Gültigkeit der H_0 sollten die Differenzen nahe bei 0 liegen.

Auf die auftretenden Differenzen wird dann der Einstichproben t-Test angewendet, mit dem Wert $\mu_0 = 0$.

Um den t-Test für gepaarte Stichproben durchführen zu können, muss die **Differenz** der beiden Stichproben normalverteilt sein.

$\rightarrow$ Es reicht nicht aus zu zeigen, dass die beiden Originalvariablen einer Normalverteilung folgen!!

Überprüfung der Voraussetzungen

$\rightarrow$ *Transformieren*

$\rightarrow$ *Variable berechnen*

$\rightarrow$ Erstelle neue Variable aus Differenz der beiden untersuchten Variablen

$\rightarrow$ Teste die neu erzeugte Variable auf Normalverteilung.

Der t-Test für gepaarte Stichproben

Der t-Test für gepaarte Stichproben in SPSS

$\rightarrow$ *Analysieren*

$\rightarrow$ *Mittelwerte vergleichen*

$\rightarrow$ *T-Test bei verbundenen Stichproben*

$\rightarrow$ *Variablenpaar auswählen:* Schiebe erste Variablen des Paares in Spalte mit *Variable 1*, die zweite Variable das Paares in Spalte mit *Variable 2*

Der Wilcoxon-Test

Voraussetzungen

Gegeben sind n unabhängige Wiederholungen eines Zufallspaares $(X_i, Y_i), i = 1, \ldots, n$.

Die Nullhypothese zum **Wilcoxon-Test** lautet:

$$H_0 : X_i - Y_i \text{ hat den Median } 0.$$

Vorgehen:

- Berechne die Differenzen $d_1 = X_1 - Y_1, \ldots, d_n = X_n - Y_n$.
- Berechne die Ränge R_i der absoluten Beträge $|d_1|, \ldots, |d_n|$.
- Bilde die Summe R_+ der Rangwerte, die zu positiven d-Werten gehören und die Summe R_- der Rangwerte, die zu negativen d-Werten gehören.

Der Wilcoxon-Test

Grundlegender Gedanke:
Unter der Nullhypothese, sollten die Differenzen d_i der n Beobachtungen nicht allzu stark voneinander abweichen. Demzufolge sollten auch die Vorzeichen der d-Werte in etwa mit der gleichen Häufigkeit auftreten. Überwiegt bei den d-Werten aber ein Vorzeichen zu stark, dann wird je nach dem entweder R_+ oder R_- zu groß, woraufhin der Test dann verwirft.

Die Teststatistik

$$Z := \min\{R_+, R_-\}$$

ist unter H_0 für eine Stichprobengröße $n > 25$ annähernd $N(\frac{n(n+1)}{4}, \frac{n(n+1)(n+2)}{24})$-verteilt.

Der Wilcoxon-Test in SPSS

→ *Analysieren*

→ *Nichtparametrische Tests*

→ *Alte Dialogfelder*

→ *Zwei verbundene Stichproben*

→ Übertrage die beiden zu testenden Variablen in das Feld *Testpaare* – eine Variable in die Spalte *Variable1*, die andere Variable in die Spalte *Variable2*

Der Wilcoxon-Test in SPSS

→ *Analysieren*

→ *Nichtparametrische Tests*

→ *Verbundene Stichproben...*

→ Aktiviere das Feld *Felder*

→ Übertrage die zu testenden Variablen in das Feld *Testfelder* und bestätige mit *Ausführen*

Aufgaben zur Vertiefung I

Aufgaben zum Datensatz `Kino.sav`

Überprüfe die beiden folgenden Nullhypothesen mit einem geeigneten Signifikanztest:

- H_0 : Männer und Frauen sind gleich alt
- H_0 : Männer und Frauen gehen gleich häufig ins Kino

Aufgaben zum Datensatz `MannFrau.sav`

Überprüfe die beiden folgenden Nullhypothesen mit dem korrekten Signifikanztest:

- H_0 : Männer und Frauen sind gleich groß
- H_0 : Männer und Frauen sind gleich alt

Aufgaben zur Vertiefung II

Aufgabe zum Datensatz `Arbeitsbeschaffung.sav`

Der Datensatz enthält das Bruttoeinkommen von Erwerbslosen vor und während einer Arbeitsbeschaffungsmaßnahme (ABM). Untersuche die Frage ob sich das Einkommen durch die ABM signifikant verändert, d.h.
$H_0 : \text{Einkommen}_{\text{vorher}} = \text{Einkommen}_{\text{nachher}}$.

Aufgabe zum Datensatz `Pisa.sav`

Gibt es zwischen den drei Leistungsparametern irgendwo signifikante Unterschiede? Untersuche also die drei Nullhypothesen:

- $H_0 : \text{Leistung}_{\text{Lesen}} = \text{Leistung}_{\text{Mathe}}$
- $H_0 : \text{Leistung}_{\text{Lesen}} = \text{Leistung}_{\text{Naturwissenschaft}}$
- $H_0 : \text{Leistung}_{\text{Mathe}} = \text{Leistung}_{\text{Naturwissenschaft}}$