# Statistik – Einführung

## *Tests für zwei Grundgesamtheiten*
## *Kapitel 11*

Statistik – WU Wien

Gerhard Derflinger · Michael Hauser · Jörg Lenneis · Josef Leydold ·
Günter Tirler · Rosmarie Wakolbinger



## Lernziele

1. Testen Hypothesen über Parameter zweier Grundgesamtheiten:
   - Hypothesen über deren Mittel (Erwartungswerte)
   - Hypothesen über deren Anteile

2. Unterscheiden unabhängige Gesamtheiten und Gesamtheiten, die miteinander in Beziehung stehen.

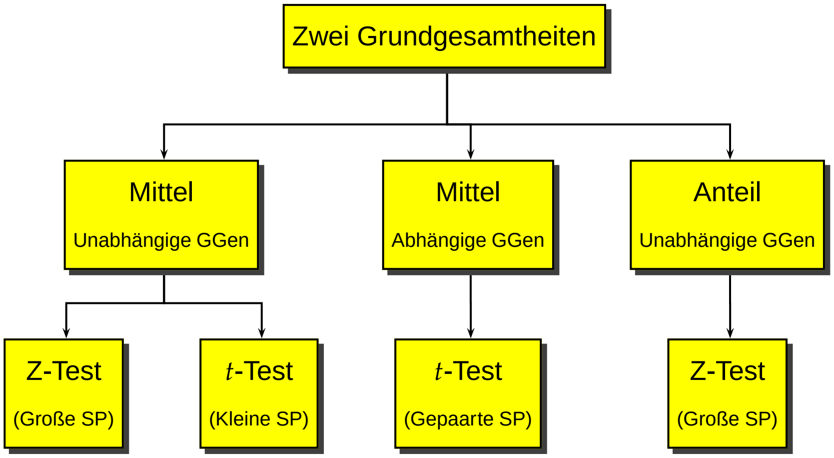


## Problem

Wie würden sie diese Fragen beantworten?

- Wer bekommt bessere Noten:
  Männer oder Frauen?

- Welches Betriebssystem ist schneller zu erlernen:
  Windows oder Linux?



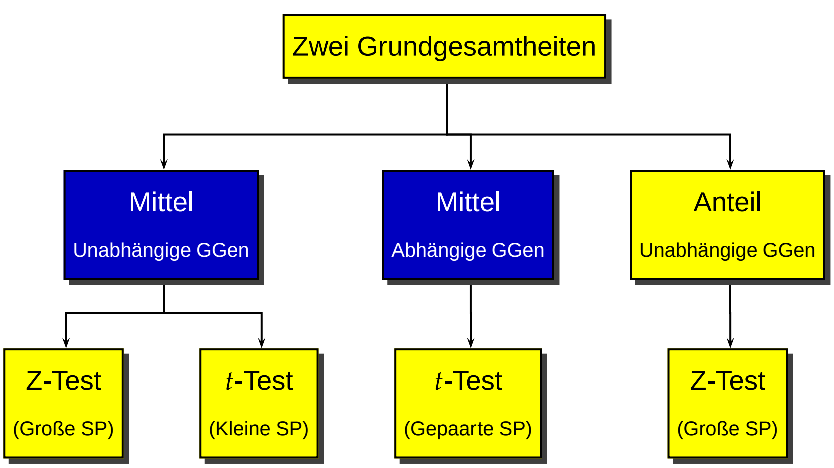



# Test zweier Mittel

## Unabhängiges Ziehen von Stichproben und Experimente mit gepaarten Differenzen



## Tests für zwei Grundgesamtheiten



| Unabhängig | Abhängig |
|---|---|
| 1. Verschiedene Datenquellen<br>• Ohne Beziehung<br>• Unabhängig | 1. Selbe Datenquelle<br>• Gepaart oder gematched<br>• Wiederholte Messungen (vorher / nachher) |
| 2. Verwende die Differenz zwischen den beiden Stichprobenmitteln.<br>• $\bar{X}_1 - \bar{X}_2$ | 2. Verwende die Differenz zwischen jedem Beobachtungspaar.<br>• $D_i = X_{1i} - X_{2i}$ |



## Zwei unabhängige GGen // Beispiele

1. Ein Ökonom möchte angeben, ob ein Unterschied im Mittel der Familieneinkommen von Haushalten aus 2 sozio-ökonomischen Gruppen besteht.
2. Der Studiendekan der WU möchte die Mittel der Mathematiknoten der Handelsakademiker und der AHS-Maturanten vergleichen.



## Zwei abhängige GGen // Beispiele

1. Nike möchte feststellen, ob ein Unterschied in der Haltbarkeit von 2 verschiedenen Sohlenmaterialien besteht. Das eine Material wird auf dem linken Schuh das andere auf dem rechten eines Paares aufgebracht.
2. Sie vergleichen die Anfangsgehälter von Männern und Frauen mit Universitätsabschluß. Um die Vergleichbarkeit zu verbesseren wird zu jedem Mann mit einem spezifischen Abschluß und Notendurchschnitt eine passende Frau gesucht. Diese Paare werden dann zum Vergleich der Gehälter herangezogen.
   Effekt: Verzerrungen durch die unterschiedliche Studien-präferenzen von Männern und Frauen werden vermieden.

**Unabhängig oder gepaart?**

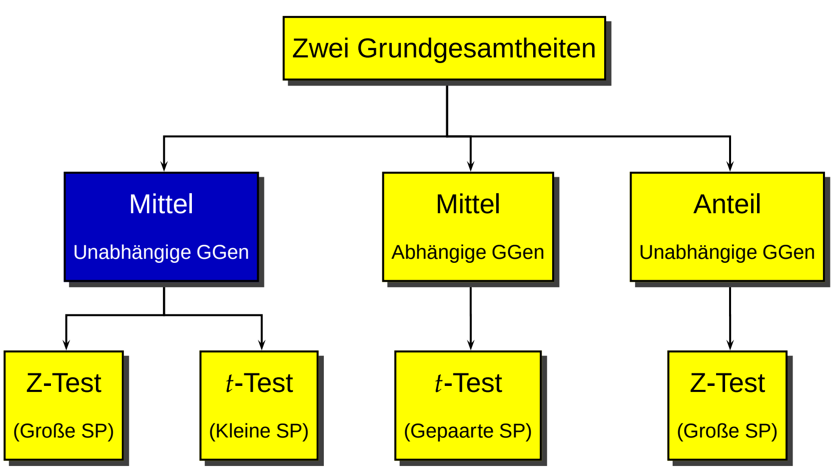
1. Benzinverbrauch von PKWs vor und nach der Montage von Radialreifen.
2. Die Brenndauer von Glühbirnen, die in 2 verschieden Fabriken hergestellt wurden.
3. Unterschied in der Härte von 2 Metalllegierungen: Eines enthält eine bestimmte Beimischung, das andere nicht.
4. Profilabnutzung von Motorradreifen: Vorderrad und Hinterrad einer Maschine.



# Testen zweier unabhängiger Mittelwerte



## Tests für zwei Grundgesamtheiten

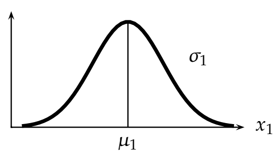


| | Forschungsfrage | | |
|---|---|---|---|
| | Kein Unterschied | GG1 $\geq$ GG2 | GG1 $\leq$ GG2 |
| Hypothese | Unterschied | GG1 $<$ GG2 | GG1 $>$ GG2 |
| $H_0$ | $\mu_1 - \mu_2 = 0$ | $\mu_1 - \mu_2 \geq 0$ | $\mu_1 - \mu_2 \leq 0$ |
| $H_A$ | $\mu_1 - \mu_2 \neq 0$ | $\mu_1 - \mu_2 < 0$ | $\mu_1 - \mu_2 > 0$ |

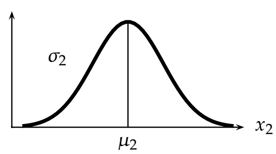


## Stichprobenverteilung

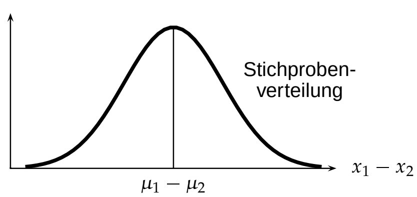
Grundgesamtheit 1




Grundgesamtheit 2

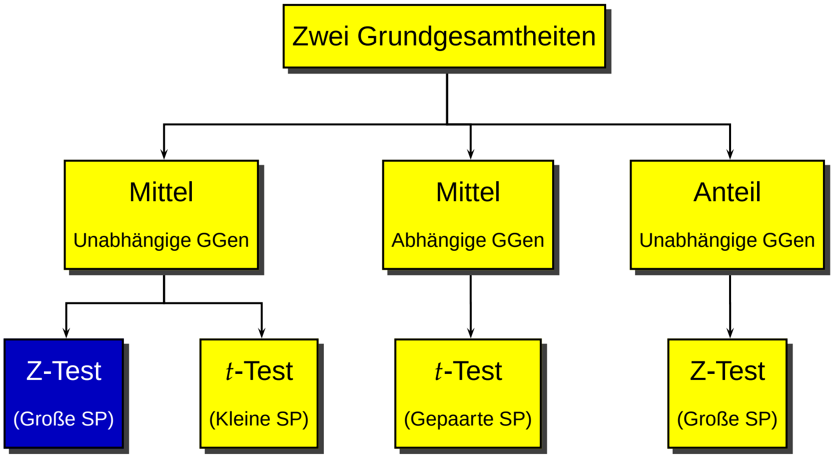



Zufallsstichprobe vom Umfang $n_1$. Berechne $\bar{x}_1$. $\rightarrow$ Berechne $\bar{x}_1 - \bar{x}_2$ für jedes mögliche Stichprobenpaar. $\leftarrow$ Zufallsstichprobe vom Umfang $n_2$. Berechne $\bar{x}_2$.






# Große Stichproben: Z-Test für 2 unabhängige Mittel






# Große SPen: Z-Test für 2 unabhängige Mittel

1. Annahmen
   - Unabhängige, zufällige Stichproben
   - Grundgesamtheiten normalverteilt, oder
   - näherungsweise normalverteilt falls $n_1 \geq 30$ und $n_2 \geq 30$.

2. Z-Statistik für zwei unabhängige Stichproben (standard-normalverteilt)

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$



# Große Stichproben Z-Test // Beispiel

Sie sind ein Finanzanalyst bei Meryll Lynch. Sie untersuchen, ob es einen Unterschied in den Dividendenerträgen zwischen den Aktien, die an der NYSE bzw. der NASDAQ gelistet sind, gibt. Sie sammeln folgende Daten:

| | NYSE | NASDAQ |
|---|---|---|
| Anzahl | 121 | 125 |
| Mittel | 3.27 | 2.53 |
| Std.Abw. | 1.30 | 1.16 |

Liegt ein Unterschied im durchschnittlichen Ertrag vor? ($\alpha = 0.05$)

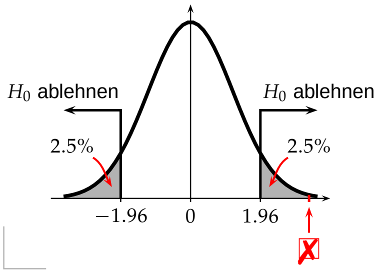
$H_0$: $\mu_1 - \mu_2 = 0 \quad (\mu_1 = \mu_2)$

$H_A$: $\mu_1 - \mu_2 \neq 0 \quad (\mu_1 \neq \mu_2)$

$\alpha = 0.05$

$n_1 = 121$, $n_2 = 125$

Kritische Werte: $\pm 1.96$




Teststatistik:

$$Z = \frac{(3.27 - 2.53)}{\sqrt{\frac{1.69}{121} + \frac{1.3456}{125}}} = 4.705$$

Entscheidung:

$H_0$ wird abgelehnt.

Interpretation:

Es gibt Evidenz auf einen Unterschied in den Mitteln.



# Große Stichproben Z-Test // Beispiel

Sie sind ein Ökonom in einem Bildungsministerium. Sie wollen wissen, ob sich die Ausgaben von BW-Studenten und Jus-Studenten unterscheiden. Sie erheben folgende Daten:

| | BW | Jus |
|---|---|---|
| Zahl | 35 | 35 |
| Mittel | 369 € | 380 € |
| Std.Abw. | 32.3 € | 36.6 € |

Liegt ein Unterschied in den durchschnittlichen Ausgaben vor? ($\alpha = 0.10$)



# Große Stichproben Z-Test // Lösung

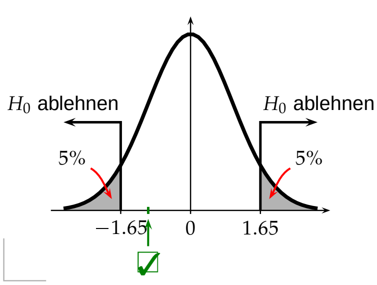
$H_0$: $\mu_1 - \mu_2 = 0 \quad (\mu_1 = \mu_2)$

$H_A$: $\mu_1 - \mu_2 \neq 0 \quad (\mu_1 \neq \mu_2)$

$\alpha = 0.10$

$n_1 = 35$, $n_2 = 35$

Kritische Werte: $\pm 1.645$




Teststatistik:

$$Z = \frac{(363 - 380)}{\sqrt{\frac{32.3^2}{35} + \frac{36.6^2}{35}}} = -1.333$$

Entscheidung:

$H_0$ wird nicht abgelehnt.

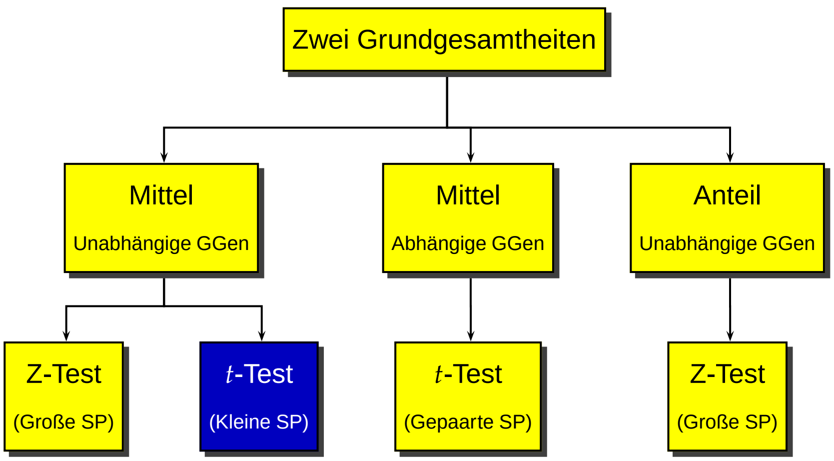
Interpretation:

Es gibt keinen signifikanten Unterschied in den Mitteln.

# Kleine Stichproben:
# $t$-Test für 2 unabhängige Mittel



## Tests für zwei Grundgesamtheiten






## Kleine SPen $t$-Test für 2 unabhängige Mittel

1. Testet die Mittel von 2 unabhängigen Grundgesamtheiten, die gleiche Varianzen haben.
2. Annahmen:
   - Unabhängige, zufällige Stichproben
   - Beide Grundgesamtheiten sind normal verteilt.
   - Die Varianzen in den Grundgesamtheiten sind unbekannt, aber gleich.

3. $t$-Statistik für zwei unabhängige Stichproben ($t$-verteilt mit $\nu$ Feiheitsgraden)

$$\boxed{t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}}$$

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

Freiheitsgrade: $\nu = n_1 + n_2 - 2$. $S_p^2$ heißt gepoolte Varianz.
(Beide Stichproben werden zu einer zusammengefasst, gepoolt.)



# Kleine Stichproben $t$-Test // Beispiel

Sie sind ein Finanzanalyst bei Meryll Lynch. Sie untersuchen, ob es einen Unterschied in den Dividendenerträgen zwischen den Aktien, die an der NYSE bzw. der NASDAQ gelistet sind, gibt. Sie sammeln folgende Daten:

| | NYSE | NASDAQ |
|---|---|---|
| Anzahl | 21 | 25 |
| Mittel | 3.27 | 2.53 |
| Std.Abw. | 1.30 | 1.16 |

Liegt ein Unterschied im durchschnittlichen Ertrag vor? ($\alpha = 0.05$)



# Kleine Stichproben $t$-Test // Lösung

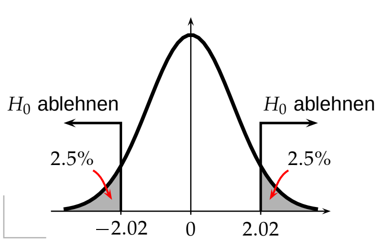
$H_0$: $\mu_1 - \mu_2 = 0 \quad (\mu_1 = \mu_2)$

$H_A$: $\mu_1 - \mu_2 \neq 0 \quad (\mu_1 \neq \mu_2)$

$\alpha = 0.05$

$n_1 = 21$, $n_2 = 25$
$\nu = 21 + 25 - 2 = 44$

Krit.W.: $\pm t_{44,0.975} = \pm 2.0154$




Teststatistik:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2(\frac{1}{n_1} + \frac{1}{n_2})}} = \frac{(3.27 - 2.53) - (0)}{\sqrt{1.502(\frac{1}{21} + \frac{1}{25})}} = 2.04$$

$$\begin{aligned} s_p^2 &= \frac{(n_1 - 1)\, s_1^2 + (n_2 - 1)\, s_2^2}{n_1 + n_2 - 2} = \frac{(21 - 1)\, 1.30^2 + (25 - 1)\, 1.16^2}{21 + 25 - 2} \\ &= 1.502 \end{aligned}$$



# Kleine Stichproben $t$-Test // Lösung

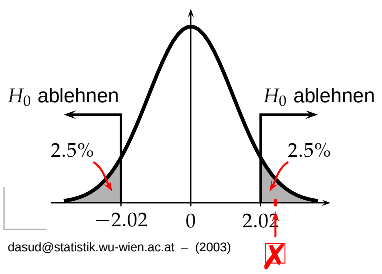
$H_0$: $\mu_1 - \mu_2 = 0 \quad (\mu_1 = \mu_2)$

$H_A$: $\mu_1 - \mu_2 \neq 0 \quad (\mu_1 \neq \mu_2)$

$\alpha = 0.05$

$n_1 = 21$, $n_2 = 25$
$\nu = 21 + 25 - 2 = 44$

Krit.W.: $\pm t_{44,0.975} = \pm 2.0154$




Teststatistik:

$$t = \frac{(3.27 - 2.53)}{\sqrt{1.502\left(\frac{1}{21} + \frac{1}{25}\right)}} = 2.04$$

Entscheidung:

$H_0$ wird abgelehnt.

Interpretation:

Es gibt Evidenz auf einen Unterschied in den Mitteln.



# Kleine Stichproben $t$-Test // Beispiel

Sie sind ein Forschungsassistent bei Peugeot. Bei Testfahrten werden folgende Daten über den Bezinverbrauch erhoben:

| | 405 | 406 |
|---|---|---|
| Zahl | 15 | 11 |
| Mittel | 8.7 | 7.8 |
| Std.Abw. | 1.9 | 1.5 |

Angenommen Standardabweichungen in den Grundgesamtheiten sind gleich. Gibt es einen Unterschied im Treibstoffverbrauch (Liter pro 100km) bei den zwei verschiedenen Modellen ? ($\alpha = 0.05$)

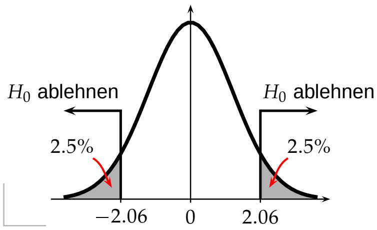
$H_0\colon \mu_1 - \mu_2 = 0 \quad (\mu_1 = \mu_2)$

$H_A\colon \mu_1 - \mu_2 \neq 0 \quad (\mu_1 \neq \mu_2)$

$\alpha = 0.05$

$n_1 = 15,\ n_2 = 11$
$\nu = 15 + 11 - 2 = 24$

Krit.W.: $\pm t_{24,0.975} = \pm 2.0639$




Teststatistik:



# Kleine Stichproben *t*-Test // Lösung

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(8.7 - 7.8) - (0)}{\sqrt{3.043\left(\frac{1}{15} + \frac{1}{11}\right)}} = 1.30$$

$$s_p^2 = \frac{(n_1 - 1)\, s_1^2 + (n_2 - 1)\, s_2^2}{n_1 + n_2 - 2} = \frac{(15-1)\,1.9^2 + (11-1)\,1.5^2}{15 + 11 - 2}$$
$$= 3.043$$



# Kleine Stichproben *t*-Test // Lösung

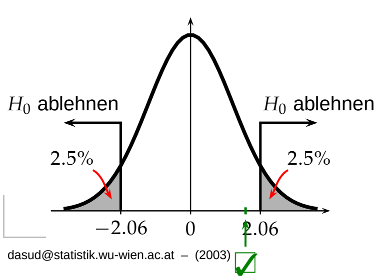
$H_0\colon \mu_1 - \mu_2 = 0 \quad (\mu_1 = \mu_2)$

$H_A\colon \mu_1 - \mu_2 \neq 0 \quad (\mu_1 \neq \mu_2)$

$\alpha = 0.05$

$n_1 = 15,\ n_2 = 11$
$\nu = 15 + 11 - 2 = 24$

Krit.W.: $\pm t_{24,0.975} = \pm 2.0639$




Teststatistik:

$$t = \frac{(8.7 - 7.8)}{\sqrt{3.043\left(\frac{1}{15} + \frac{1}{11}\right)}} = 1.30$$

Entscheidung:

$H_0$ wird nicht abgelehnt.

Interpretation:

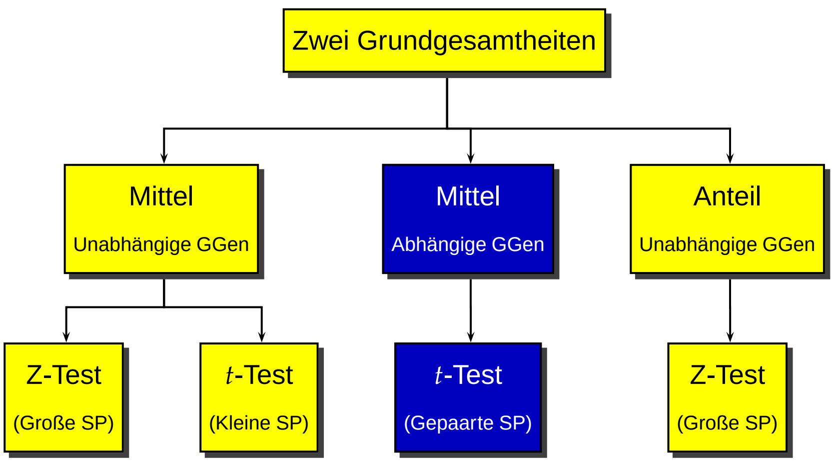
Es gibt keinen signifikanten Unterschied in den Mitteln.

# Gepaarte Stichproben *t*-Test

## Experimente mit gepaarten Differenzen



## Tests für zwei Grundgesamtheiten






## Gepaarte SPen *t*-Test für die Differenz von Mittel

1. Testet die Mittel von 2 in Beziehung stehenden Grundgesamtheiten.
   - Gepaart oder gematched
   - Wiederholte Messungen (vorher / nachher)
2. Eliminiert die Variation zwischen den Subjekten.
3. Annahmen:
   - Beide Grundgesamtheiten sind normal verteilt,
   - oder näherungsweise normalverteilt ($n_1 \geq 30$ und $n_2 \geq 30$).

| Hypothese | Forschungsfrage | | |
|---|---|---|---|
| | Kein Unterschied<br>Unterschied | GG1 $\geq$ GG2<br>GG1 $<$ GG2 | GG1 $\leq$ GG2<br>GG1 $>$ GG2 |
| $H_0$ | $\mu_D = 0$ | $\mu_D \geq 0$ | $\mu_D \leq 0$ |
| $H_A$ | $\mu_D \neq 0$ | $\mu_D < 0$ | $\mu_D > 0$ |

$\mu_D$ ist der Erwartungswert der Differenz $D$ der Zufallsvariablen $X_1$ und $X_2$. $\mu_D = E(D)$.

Differenz der $i$-ten Beobachtung: $D_i = X_{1i} - X_{2i}$



# Gepaarte SPen $t$-Test // Datentabelle

| Beobachtung | Gruppe 1 | Gruppe 2 | Differenz |
|---|---|---|---|
| 1 | $x_{11}$ | $x_{21}$ | $d_1 = x_{11} - x_{21}$ |
| 2 | $x_{12}$ | $x_{22}$ | $d_2 = x_{12} - x_{22}$ |
| ... | ... | ... | ... |
| $i$ | $x_{1i}$ | $x_{2i}$ | $d_i = x_{1i} - x_{2i}$ |
| ... | ... | ... | ... |
| $n$ | $x_{1n}$ | $x_{2n}$ | $d_n = x_{1n} - x_{2n}$ |



# Gepaarte SPen $t$-Test // Teststatistik

Teststatistik ($t$-verteilt mit $\nu$ FG):

$$\boxed{t = \frac{\bar{D} - \mu_D}{\frac{S_D}{\sqrt{n_D}}} \qquad \nu = n_D - 1}$$

Stichprobenmittel:

$$\bar{D} = \frac{1}{n_D} \sum_{i=1}^{n_D} D_i = \bar{X}_1 - \bar{X}_2$$

Stichprobenstandardabweichung:

$$S_D = \sqrt{\frac{1}{n_D - 1} \sum_{i=1}^{n_D} (D_i - \bar{D})^2}$$

Sie arbeiten in einer Personalabteilung, und wollen wissen, ob ein Trainingsprogramm effektiv ist. Sie sammeln folgende Testergebnisse:

| Name | vorher(1) | nachher(2) |
|---|---|---|
| Max | 85 | 94 |
| Sandra | 94 | 87 |
| Hugo | 78 | 79 |
| Marianne | 87 | 88 |

Ist das Trainingsprogramm wirksam? (Signifikanzniveau 10%)



# Rechentabelle

| Name | vorher(1) | nachher(2) | Differenz |
|---|---|---|---|
| Max | 85 | 94 | $-9$ |
| Sandra | 94 | 87 | 7 |
| Hugo | 78 | 79 | $-1$ |
| Marianne | 87 | 88 | $-1$ |
| Gesamt | | | $-4$ |



# Nullhypothese

1. War das Training effektiv?
2. Effektiv bedeutet:
   “Punkte-Nachher” > “Punkte-Vorher”.
3. Statistisch heißt dies:
   $\mu_2 > \mu_1$.
4. Umgeformt:
   $0 > \mu_1 - \mu_2$.
5. Wir definieren $\mu_D = \mu_1 - \mu_2$ und substituieren in (4):
   $0 > \mu_D$ bzw. $\mu_D < 0$.
6. Die **Alternativhypothese** lautet
   $H_A$: $\mu_D < 0$.

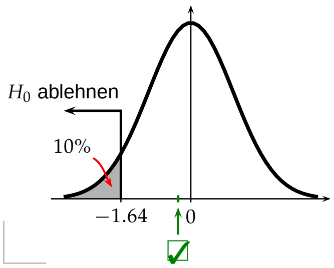
$H_0\colon \mu_D = 0 \quad (\mu_D = \mu_1 - \mu_2)$

$H_A\colon \mu_D < 0$

$\alpha = 0.10$

$n_D = 4, \quad \nu = 4 - 1 = 3$

Krit.W.: $-t_{3,0.90} = -1.6377$




Teststatistik:

$$t = \frac{\bar{d} - \mu_D}{\frac{s_D}{\sqrt{n_D}}} = \frac{-1 - 0}{\frac{6.53}{\sqrt{4}}} \approx -0.306$$

Entscheidung:

$H_0$ wird nicht abgelehnt.

Interpretation:

Es gibt keine Evidenz auf eine Wirksamkeit des Trainings.



# Gepaarte SPen $t$-Test // Beispiel

Sie sind ein Marktforscher und wollen die Preise eines Produktes ihres Klienten mit denen eines Konkurrenzprodukts vergleichen. Dazu erheben sie die Preise in 8 Einzelhandelsgeschäften. Wird das Produkt ihres Klienten billiger verkauft als das der Konkurrenz? (Signifikanzniveau 1%)

| Geschäft | (1) Klient € | (2) Konkurrenz € |
|---|---|---|
| 1 | 10 | 11 |
| 2 | 8 | 11 |
| 3 | 7 | 10 |
| 4 | 9 | 12 |
| 5 | 11 | 11 |
| 6 | 10 | 13 |
| 7 | 9 | 12 |
| 8 | 8 | 10 |



# Gepaarte SPen $t$-Test // Lösung

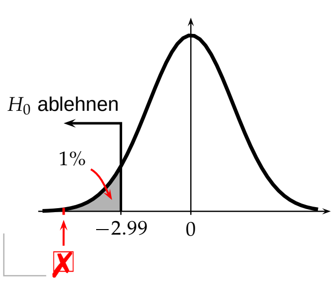
$H_0\colon \mu_D = 0 \quad (\mu_D = \mu_1 - \mu_2)$

$H_A\colon \mu_D < 0$

$\alpha = 0.01$

$n_D = 8, \quad \nu = 8 - 1 = 7$

Krit.W.: $-t_{7,0.99} = -2.998$




Teststatistik:

$$t = \frac{\bar{d} - \mu_D}{\frac{s_D}{\sqrt{n_D}}} = \frac{-2.25 - 0}{\frac{1.16}{\sqrt{8}}} \approx -5.486$$

Entscheidung:

$H_0$ wird abgelehnt.

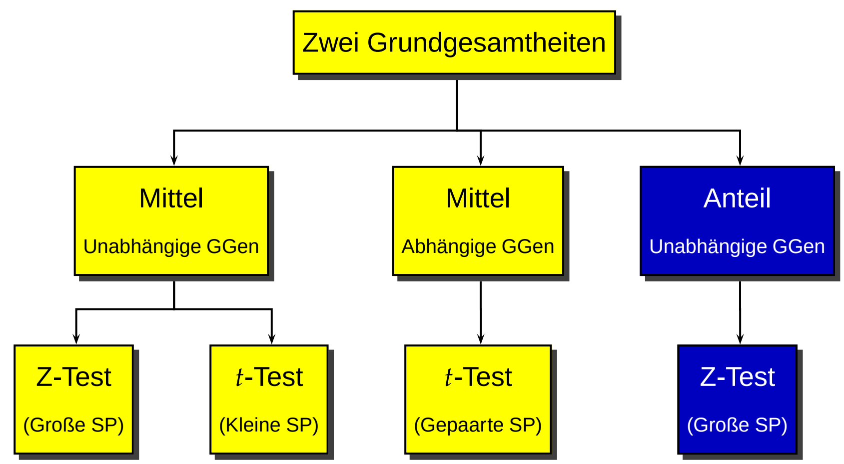
Interpretation:

Es gibt Evidenz, dass das Produkt des Klienten billiger verkauft wird.

# Z-Test für die Differenz von Anteilen



## Tests für zwei Grundgesamtheiten






## Z-Test für die Differenz von 2 Anteilen

1. Annahmen:
   - Die Grundgesamtheiten sind unabhängig.
   - Die Grundgesamtheiten folgen einer Binomialverteilung.
   - Die Verteilungen können durch die Normalverteilung approximiert werden. Das ist in der Regel der Fall, wenn $n_1\, p_1\,(1-p_1) \geq 9$ und $n_2\, p_2\,(1-p_2) \geq 9$.
2. Z-Teststatistik für 2 Anteile (standard-normalverteilt)

$$Z = \frac{(p_1 - p_2) - (\theta_1 - \theta_2)}{\sqrt{p\,(1-p)\,\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad \text{wobei} \quad p = \frac{X_1 + X_2}{n_1 + n_2}$$

$X_1$ bezeichnet die Anzahl der ausgewählten Einheiten in Stichprobe 1 (mit Umfang $n_1$). Analog für $X_2$.

| Hypothese | Forschungsfrage | | |
|---|---|---|---|
| | Kein Unterschied | GG1 $\geq$ GG2 | GG1 $\leq$ GG2 |
| | Unterschied | GG1 $<$ GG2 | GG1 $>$ GG2 |
| $H_0$ | $\theta_1 - \theta_2 = 0$ | $\theta_1 - \theta_2 \geq 0$ | $\theta_1 - \theta_2 \leq 0$ |
| $H_A$ | $\theta_1 - \theta_2 \neq 0$ | $\theta_1 - \theta_2 < 0$ | $\theta_1 - \theta_2 > 0$ |



## Z-Test für 2 Anteile // Beispiel

Als Leiter der Personalabteilung wollen sie zwei Personal-bewertungsmethoden in Hinblick auf die wahrgenommene Fairness testen. 63 von 78 Beschäftigten werten Methode 1 als fair, 49 von 82 die Methode 2.

Liegt eine Unterschied in der Bewertung bei einem Signifikanz-niveau von 0.01 vor?



## Z-Test für 2 Anteile // Lösung

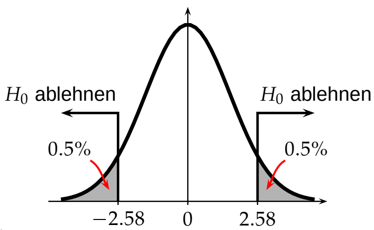
$H_0$: $\theta_1 - \theta_2 = 0 \quad (\theta_1 = \theta_2)$

$H_A$: $\theta_1 - \theta_2 \neq 0$

$\alpha = 0.01$

$n_1 = 78$, $n_2 = 82$

Kritische Werte: $\pm 2.58$




Teststatistik:

$$p_1 = \frac{X_1}{n_1} = \frac{63}{78} = 0.808 \qquad p_2 = \frac{X_2}{n_2} = \frac{49}{82} = 0.598$$

$$p = \frac{X_1 + X_2}{n_1 + n_2} = \frac{63 + 49}{78 + 82} = 0.70$$

$$Z = \frac{(p_1 - p_2) - (\theta_1 - \theta_2)}{\sqrt{p\,(1-p)\,\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.808 - 0.598) - (0)}{\sqrt{0.70\,(1-0.70)\,\left(\frac{1}{78} + \frac{1}{82}\right)}} = 2.897$$



## Z-Test für 2 Anteile // Lösung

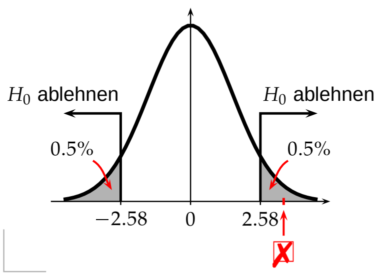
$H_0 \colon \theta_1 - \theta_2 = 0 \quad (\theta_1 = \theta_2)$

$H_A \colon \theta_1 - \theta_2 \neq 0$

$\alpha = 0.01$

$n_1 = 78,\ n_2 = 82$

Kritische Werte: $\pm 2.58$




Teststatistik:

$Z = 2.897$

Entscheidung:

$H_0$ wird abgelehnt.

Interpretation:

Es gibt Evidenz, dass die beiden Bewertungsmethoden unterschiedlich wahrgenommen werden.



## Z-Test für 2 Anteile // Beispiel

Sie sind Volkswirt im Arbeitsministerium. Sie studieren die Arbeitslosenraten der EU-Erhebung im Burgenland und in der Steiermark. Im Burgenland sind 74 von 1500 Befragten arbeitslos (suchen Arbeit und haben keine), in der Steiermark 129 von 1500.

Ist die Arbeitslosenrate (bei einem Signifikanzniveau von 0.05) im Burgenland kleiner als in der Steiermark?

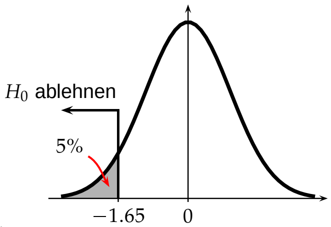
$H_0: \theta_B - \theta_S = 0 \quad (\theta_B = \theta_S)$

$H_A: \theta_B - \theta_S < 0 \quad (\theta_B < \theta_S)$

$\alpha = 0.05$

$n_B = 1500,\ n_S = 1500$

Kritischer Wert: $-1.645$




Teststatistik:



## Z-Test für 2 Anteile // Lösung

$$p_B = \frac{X_B}{n_B} = \frac{74}{1500} = 0.0493 \qquad p_S = \frac{X_S}{n_S} = \frac{129}{1500} = 0.0860$$

$$p = \frac{X_B + X_S}{n_B + n_S} = \frac{74+129}{1500+1500} = 0.0677$$

$$Z = \frac{(p_1 - p_2) - (\theta_1 - \theta_2)}{\sqrt{p\,(1-p)\,\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$= \frac{(0.0493 - 0.0860) - (0)}{\sqrt{0.0677\,(1 - 0.0677)\,\left(\frac{1}{1500} + \frac{1}{1500}\right)}} = -4.00$$



## Z-Test für 2 Anteile // Lösung

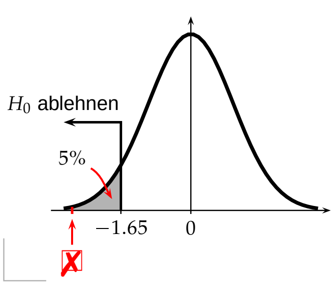
$H_0: \theta_B - \theta_S = 0 \quad (\theta_B = \theta_S)$

$H_A: \theta_B - \theta_S < 0 \quad (\theta_B < \theta_S)$

$\alpha = 0.05$

$n_B = 1500,\ n_S = 1500$

Kritischer Wert: $-1.645$




Teststatistik:

$Z = -4.00$

Entscheidung:

$H_0$ wird abgelehnt.

Interpretation:

Die Arbeitslosenrate ist im Burgenland signifikant niedriger als in der Steiermark.

1. Testeten Hypothesen über Parameter zweier Grundgesamtheiten:
   - Hypothesen über deren Mittel (Erwartungswerte)
   - Hypothesen über deren Anteile

2. Unterschieden unabhängige Gesamtheiten und Gesamtheiten, die miteinander in Beziehung stehen.