# Mittelwertvergleiche

- *Nächste Anwendung:* Vergleich der Mittelwerte zweier *normalverteilter* Zufallsvariablen $Y^A$ und $Y^B$
  1. auf **derselben** Grundgesamtheit durch Beobachtung von Realisationen $(x_1^A, x_1^B), \ldots, (x_n^A, x_n^B)$ einer (gemeinsamen) einfachen Stichprobe $(X_1^A, X_1^B), \ldots, (X_n^A, X_n^B)$ zur **zweidimensionalen** Zufallsvariablen $(Y^A, Y^B)$, insbesondere von Realisationen von $Y^A$ und $Y^B$ für **dieselben** Elemente der Grundgesamtheit („verbundene Stichprobe"),
  2. auf **derselben oder unterschiedlichen** Grundgesamtheit(en) durch Beobachtung von Realisationen $x_1^A, \ldots, x_{n_A}^A$ und $x_1^B, \ldots, x_{n_B}^B$ zu zwei **unabhängigen** einfachen Stichproben $X_1^A, \ldots, X_{n_A}^A$ und $X_1^B, \ldots, X_{n_B}^B$ (möglicherweise mit $n_A \neq n_B$) zu den beiden Zufallsvariablen $Y^A$ und $Y^B$.
- Anwendungsbeispiele für beide Fragestellungen:
  1. Vergleich der Montagezeiten zweier unterschiedlicher Montageverfahren auf Grundlage von Zeitmessungen beider Verfahren *für dieselbe (Stichproben-)Auswahl von Arbeitern*.
  2. Vergleich der in Eignungstests erreichten Punktzahlen von männlichen und weiblichen Bewerbern (auf Basis zweier unabhängiger einfacher Stichproben).

# $t$-Differenzentest bei verbundener Stichprobe

- Idee für Mittelwertvergleich bei verbundenen Stichproben:
  - Ein Vergleich der Mittelwerte von $Y^A$ und $Y^B$ kann anhand des Mittelwerts $\mu := \mathrm{E}(Y)$ der Differenz $Y := Y^A - Y^B$ erfolgen, denn mit $\mu_A := \mathrm{E}(Y^A)$ und $\mu_B := \mathrm{E}(Y^B)$ gilt offensichtlich $\mu = \mu_A - \mu_B$ und damit:

    $$\mu < 0 \iff \mu_A < \mu_B \qquad \mu = 0 \iff \mu_A = \mu_B \qquad \mu > 0 \iff \mu_A > \mu_B$$
  - Mit $x_1 := x_1^A - x_1^B, \ldots, x_n := x_n^A - x_n^B$ liegt eine Realisation einer einfachen Stichprobe $X_1 := X_1^A - X_1^B, \ldots, X_n := X_n^A - X_n^B$ vom Umfang $n$ zu $Y = Y^A - Y^B$ vor.
  - Darüberhinaus gilt: Ist $(Y^A, Y^B)$ gemeinsam (zweidimensional) normalverteilt, so ist auch die Differenz $Y = Y^A - Y^B$ normalverteilt.
- Es liegt also nahe, die gemeinsame Stichprobe zu $(Y^A, Y^B)$ zu „einer" Stichprobe zu $Y = Y^A - Y^B$ zusammenzufassen und den bekannten $t$-Test für den Mittelwert einer (normalverteilten) Zufallsvariablen bei unbekannter Varianz auf der Grundlage der einfachen Stichprobe $X_1, \ldots, X_n$ zu $Y$ durchzuführen.
- Prinzipiell wäre bei bekannter Varianz von $Y = Y^A - Y^B$ auch ein entsprechender Gauß-Test durchführbar; Anwendungen hierfür sind aber selten.

# Zusammenfassung: *t*-Differenzentest

| | | | |
|---|---|---|---|
| Anwendungs-voraussetzungen | exakt: $(Y^A, Y^B)$ gemeinsam (zweidimensional) normalverteilt, $\mathrm{E}(Y^A) = \mu_A, \mathrm{E}(Y^B) = \mu_B$ sowie Varianzen/Kovarianz unbekannt<br>approx.: $\mathrm{E}(Y^A) = \mu_A, \mathrm{E}(Y^B) = \mu_B, \mathrm{Var}(Y^A), \mathrm{Var}(Y^B)$ unbek.<br>$(X_1^A, X_1^B), \ldots, (X_n^A, X_n^B)$ einfache Stichprobe zu $(Y^A, Y^B)$ | | |
| Nullhypothese<br>Gegenhypothese | $H_0 : \mu_A = \mu_B$<br>$H_1 : \mu_A \neq \mu_B$ | $H_0 : \mu_A \leq \mu_B$<br>$H_1 : \mu_A > \mu_B$ | $H_0 : \mu_A \geq \mu_B$<br>$H_1 : \mu_A < \mu_B$ |
| Teststatistik | $t = \frac{\overline{X}}{S}\sqrt{n}$ | | |
| Verteilung ($H_0$) | $t$ für $\mu_A = \mu_B$ (näherungsweise) $t(n-1)$-verteilt | | |
| Benötigte Größen | $X_i = X_i^A - X_i^B$ für $i \in \{1, \ldots, n\}$, $\quad \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$<br>$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2} = \sqrt{\frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2 - n\overline{X}^2\right)}$ | | |
| Kritischer Bereich zum Niveau $\alpha$ | $(-\infty, -t_{n-1;1-\frac{\alpha}{2}}) \cup (t_{n-1;1-\frac{\alpha}{2}}, \infty)$ | $(t_{n-1;1-\alpha}, \infty)$ | $(-\infty, -t_{n-1;1-\alpha})$ |
| $p$-Wert | $2 \cdot (1 - F_{t(n-1)}(\lvert t \rvert))$ | $1 - F_{t(n-1)}(t)$ | $F_{t(n-1)}(t)$ |

# Beispiel: Montagezeiten von zwei Verfahren

- Untersuchungsgegenstand: Ist ein neu vorgeschlagenes Montageverfahren besser (im Sinne einer im Mittel kürzeren Bearbeitungsdauer $Y^B$) als das zur Zeit eingesetzte Montageverfahren (mit Bearbeitungsdauer $Y^A$)?
- Stichprobeninformation: Zeitmessungen der Montagedauern $x_i^A$ für Verfahren $A$ und $x_i^B$ für Verfahren $B$ bei **denselben** $n = 7$ Arbeitern:

| Arbeiter $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| $x_i^A$ | 64 | 71 | 68 | 66 | 73 | 62 | 70 |
| $x_i^B$ | 60 | 66 | 66 | 69 | 63 | 57 | 62 |

- Annahme: $(Y^A, Y^B)$ gemeinsam normalverteilt, $(X_1^A, X_1^B), \ldots, (X_n^A, X_n^B)$ einfache Stichprobe zu $(Y^A, Y^B)$.
- Gewünschtes Signifikanzniveau: $\alpha = 0.05$

Geeigneter Test: Exakter $t$-**Differenzentest** für verbundene Stichproben

1. **Hypothesen:**

   $H_0 : \mu_A \leq \mu_B \qquad \text{gegen} \qquad H_1 : \mu_A > \mu_B$

2. **Teststatistik:**

   $t = \frac{\overline{X}}{S}\sqrt{n}$ ist unter $H_0$ $t(n-1)$-verteilt (für $\mu_A{=}\mu_B$).

3. **Kritischer Bereich zum Niveau $\alpha = 0.05$:**

   $K = (t_{n-1;1-\alpha}, +\infty) = (t_{6;0.95}, +\infty) = (1.943, +\infty)$

4. **Berechnung der realisierten Teststatistik:**

| Arbeiter $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| $x_i^A$ | 64 | 71 | 68 | 66 | 73 | 62 | 70 |
| $x_i^B$ | 60 | 66 | 66 | 69 | 63 | 57 | 62 |
| $x_i = x_i^A - x_i^B$ | 4 | 5 | 2 | $-3$ | 10 | 5 | 8 |

   Mit $\overline{x} = \frac{1}{7}\sum_{i=1}^{7} x_i = 4.4286$ und $s = \sqrt{\frac{1}{7-1}\sum_{i=1}^{7}(x_i - \overline{x})^2} = 4.1975$:

   $$t = \frac{\overline{x}}{s}\sqrt{n} = \frac{4.4286}{4.1975}\sqrt{7} = 2.7914$$

5. **Entscheidung:**

   $t = 2.7914 \in (1.943, +\infty) = K \quad \Rightarrow \quad H_0$ wird abgelehnt!

   ($p$-Wert: $1 - F_{t(6)}(t) = 1 - F_{t(6)}(2.7914) = 1 - 0.9842 = 0.0158$)

Der Test kommt also zur Entscheidung, dass das neue Montageverfahren eine im Mittel signifikant kürzere Montagedauer aufweist.

# Mittelwertvergleiche bei zwei unabhängigen Stichproben

- Liegen zwei unabhängige Stichproben $X_1^A, \ldots, X_{n_A}^A$ und $X_1^B, \ldots, X_{n_B}^B$ zu jeweils normalverteilten Zufallsvariablen $Y^A$ und $Y^B$ vor, kann eine „Aggregation" zu einer einzigen Stichprobe wie beim Vorliegen verbundener Stichproben so nicht durchgeführt werden.
- Verglichen werden nun nicht mehr Beobachtungspaare, sondern die (getrennt) berechneten Mittelwerte $\overline{X^A}$ und $\overline{X^B}$ der beiden Stichprobenrealisationen zu $Y^A$ bzw. $Y^B$.
- Wir setzen zunächst die *Normalverteilungsannahme für $Y^A$ und $Y^B$* voraus!
- Die Differenz $\overline{X^A} - \overline{X^B}$ ist wegen der Unabhängigkeit der Stichproben dann offensichtlich normalverteilt mit Erwartungswert $\mu_A - \mu_B$ (für $\mu_A = \mu_B$ gilt also gerade $\mathrm{E}(\overline{X^A} - \overline{X^B}) = 0$) und Varianz

$$\mathrm{Var}(\overline{X^A} - \overline{X^B}) = \mathrm{Var}(\overline{X^A}) + \mathrm{Var}(\overline{X^B}) = \frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B} \ .$$

- Sind die beteiligten Varianzen bekannt, kann zum Vergleich von $\mu_A$ und $\mu_B$ somit unmittelbar ein exakter Gauß-Test konstruiert werden.

# Zusammenfassung: 2-Stichproben-Gauß-Test

bei bekannten Varianzen

| | | | |
|---|---|---|---|
| Anwendungs-<br>voraussetzungen | exakt: $Y^A \sim N(\mu_A, \sigma_A^2)$, $Y^B \sim N(\mu_B, \sigma_B^2)$, $\sigma_A^2, \sigma_B^2$ bekannt<br>$X_1^A, \ldots, X_{n_A}^A$ einfache Stichprobe zu $Y^A$, unabhängig von einfacher Stichprobe $X_1^B, \ldots, X_{n_B}^B$ zu $Y^B$. | | |
| Nullhypothese<br>Gegenhypothese | $H_0 : \mu_A = \mu_B$<br>$H_1 : \mu_A \neq \mu_B$ | $H_0 : \mu_A \leq \mu_B$<br>$H_1 : \mu_A > \mu_B$ | $H_0 : \mu_A \geq \mu_B$<br>$H_1 : \mu_A < \mu_B$ |
| Teststatistik<br><br>Verteilung ($H_0$) | $N = \dfrac{\overline{X^A} - \overline{X^B}}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}}$<br>$N$ für $\mu_A = \mu_B$ $N(0,1)$-verteilt | | |
| Benötigte Größen | $\overline{X^A} = \frac{1}{n_A} \sum_{i=1}^{n_A} X_i^A, \quad \overline{X^B} = \frac{1}{n_B} \sum_{i=1}^{n_B} X_i^B$ | | |
| Kritischer Bereich<br>zum Niveau $\alpha$ | $(-\infty, -N_{1-\frac{\alpha}{2}})$<br>$\cup(N_{1-\frac{\alpha}{2}}, \infty)$ | $(N_{1-\alpha}, \infty)$ | $(-\infty, -N_{1-\alpha})$ |
| $p$-Wert | $2 \cdot (1 - \Phi(\|N\|))$ | $1 - \Phi(N)$ | $\Phi(N)$ |

- Sind die Varianzen $\sigma_A^2$ und $\sigma_B^2$ unbekannt, so ist zu unterscheiden, ob man wenigstens $\sigma_A^2 = \sigma_B^2$ annehmen kann oder nicht.
- Im Fall übereinstimmender Varianzen $\sigma_A^2 = \sigma_B^2$ wird diese mit Hilfe eines gewichteten Mittelwerts $S^2$ der Stichprobenvarianzen

$$S_{Y^A}^2 = \frac{1}{n_A - 1}\sum_{i=1}^{n_A}(X_i^A - \overline{X^A})^2 \qquad \text{und} \qquad S_{Y^B}^2 = \frac{1}{n_B - 1}\sum_{j=1}^{n_B}(X_j^B - \overline{X^B})^2$$

in der Form

$$S^2 = \frac{(n_A - 1)S_{Y^A}^2 + (n_B - 1)S_{Y^B}^2}{n_A + n_B - 2} = \frac{\sum_{i=1}^{n_A}(X_i^A - \overline{X^A})^2 + \sum_{j=1}^{n_B}(X_j^B - \overline{X^B})^2}{n_A + n_B - 2}$$

geschätzt, ein exakter $t$-Test ist damit konstruierbar.
- Für $n_A = n_B$ erhält man die einfachere Darstellung $S^2 = \frac{S_{Y^A}^2 + S_{Y^B}^2}{2}$.

# Zusammenfassung: 2-Stichproben-$t$-Test

bei unbekannten, aber übereinstimmenden Varianzen

| | | | |
|---|---|---|---|
| Anwendungs-voraussetzungen | exakt: $Y^A \sim N(\mu_A, \sigma_A^2)$, $Y^B \sim N(\mu_B, \sigma_B^2), \mu_A, \mu_B, \sigma_A^2 = \sigma_B^2$ unbek.<br>approx.: $\text{E}(Y^A) = \mu_A, \text{E}(Y^B) = \mu_B, \text{Var}(Y^A) = \text{Var}(Y^B)$ unbekannt<br>$X_1^A, \ldots, X_{n_A}^A$ einfache Stichprobe zu $Y^A$, unabhängig von einfacher Stichprobe $X_1^B, \ldots, X_{n_B}^B$ zu $Y^B$. | | |
| Nullhypothese<br>Gegenhypothese | $H_0 : \mu_A = \mu_B$<br>$H_1 : \mu_A \neq \mu_B$ | $H_0 : \mu_A \leq \mu_B$<br>$H_1 : \mu_A > \mu_B$ | $H_0 : \mu_A \geq \mu_B$<br>$H_1 : \mu_A < \mu_B$ |
| Teststatistik | $t = \dfrac{\overline{X^A} - \overline{X^B}}{\sqrt{\frac{S^2}{n_A} + \frac{S^2}{n_B}}} = \dfrac{\overline{X^A} - \overline{X^B}}{S}\sqrt{\dfrac{n_A \cdot n_B}{n_A + n_B}}$ | | |
| Verteilung ($H_0$) | $t$ für $\mu_A = \mu_B$ (näherungsweise) $t(n_A + n_B - 2)$-verteilt | | |
| Benötigte Größen | $\overline{X^A} = \frac{1}{n_A}\sum_{i=1}^{n_A} X_i^A, \quad \overline{X^B} = \frac{1}{n_B}\sum_{i=1}^{n_B} X_i^B,$<br>$S = \sqrt{\frac{(n_A-1)S_{Y^A}^2+(n_B-1)S_{Y^B}^2}{n_A+n_B-2}} = \sqrt{\frac{\sum_{i=1}^{n_A}(X_i^A-\overline{X^A})^2+\sum_{i=1}^{n_B}(X_i^B-\overline{X^B})^2}{n_A+n_B-2}}$ | | |
| Kritischer Bereich zum Niveau $\alpha$ | $(-\infty, -t_{n_A+n_B-2;1-\frac{\alpha}{2}})$<br>$\cup(t_{n_A+n_B-2;1-\frac{\alpha}{2}}, \infty)$ | $(t_{n_A+n_B-2;1-\alpha}, \infty)$ | $(-\infty, -t_{n_A+n_B-2;1-\alpha})$ |
| $p$-Wert | $2 \cdot (1 - F_{t(n_A+n_B-2)}(\|t\|))$ | $1 - F_{t(n_A+n_B-2)}(t)$ | $F_{t(n_A+n_B-2)}(t)$ |

# Beispiel: Absatzwirkung einer Werbeaktion

- Untersuchungsgegenstand: Hat eine spezielle Sonderwerbeaktion positiven Einfluss auf den mittleren Absatz?
- Stichprobeninformation: Messung der prozentualen Absatzänderungen $x_1^A, \ldots, x_{10}^A$ in $n_A = 10$ Supermärkten **ohne** Sonderwerbeaktion und $x_1^B, \ldots, x_5^B$ in $n_B = 5$ Supermärkten **mit** Sonderwerbeaktion.
- Annahme: Für prozentuale Absatzänderungen $Y^A$ ohne bzw. $Y^B$ mit Sonderwerbeaktion gilt $Y^A \sim N(\mu_A, \sigma_A^2)$, $Y^B \sim N(\mu_B, \sigma_B^2)$, $\mu_A, \mu_B, \sigma_A^2 = \sigma_B^2$ unbekannt, $X_1^A, \ldots, X_{10}^A$ einfache Stichprobe zu $Y^A$, unabhängig von einfacher Stichprobe $X_1^B, \ldots, X_5^B$ zu $Y^B$.
- (Zwischen-)Ergebnisse aus Stichprobenrealisation:

$$\overline{x^A} = 6.5, \quad \overline{x^B} = 8, \quad s_{Y^A}^2 = 20.25, \quad s_{Y^B}^2 = 23.04$$

$$\Rightarrow s = \sqrt{\frac{(n_A - 1)s_{Y^A}^2 + (n_B - 1)s_{Y^B}^2}{n_A + n_B - 2}} = \sqrt{\frac{9 \cdot 20.25 + 4 \cdot 23.04}{13}} = 4.5944$$

- Gewünschtes Signifikanzniveau: $\alpha = 0.05$

Geeigneter Test:

**2-Stichproben-$t$-Test bei übereinstimmenden, aber unbekannten Varianzen**

1. **Hypothesen:**

   $$H_0 : \mu_A \geq \mu_B \qquad \text{gegen} \qquad H_1 : \mu_A < \mu_B$$

2. **Teststatistik:**

   $$t = \frac{\overline{X^A} - \overline{X^B}}{S}\sqrt{\frac{n_A \cdot n_B}{n_A + n_B}} \text{ ist unter } H_0 \ t(n_A + n_B - 2)\text{-verteilt (für } \mu_A = \mu_B).$$

3. **Kritischer Bereich zum Niveau $\alpha = 0.05$:**

   $$K = (-\infty, -t_{n_A+n_B-2;1-\alpha}) = (-\infty, -t_{13;0.95}) = (-\infty, -1.771)$$

4. **Berechnung der realisierten Teststatistik:**

   $$t = \frac{\overline{x^A} - \overline{x^B}}{s}\sqrt{\frac{n_A \cdot n_B}{n_A + n_B}} = \frac{6.5 - 8}{4.5944}\sqrt{\frac{10 \cdot 5}{10 + 5}} = -0.5961$$

5. **Entscheidung:**

   $$t = -0.5961 \notin (-\infty, -1.771) = K \quad \Rightarrow \quad H_0 \text{ wird nicht abgelehnt!}$$

   ($p$-Wert: $F_{t(13)}(t) = F_{t(13)}(-0.5961) = 0.2807$)

Der Test kommt also zur Entscheidung, dass eine positive Auswirkung der Sonderwerbeaktion auf die mittlere prozentuale Absatzänderung nicht bestätigt werden kann.

# Sonderfall: Vergleich von Anteilswerten

- Ein Sonderfall des (approximativen) 2-Stichproben-$t$-Test bei unbekannten, aber übereinstimmenden Varianzen liegt vor, wenn zwei Anteilswerte miteinander verglichen werden sollen.
- Es gelte also speziell $Y^A \sim B(1, p_A)$ und $Y^B \sim B(1, p_B)$ für $p_A \in (0,1)$ und $p_B \in (0,1)$, außerdem seien $X_1^A, \ldots, X_{n_A}^A$ sowie $X_1^B, \ldots, X_{n_B}^B$ unabhängige einfache Stichproben vom Umfang $n_A$ zu $Y^A$ bzw. vom Umfang $n_B$ zu $Y^B$.
- Zur Überprüfung stehen die Hypothesenpaare:

| | | | |
|---|---|---|---|
| | $H_0 : p_A = p_B$ | $H_0 : p_A \leq p_B$ | $H_0 : p_A \geq p_B$ |
| gegen | $H_1 : p_A \neq p_B$ | $H_1 : p_A > p_B$ | $H_1 : p_A < p_B$ |

- Für die Varianzen von $Y^A$ und $Y^B$ gilt bekanntlich $\text{Var}(Y^A) = p_A \cdot (1 - p_A)$ bzw. $\text{Var}(Y^B) = p_B \cdot (1 - p_B)$, d.h. die Varianzen sind zwar unbekannt, unter $H_0$ — genauer für $p_A = p_B$ — jedoch gleich.
- Mit den üblichen Schreibweisen $\widehat{p}_A := \frac{1}{n_A} \sum_{i=1}^{n_A} X_i^A$ bzw. $\widehat{p}_B := \frac{1}{n_B} \sum_{i=1}^{n_B} X_i^B$ erhält man für $S^2$ in Abhängigkeit von $\widehat{p}_A$ und $\widehat{p}_B$ die Darstellung:

$$S^2 = \frac{n_A \cdot \widehat{p}_A \cdot (1 - \widehat{p}_A) + n_B \cdot \widehat{p}_B \cdot (1 - \widehat{p}_B)}{n_A + n_B - 2}$$

- Approximation vernünftig, falls $5 \leq n_A \widehat{p}_A \leq n_A - 5$ und $5 \leq n_B \widehat{p}_B \leq n_B - 5$ .

# Zusammenfassung: 2-Stichproben-$t$-Test für Anteilswerte

| | | | |
|---|---|---|---|
| Anwendungs- voraussetzungen | approx.: $Y^A \sim B(1, p_A)$, $Y^B \sim B(1, p_B)$, $p_A, p_B$ unbekannt $X_1^A, \ldots, X_{n_A}^A$ einfache Stichprobe zu $Y^A$, unabhängig von einfacher Stichprobe $X_1^B, \ldots, X_{n_B}^B$ zu $Y^B$. | | |
| Nullhypothese Gegenhypothese | $H_0 : p_A = p_B$ $H_1 : p_A \neq p_B$ | $H_0 : p_A \leq p_B$ $H_1 : p_A > p_B$ | $H_0 : p_A \geq p_B$ $H_1 : p_A < p_B$ |
| Teststatistik | $t = \frac{\widehat{p}_A - \widehat{p}_B}{\sqrt{\frac{S^2}{n_A} + \frac{S^2}{n_B}}} = \frac{\widehat{p}_A - \widehat{p}_B}{S}\sqrt{\frac{n_A \cdot n_B}{n_A + n_B}}$ | | |
| Verteilung ($H_0$) | $t$ für $p_A = p_B$ näherungsweise $t(n_A + n_B - 2)$-verteilt (Näherung ok, falls $5 \leq n_A\widehat{p}_A \leq n_A - 5$ und $5 \leq n_B\widehat{p}_B \leq n_B - 5$) | | |
| Benötigte Größen | $\widehat{p}_A = \frac{1}{n_A}\sum_{i=1}^{n_A} X_i^A, \quad \widehat{p}_B = \frac{1}{n_B}\sum_{i=1}^{n_B} X_i^B,$ $S = \sqrt{\frac{n_A \cdot \widehat{p}_A \cdot (1-\widehat{p}_A) + n_B \cdot \widehat{p}_B \cdot (1-\widehat{p}_B)}{n_A + n_B - 2}}$ | | |
| Kritischer Bereich zum Niveau $\alpha$ | $(-\infty, -t_{n_A+n_B-2;1-\frac{\alpha}{2}}) \cup (t_{n_A+n_B-2;1-\frac{\alpha}{2}}, \infty)$ | $(t_{n_A+n_B-2;1-\alpha}, \infty)$ | $(-\infty, -t_{n_A+n_B-2;1-\alpha})$ |
| $p$-Wert | $2 \cdot (1 - F_{t(n_A+n_B-2)}(\lvert t \rvert))$ | $1 - F_{t(n_A+n_B-2)}(t)$ | $F_{t(n_A+n_B-2)}(t)$ |

# Beispiel: Vergleich von zwei Fehlerquoten

mit approximativem 2-Stichproben-$t$-Test für Anteilswerte

- Untersuchungsgegenstand: Vergleich von Fehlerquoten zweier Sortiermaschinen
- Für einen automatisierten Sortiervorgang werden eine günstige ($A$) sowie eine hochpreisige Maschine ($B$) angeboten. Es soll anhand von 2 (unabhängigen) Testläufen mit jeweils $n_A = n_B = 1000$ Sortiervorgängen überprüft werden, ob die Fehlerquote $p_A$ bei der günstigen Maschine $A$ höher ist als die Fehlerquote $p_B$ der hochpreisigen Maschine $B$.
- Resultat der Testläufe soll jeweils als Realisation einer einfachen Stichprobe aufgefasst werden können.
- Stichprobeninformation: Bei Maschine $A$ traten 29 Fehler auf, bei Maschine $B$ 21 Fehler.
- (Zwischen-) Ergebnisse aus Stichprobenrealisation: $\widehat{p}_A = \frac{29}{1000} = 0.029$, $\widehat{p}_B = \frac{21}{1000} = 0.021$, $s = \sqrt{\frac{1000 \cdot 0.029 \cdot (1-0.029) + 1000 \cdot 0.021 \cdot (1-0.021)}{1000+1000-2}} = 0.156$
- Gewünschtes Signifikanzniveau $\alpha = 0.05$.

1. **Hypothesen:**

   $$H_0 : p_A \leq p_B \qquad \text{gegen} \qquad H_1 : p_A > p_B$$

2. **Teststatistik:**

   $t = \frac{\widehat{p}_A - \widehat{p}_B}{S}\sqrt{\frac{n_A \cdot n_B}{n_A + n_B}}$ ist unter $H_0$ näherungsweise $t(n_A + n_B - 2)$-verteilt (für $p_A = p_B$). Näherung ok, da $5 \leq 29 \leq 995$ und $5 \leq 21 \leq 995$.

3. **Kritischer Bereich zum Niveau $\alpha = 0.05$:**

   $$K = (t_{n_A+n_B-2;1-\alpha}, +\infty) = (t_{1998;0.95}, +\infty) = (1.646, +\infty)$$

4. **Berechnung der realisierten Teststatistik:**

   $$t = \frac{\widehat{p}_A - \widehat{p}_B}{s}\sqrt{\frac{n_A \cdot n_B}{n_A + n_B}} = \frac{0.029 - 0.021}{0.1562}\sqrt{\frac{1000 \cdot 1000}{1000 + 1000}} = 1.1452$$

5. **Entscheidung:**

   $$t = 1.1452 \notin (1.646, +\infty) = K \quad \Rightarrow \quad H_0 \text{ wird nicht abgelehnt!}$$

   ($p$-Wert: $1 - F_{t(1998)}(t) = 1 - F_{t(1998)}(1.1452) = 1 - 0.8739 = 0.1261$)

Der Test kommt also zum Ergebnis, dass eine höhere Fehlerquote der günstigen Maschine nicht bestätigt werden kann.

# Approximativer 2-Stichproben-Gauß-Test

für Mittelwertvergleiche, wenn Gleichheit der Varianzen ungewiss

- Kann in der Situation des exakten 2-Stichproben-$t$-Test ($Y^A$ und $Y^B$ sind normalverteilt mit unbekannten Varianzen) auch unter $H_0$ keine Gleichheit der Varianzen vorausgesetzt werden, müssen andere Testverfahren verwendet werden, z.B. der **Welch-Test** (hier nicht besprochen).
- Als approximativer Test lässt sich (zumindest bei hinreichend großen Stichprobenumfängen, „Daumenregel" $n_A > 30$ und $n_B > 30$) auch eine leichte Modifikation des 2-Stichproben-Gauß-Tests aus Folie 187 verwenden.
- Anstelle der (dort als bekannt vorausgesetzten) Varianzen $\sigma_A^2$ und $\sigma_B^2$ sind die erwartungstreuen Schätzfunktionen $S_{Y^A}^2$ und $S_{Y^B}^2$ einzusetzen und der Test als approximativer Test durchzuführen.
- Die Teststatistik nimmt damit die Gestalt

$$N = \frac{\overline{X^A} - \overline{X^B}}{\sqrt{\frac{S_{Y^A}^2}{n_A} + \frac{S_{Y^B}^2}{n_B}}}$$

an und ist unter $H_0$ näherungsweise standardnormalverteilt.