Mathtrace – Handbuch

V. Statistik

Inhalt

V. Statistik

41. Statistik

41. Statistik

Daten schreiben Sie als data(…): die Werte durch Semikolon, Leerzeichen oder Zeilenumbruch getrennt, ein Dezimalkomma gilt wie sonst (data(1,5 2,5 3)). Eine Tabelle – eine Zeile pro Beobachtung, eine Spalte pro Variable – Zeile für Zeile durch Semikolon getrennt; die erste Zeile darf die Namen der Variablen enthalten. Bequemer: die Gittertaste auf dem Reiter f(x) öffnet eine Tabelle zum Ausfüllen, Einfügen aus der Zwischenablage (Zellen aus Excel oder Google Tabellen) oder Laden aus einer CSV-Datei, und die Schaltfläche „Als Daten“ fügt sie für die Statistik ein – auch eine Tabelle, die größer als das Gitter ist.

Bei einer Variablen wählt ein Chip, was berechnet wird. Deskriptive Statistik: arithmetisches Mittel, Median, Modalwert, Quartile (Mediane der unteren und oberen Hälfte), Spannweite und Interquartilsabstand, Varianz und Standardabweichung der Stichprobe (n − 1) und der Population (n), Variationskoeffizient und Ausreißer (außerhalb $Q_1 - 1{,}5 \cdot IQR$ und $Q_3 + 1{,}5 \cdot IQR$), mit Boxplot. Häufigkeiten und Histogramm: die Tabelle der absoluten, relativen und kumulierten Häufigkeiten – bei ganzen Zahlen mit wenigen Werten nach Werten, sonst nach Klassen (ihre Anzahl nach der Sturges-Regel). z-Werte: jeder Wert mit Mittelwert und Standardabweichung der Daten standardisiert, mit T-Wert.

Zwei Variablen (zwei Spalten) ergeben Korrelation und Regression: Summen, Kovarianz, den Korrelationskoeffizienten r nach Pearson mit der Stärke des Zusammenhangs in Worten, die Regressionsgerade nach der Methode der kleinsten Quadrate und das Bestimmtheitsmaß r^2, mit Streudiagramm. Drei und mehr Variablen erhalten ihre deskriptive Statistik nebeneinander mit Boxplots.

Schließende Statistik aus Daten: das Konfidenzintervall des Mittelwerts (Chip „Konfidenzintervall des Mittelwerts“, 90, 95 oder 99 %) über die t-Verteilung; der t-Test (Chip „t-Test“) – eine Variable gegen den Testwert μ_0, zwei Variablen als unabhängige Stichproben (Welch-Test oder mit gleichen Varianzen) oder als gepaarte Werte (Differenzen der Paare); der χ^2-Test (Chip „χ^2-Test“) – eine Spalte mit Häufigkeiten als Anpassungstest an die Gleichverteilung, eine Tafel als Unabhängigkeitstest mit Cramérs V. Unter den Chips geben Sie das Signifikanzniveau α, die Alternative (zweiseitig, linksseitig, rechtsseitig) und μ_0 ein. Der Lösungsweg: Hypothesen, Teststatistik, kritischer Wert und Bereich, p-Wert und Entscheidung; das Diagramm zeigt die Verteilung mit schattiertem kritischem Bereich. Zwei unabhängige Stichproben verschiedenen Umfangs schreiben Sie als benannte Spalten – data(A = 23 25 28; B = 20 22 21 25).

Gruppen vergleichen: Der Chip „Rangtest (Mann-Whitney, Wilcoxon, Kruskal-Wallis)“ vergleicht zwei Gruppen mit dem Mann-Whitney-Test (mit „Gepaart“ gepaarte Werte mit dem Wilcoxon-Test) und drei und mehr Gruppen mit dem Kruskal-Wallis-Test – über die Ränge der Werte, ohne Normalverteilung vorauszusetzen; bei kleinen Stichproben ohne Bindungen ist der p-Wert exakt. Der Chip „ANOVA (Varianzanalyse)“ vergleicht die Mittelwerte der Gruppen mit der einfaktoriellen Varianzanalyse – ANOVA-Tabelle, F-Test und Effektstärke η^2. Die Gruppen kommen in Spalten, beliebig lang (benannte Spalten); das Diagramm zeigt die Boxplots der Gruppen und die Verteilung der Prüfgröße. Beim χ^2-Test

einer Spalte mit Häufigkeiten nimmt das Feld „Erwartete Anteile“ eigene erwartete Anteile – 1:2:1, 0,25 0,5 0,25 oder in Prozent.

Die Regression zweier Variablen hat unter ihrem Chip die Wahl des Modells: linear (mit Signifikanztest der Korrelation), quadratisch (Normalgleichungen der Methode der kleinsten Quadrate), exponentiell y = a·e^(bx), Potenz y = a·x^b und logarithmisch y = a + b·ln x (durch Linearisierung über Logarithmen) – mit dem Bestimmtheitsmaß R^2 und einem Diagramm der angepassten Kurve. Im Formular „Statistik und Psychometrie“ gibt es zusätzlich Konfidenzintervall und t-Tests aus zusammengefassten Werten (Umfang, Mittelwert, Standardabweichung) – auch den z-Test bei bekanntem σ und den t-Test für zwei Stichproben.

Die Normalverteilung finden Sie im Formular „Statistik und Psychometrie“ (Abschnitt „Anwendungen“ auf dem Startbildschirm, Fachgebiet Statistik): die Wahrscheinlichkeiten P(X < x), P(X > x), P(a < X < b) und das Quantil – der Wert, unter dem ein gegebener Anteil liegt. Der Lösungsweg standardisiert, z = (x − μ)/σ, und nutzt die Verteilungsfunktion Φ, ein negatives z als 1 − Φ(|z|) wie in den Tabellen; der Graph zeigt die Dichte mit der gefärbten Fläche.

Eingabe

`data(3; 5; 7; 7; 9)`	Eine Variable – die Werte durch Semikolon, Leerzeichen oder untereinander.
`data(x y; 1 52; 2 58; 3 61)`	Eine Tabelle zeilenweise, die erste Zeile mit den Namen der Variablen (optional).
`data(x = 1 2 3 4; y = 2 4 5 4)`	Benannte Spalten, jede mit ihren Werten.

Beispiele

Deskriptive Statistik von acht Werten.

Eingabe: `data(2; 4; 4; 4; 5; 5; 7; 9)`

Ergebnis

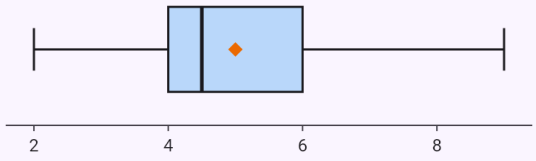
$$\bar{x} = 5$$
$$\tilde{x} = 4{,}5$$
$$s \approx 2{,}138$$

Lösungsweg

- Anzahl der Werte $n = 8$, sortierte Daten:
  $2,\ 4,\ 4,\ 4,\ 5,\ 5,\ 7,\ 9$
- Arithmetisches Mittel: $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{40}{8} = 5$
- Der Median ist der mittlere Wert der sortierten Daten:
  $\tilde{x} = \frac{x_{(4)} + x_{(5)}}{2} = \frac{4+5}{2} = 4{,}5$
- Modalwert (häufigster Wert, 3×): $\hat{x} = 4$
- Die Quartile sind die Mediane der unteren und der oberen Hälfte: $Q_1 = 4, \quad Q_2 = \tilde{x} = 4{,}5, \quad Q_3 = 6$
- Spannweite und Interquartilsabstand
  $R = x_{max} - x_{min} = 9 - 2 = 7$
  $IQR = Q_3 - Q_1 = 6 - 4 = 2$
- Summe der quadrierten Abweichungen vom Mittelwert:
  $\sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum x_i^2 - n\bar{x}^2 = 32$
- Varianz – Stichprobe s² und Population σ²
  $s^2 = \frac{1}{n-1}\sum (x_i - \bar{x})^2 = \frac{32}{7} \approx 4{,}571$
  $\sigma^2 = \frac{1}{n}\sum (x_i - \bar{x})^2 = \frac{32}{8} = 4$
- Standardabweichung
  $s = \sqrt{s^2} = \sqrt{4{,}571} \approx 2{,}138$
  $\sigma = \sqrt{4} = 2$
- Variationskoeffizient:
  $v = \frac{s}{\bar{x}} = \frac{2{,}138}{5} \approx 0{,}4276 = 42{,}8\,\%$
- Die Zäune für Ausreißer
  $Q_1 - 1{,}5\,IQR = 1, \quad Q_3 + 1{,}5\,IQR = 9$ – kein Wert liegt außerhalb.
- Die Stichprobenvarianz s² teilt durch n − 1 und schätzt die Varianz der Population – die übliche Wahl. Die Populationsvarianz σ² teilt durch n, für Daten der ganzen Population.
- Achtung: wenige Daten, die Schätzungen (vor allem Varianz und Quartile) sind unsicher.

Graph

Boxplot (Median, Quartile, Ausreißer; ◆ Mittelwert)



Der Ausreißer 31 im Boxplot.

Eingabe: `data(12 15 14 13 16 15 14 31)`

Ergebnis

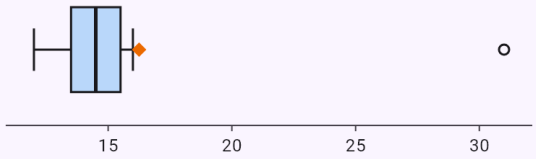
$$\bar{x} = 16{,}25$$
$$\tilde{x} = 14{,}5$$
$$s \approx 6{,}089$$

Lösungsweg

- Anzahl der Werte $n = 8$, sortierte Daten:
  $12,\ 13,\ 14,\ 14,\ 15,\ 15,\ 16,\ 31$
- Arithmetisches Mittel: $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{130}{8} = 16{,}25$
- Der Median ist der mittlere Wert der sortierten Daten:
  $\tilde{x} = \frac{x_{(4)} + x_{(5)}}{2} = \frac{14 + 15}{2} = 14{,}5$
- Modalwert (häufigster Wert, 2×): $\hat{x} = 14,\ 15$
- Die Quartile sind die Mediane der unteren und der oberen Hälfte: $Q_1 = 13{,}5, \quad Q_2 = \tilde{x} = 14{,}5, \quad Q_3 = 15{,}5$
- Spannweite und Interquartilsabstand
  $R = x_{max} - x_{min} = 31 - 12 = 19$
  $IQR = Q_3 - Q_1 = 15{,}5 - 13{,}5 = 2$
- Summe der quadrierten Abweichungen vom Mittelwert:
  $\sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum x_i^2 - n\bar{x}^2 = 259{,}5$
- Varianz – Stichprobe s² und Population σ²
  $s^2 = \frac{1}{n-1}\sum (x_i - \bar{x})^2 = \frac{259{,}5}{7} \approx 37{,}07$
  $\sigma^2 = \frac{1}{n}\sum (x_i - \bar{x})^2 = \frac{259{,}5}{8} \approx 32{,}44$
- Standardabweichung
  $s = \sqrt{s^2} = \sqrt{37{,}07} \approx 6{,}089$
  $\sigma = \sqrt{32{,}44} \approx 5{,}695$
- Variationskoeffizient:
  $v = \frac{s}{\bar{x}} = \frac{6{,}089}{16{,}25} \approx 0{,}3747 = 37{,}5\,\%$
- Ausreißer liegen außerhalb der Zäune
  $Q_1 - 1{,}5\,IQR = 10{,}5, \quad Q_3 + 1{,}5\,IQR = 18{,}5$: 31
- Die Stichprobenvarianz s² teilt durch n − 1 und schätzt die Varianz der Population – die übliche Wahl. Die Populationsvarianz σ² teilt durch n, für Daten der ganzen Population.
- Achtung: wenige Daten, die Schätzungen (vor allem Varianz und Quartile) sind unsicher.

Graph

Boxplot (Median, Quartile, Ausreißer; ◆ Mittelwert)



Häufigkeitstabelle von Noten und ihr Histogramm.

Eingabe: `data(1 2 2 3 3 3 3 4 4 5 1 3 2 4 3)` Häufigkeiten und Histogramm

Ergebnis

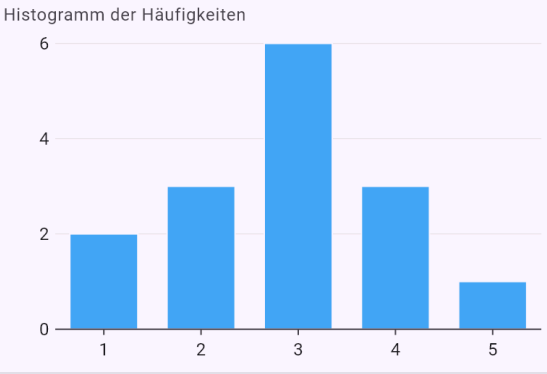
$$n = 15$$
$$\hat{x} = 3$$

Lösungsweg

- Anzahl der Werte $n = 15$, sortierte Daten:
  $1,\ 1,\ 2,\ 2,\ 2,\ 3,\ 3,\ 3,\ 3,\ 3,\ 3,\ 4,\ 4,\ 4,\ 5$
- Häufigkeitstabelle – absolute n, relative f, kumulierte N und F

x_i	n_i	f_i	N_i	F_i
1	2	0,133	2	0,133
2	3	0,200	5	0,333
3	6	0,400	11	0,733
4	3	0,200	14	0,933
5	1	0,067	15	1,000
$\sum$	15	1		

Graph



Stetige Daten in Klassen eingeteilt.

Eingabe: `data(12.5 14.1 15.8 17.2 18.9 21.3 22.0 23.7 25.1 26.4 28.8 31.2)`
Häufigkeiten und Histogramm

Ergebnis

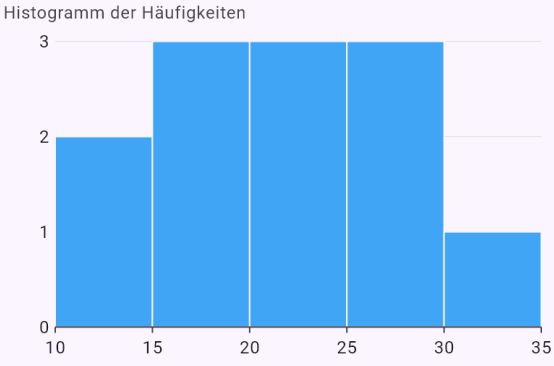
$$n = 12$$
$$k = 5, \quad h = 5$$

Lösungsweg

- Anzahl der Werte $n = 12$, sortierte Daten
  $12{,}5,\ 14{,}1,\ 15{,}8,\ 17{,}2,\ 18{,}9,\ 21{,}3,\ 22,\ 23{,}7,\ 25{,}1,\ 26,$
- Es gibt viele Werte – wir teilen sie in 5 Klassen (Sturges-Regel)
  $k \approx 1 + \log_2 n = 1 + \log_2 12 \approx 4{,}585 \Rightarrow k = 5$
  $h \approx \frac{R}{k} = \frac{18{,}7}{5} \approx 3{,}74 \Rightarrow h = 5$
- Häufigkeitstabelle der Klassen (x die Klassenmitte)

I_j	x_j	n_j	f_j	N_j	F_j
$[10;\ 15)$	12,5	2	0,167	2	0,167
$[15;\ 20)$	17,5	3	0,250	5	0,417
$[20;\ 25)$	22,5	3	0,250	8	0,667
$[25;\ 30)$	27,5	3	0,250	11	0,917
$[30;\ 35]$	32,5	1	0,083	12	1,000
$\sum$		12	1		

Graph



z-Werte und T-Werte der Daten.

Eingabe: `data(2 4 4 4 5 5 7 9)` z-Werte

Ergebnis

$$\bar{x} = 5$$
$$s \approx 2{,}138$$

Lösungsweg

- Anzahl der Werte $n = 8$, sortierte Daten:
  $2,\ 4,\ 4,\ 4,\ 5,\ 5,\ 7,\ 9$
- Arithmetisches Mittel: $\bar{x} = \dfrac{40}{8} = 5$
- Standardabweichung
  $$s = \sqrt{\frac{1}{n-1}\sum (x_i - \bar{x})^2} = \sqrt{\frac{32}{7}} \approx 2{,}138$$
- Jeder Wert wird standardisiert
  $$z_i = \frac{x_i - \bar{x}}{s} = \frac{x_i - 5}{2{,}138}$$
  $$T_i = 50 + 10 z_i$$
- z-Werte und T-Werte

x_i	z_i	T_i
2	$-1{,}40$	$36{,}0$
4	$-0{,}47$	$45{,}3$
4	$-0{,}47$	$45{,}3$
4	$-0{,}47$	$45{,}3$
5	$0{,}00$	$50{,}0$
5	$0{,}00$	$50{,}0$
7	$0{,}94$	$59{,}4$
9	$1{,}87$	$68{,}7$

- Die z-Werte nutzen Mittelwert und Standardabweichung dieser Daten. Zum Vergleich mit einer Population nimmt man Mittelwert und SD der Norm (Modell „Umrechnung von Standardwerten“).

Korrelation und Regressionsgerade – Lernstunden und Testpunkte.

Eingabe: `data(x y; 1 52; 2 58; 3 61; 4 70; 5 74; 6 79)`

Ergebnis

$$r \approx 0{,}9937$$

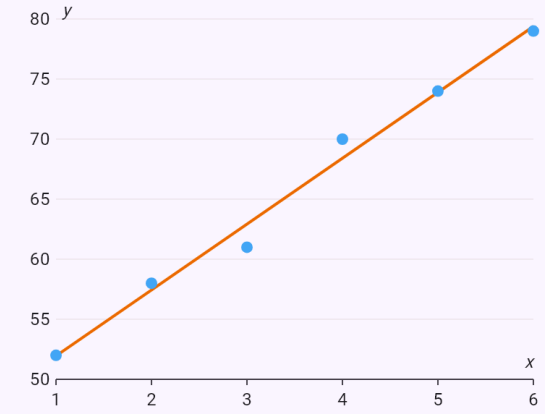
$$\hat{y} = 46{,}47 + 5{,}486x$$

Lösungsweg

- Anzahl der Werte $n = 6$, sortierte Daten
  $[1;\,52],\ [2;\,58],\ [3;\,61],\ [4;\,70],\ [5;\,74],\ [6;\,79]$
- Summen und Mittelwerte
  $\sum x_i = 21$
  $\sum y_i = 394$
  $\sum x_i^2 = 91$
  $\sum y_i^2 = 26\,410$
  $\sum x_i y_i = 1475$
  $\bar{x} = 3{,}5$
  $\bar{y} = 65{,}67$
- Standardabweichung
  $$s_x = \sqrt{\frac{\sum x_i^2 - n\bar{x}^2}{n-1}} \approx 1{,}871$$
  $$s_y = \sqrt{\frac{\sum y_i^2 - n\bar{y}^2}{n-1}} \approx 10{,}33$$
- Kovarianz
  $$s_{xy} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{n-1} = \frac{1475 - 6 \cdot 3{,}5 \cdot 65{,}67}{5} = 19{,}2$$
- Korrelationskoeffizient nach Pearson:
  $r = \frac{s_{xy}}{s_x s_y} = \frac{19{,}2}{1{,}871 \cdot 10{,}33} \approx 0{,}9937$ – der Zusammenhang ist sehr stark, positiv (mit wachsendem x wächst y).
- Koeffizienten der Regressionsgeraden nach der Methode der kleinsten Quadrate
  $b = \frac{s_{xy}}{s_x^2} = \frac{19{,}2}{3{,}5} \approx 5{,}486$
  $a = \bar{y} - b\bar{x} = 65{,}67 - 5{,}486 \cdot 3{,}5 \approx 46{,}47$
- Regressionsgerade: $\hat{y} = 46{,}47 + 5{,}486x$
- Bestimmtheitsmaß $r^2 = 0{,}9937^2 \approx 0{,}9874$ – die Gerade erklärt $98{,}7\,\%$ der Varianz von y.
- Test der Nullhypothese ρ = 0 (t-Verteilung mit 4 Freiheitsgraden) – die Korrelation ist statistisch signifikant.
  $$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0{,}9937\sqrt{4}}{\sqrt{1-0{,}9937^2}} \approx 17{,}73$$
  $p < 0{,}0001 \le \alpha = 0{,}05$
- Die Korrelation beschreibt nur einen linearen Zusammenhang: sie sagt nichts über Ursachen und kann nichtlineare übersehen. Die Regressionsgerade nur im Bereich der gemessenen x verwenden.

Graph

Streudiagramm mit Regressionsgerade



Drei Variablen nebeneinander.

Eingabe: `data(A B C; 12 15 9; 14 18 11; 11 16 10; 15 19 12; 13 17 10)`

Ergebnis

$$\bar{x}_{\mathrm{A}} = 13$$
$$\bar{x}_{\mathrm{B}} = 17$$
$$\bar{x}_{\mathrm{C}} = 10{,}4$$

Lösungsweg

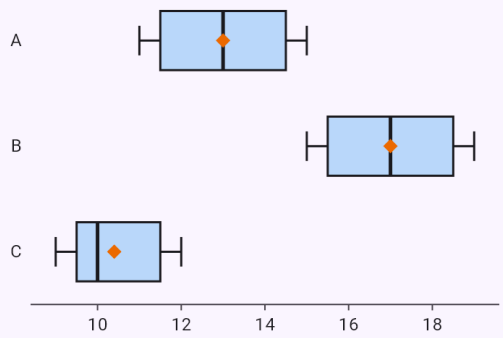
- Deskriptive Statistik der Variablen

	A	B	C
n	5	5	5
$\bar{x}$	13	17	10,4
$\tilde{x}$	13	17	10
Q_1	11,5	15,5	9,5
Q_3	14,5	18,5	11,5
s	1,581	1,581	1,14
σ	1,414	1,414	1,02
x_{min}	11	15	9
x_{max}	15	19	12

- Die Stichprobenvarianz s² teilt durch n − 1 und schätzt die Varianz der Population – die übliche Wahl. Die Populationsvarianz σ² teilt durch n, für Daten der ganzen Population.

Graph

Boxplot (Median, Quartile, Ausreißer; ◆ Mittelwert)



95-%-Konfidenzintervall des Mittelwerts von zehn Messungen.

Eingabe: `data(102 98 110 105 99 107 101 104 96 108)` Konfidenzintervall des Mittelwerts

Ergebnis

$$95\,\%: \ [99{,}71;\ 106{,}3]$$

Lösungsweg

- Umfang, Mittelwert und Stichprobenstandardabweichung:
  $n = 10, \quad \bar{x} = 103, \quad s \approx 4{,}595$
- Der Standardfehler: $SE = \frac{s}{\sqrt{n}} = \frac{4{,}595}{\sqrt{10}} \approx 1{,}453$
- Das Quantil für die gewählte Sicherheit: $t_{0{,}975}(9) \approx 2{,}262$
- Das Konfidenzintervall
  $\bar{x} \pm t \cdot SE = 103 \pm 2{,}262 \cdot 1{,}453 = 103 \pm 3{,}287 \Rightarrow [99,$
- 95 % Konfidenz bedeutet: Würde man die Stichprobe sehr oft wiederholen, enthielten etwa 95 % der so berechneten Intervalle den wahren Mittelwert μ. Es bedeutet nicht, dass μ mit 95 % Wahrscheinlichkeit in diesem konkreten Intervall liegt.
- Bei kleinen Stichproben (n < 30) setzt die t-Verteilung annähernd normalverteilte Daten voraus. Bei stark schiefen Daten oder Ausreißern ist ein nichtparametrischer Test besser geeignet.

Weicht der Mittelwert von 100 ab? Einstichproben-t-Test.

Eingabe: `data(102 98 110 105 99 107 101 104 96 108)` t-Test

Ergebnis

$$t \approx 2{,}065$$

$$p \approx 0{,}06895 > \alpha = 0{,}05$$

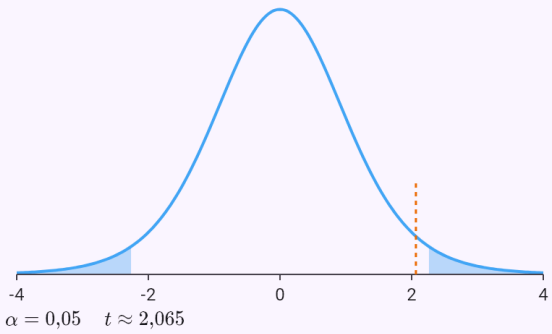
Lösungsweg

- Umfang, Mittelwert und Stichprobenstandardabweichung: $n = 10, \quad \bar{x} = 103, \quad s \approx 4{,}595$
- Nullhypothese und Alternativhypothese: $H_0: \mu = 100, \quad H_1: \mu \neq 100$
- Der Standardfehler: $SE = \frac{s}{\sqrt{n}} = \frac{4{,}595}{\sqrt{10}} \approx 1{,}453$
- Die Teststatistik – unter H_0 t-verteilt mit 9 Freiheitsgraden: $t = \frac{\bar{x} - \mu_0}{SE} = \frac{103 - 100}{1{,}453} \approx 2{,}065$
- Der kritische Wert zum Niveau α: $t_{0{,}975}(9) \approx 2{,}262$, der kritische Bereich $|t| \geq 2{,}262$
- Der p-Wert: $p = 2\,P(T \geq 2{,}065) \approx 0{,}06895$
- Da $p \approx 0{,}06895 > \alpha = 0{,}05$, lehnen wir die Nullhypothese nicht ab – die Daten liefern nicht genug Belege gegen sie.
- Der p-Wert ist die Wahrscheinlichkeit eines mindestens so extremen Ergebnisses wie des beobachteten, falls H_0 gilt – nicht die Wahrscheinlichkeit, dass H_0 gilt. H_0 nicht abzulehnen heißt nicht, dass H_0 wahr ist; statistische Signifikanz sagt nichts über Größe oder praktische Bedeutung eines Unterschieds.

- Bei kleinen Stichproben (n < 30) setzt die t-Verteilung annähernd normalverteilte Daten voraus. Bei stark schiefen Daten oder Ausreißern ist ein nichtparametrischer Test besser geeignet.

Graph

t-Verteilung – schattiert der kritische Bereich, gestrichelt der Wert der Teststatistik



Zwei unabhängige Stichproben verschiedenen Umfangs – Welch-t-Test.

Eingabe: `data(A = 23 25 28 30 26 27; B = 20 22 21 25 19 24 23)` t-Test

Ergebnis

$$t \approx 3{,}503$$

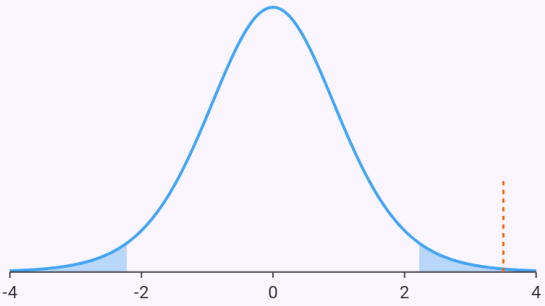
$$p \approx 0{,}005546 \leq \alpha = 0{,}05$$

Lösungsweg

- Umfang, Mittelwert und Stichprobenstandardabweichung
  $n_1 = 6,\ \bar{x}_1 = 26{,}5,\ s_1 \approx 2{,}429$
  $n_2 = 7,\ \bar{x}_2 = 22,\ s_2 \approx 2{,}16$
- Nullhypothese und Alternativhypothese:
  $H_0: \mu_1 = \mu_2, \quad H_1: \mu_1 \neq \mu_2$
- Der Standardfehler
  $$SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{2{,}429^2}{6} + \frac{2{,}16^2}{7}} \approx 1{,}285$$
- Die Freiheitsgrade nach Welch und Satterthwaite
  $$\nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}} \approx 10{,}18$$
- Die Teststatistik – unter H_0 t-verteilt mit $10{,}18$ Freiheitsgraden: $t = \frac{\bar{x}_1 - \bar{x}_2}{SE} = \frac{26{,}5 - 22}{1{,}285} \approx 3{,}503$
- Der kritische Wert zum Niveau α: $t_{0{,}975}(10{,}18) \approx 2{,}223$, der kritische Bereich $|t| \geq 2{,}223$
- Der p-Wert: $p = 2\,P(T \geq 3{,}503) \approx 0{,}005546$
- Da $p \approx 0{,}005546 \leq \alpha = 0{,}05$, lehnen wir die Nullhypothese ab – das Ergebnis ist statistisch signifikant.
- Der p-Wert ist die Wahrscheinlichkeit eines mindestens so extremen Ergebnisses wie des beobachteten, falls H_0 gilt – nicht die Wahrscheinlichkeit, dass H_0 gilt. H_0 nicht abzulehnen heißt nicht, dass H_0 wahr ist; statistische Signifikanz sagt nichts über Größe oder praktische Bedeutung eines Unterschieds.
- Der Welch-Test setzt keine gleichen Varianzen der Gruppen voraus; seine Freiheitsgrade ν müssen nicht ganzzahlig sein.

Graph

t-Verteilung – schattiert der kritische Bereich, gestrichelt der Wert der Teststatistik



$\alpha = 0{,}05 \quad t \approx 3{,}503$

Gepaarter t-Test – Messungen vorher und nachher.

Eingabe: `data(x1 x2; 72 70; 75 73; 68 69; 80 76; 77 74; 74 71)` t-Test · Gepaart

Ergebnis

$$t \approx 3{,}081$$

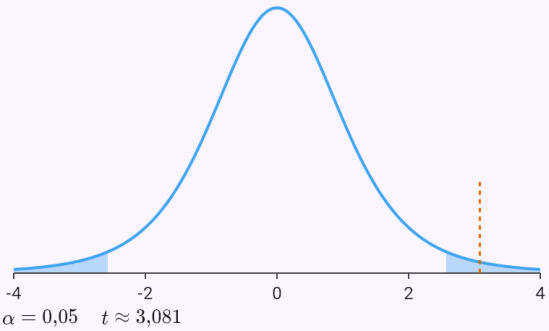
$$p \approx 0{,}02743 \leq \alpha = 0{,}05$$

Lösungsweg

- Die Differenzen der Paare d (Spalte 1 − Spalte 2): $2,\ 2,\ -1,\ 4,\ 3,\ 3;\ n = 6, \quad \bar{d} \approx 2{,}167, \quad s_d \approx 1{,}722$
- Nullhypothese und Alternativhypothese: $H_0 : \mu_d = 0, \quad H_1 : \mu_d \neq 0$
- Der Standardfehler: $SE = \frac{s}{\sqrt{n}} = \frac{1{,}722}{\sqrt{6}} \approx 0{,}7032$
- Die Teststatistik – unter H_0 t-verteilt mit 5 Freiheitsgraden: $t = \frac{\bar{d} - \mu_0}{SE} = \frac{2{,}167 - 0}{0{,}7032} \approx 3{,}081$
- Der kritische Wert zum Niveau α: $t_{0{,}975}(5) \approx 2{,}571$, der kritische Bereich $|t| \geq 2{,}571$
- Der p-Wert: $p = 2\,P(T \geq 3{,}081) \approx 0{,}02743$
- Da $p \approx 0{,}02743 \leq \alpha = 0{,}05$, lehnen wir die Nullhypothese ab – das Ergebnis ist statistisch signifikant.
- Der p-Wert ist die Wahrscheinlichkeit eines mindestens so extremen Ergebnisses wie des beobachteten, falls H_0 gilt – nicht die Wahrscheinlichkeit, dass H_0 gilt. H_0 nicht abzulehnen heißt nicht, dass H_0 wahr ist; statistische Signifikanz sagt nichts über Größe oder praktische Bedeutung eines Unterschieds.
- Bei kleinen Stichproben (n < 30) setzt die t-Verteilung annähernd normalverteilte Daten voraus. Bei stark schiefen Daten oder Ausreißern ist ein nichtparametrischer Test besser geeignet.
- Der gepaarte t-Test ist ein Einstichproben-t-Test der Differenzen (Spalte 1 − Spalte 2) – jedes Paar gehört zur selben Person (z. B. vorher und nachher).

Graph

t-Verteilung – schattiert der kritische Bereich, gestrichelt der Wert der Teststatistik



$\alpha = 0{,}05 \quad t \approx 3{,}081$

Ist der Würfel fair? χ²-Anpassungstest.

Eingabe: `data(8 12 9 11 6 14)` χ²-Test

Ergebnis

$$\chi^2 = 4{,}2$$
$$p \approx 0{,}521 > \alpha = 0{,}05$$

Lösungsweg

- Nullhypothese: alle 6 Kategorien sind gleich wahrscheinlich ($H_0: p_1 = \cdots = p_6 = \frac{1}{6}$); Alternative: mindestens eine weicht ab.
- Die erwarteten Häufigkeiten unter H_0 $E_i = \frac{N}{k} = \frac{60}{6} = 10$ und die Beiträge der Kategorien

i	O_i	E_i	$\frac{(O_i-E_i)^2}{E_i}$
1	8	10	0,4
2	12	10	0,4
3	9	10	0,1
4	11	10	0,1
5	6	10	1,6
6	14	10	1,6
$\sum$	60	60	4,2

- Die Teststatistik – unter H_0 χ²-verteilt mit 5 Freiheitsgraden

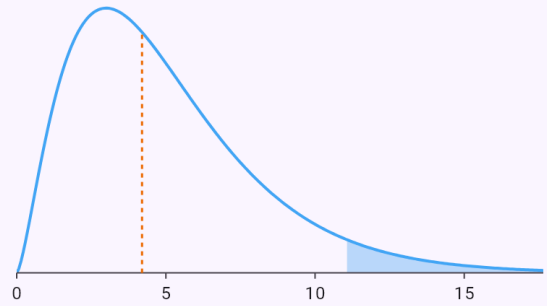
$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = 0{,}4 + 0{,}4 + 0{,}1 + 0{,}1 + 1{,}6 + 1$$

- Der kritische Wert zum Niveau α: $\chi^2_{0{,}95}(5) \approx 11{,}07$, der kritische Bereich $\chi^2 \geq 11{,}07$
- Der p-Wert: $p = P(\chi^2 \geq 4{,}2) \approx 0{,}521$
- Da $p \approx 0{,}521 > \alpha = 0{,}05$, lehnen wir die Nullhypothese nicht ab – die Daten liefern nicht genug Belege gegen sie.

- Der p-Wert ist die Wahrscheinlichkeit eines mindestens so extremen Ergebnisses wie des beobachteten, falls H_0 gilt – nicht die Wahrscheinlichkeit, dass H_0 gilt. H_0 nicht abzulehnen heißt nicht, dass H_0 wahr ist; statistische Signifikanz sagt nichts über Größe oder praktische Bedeutung eines Unterschieds.
- Der χ²-Test arbeitet mit Häufigkeiten (Anzahlen von Fällen), nicht mit Prozenten oder Mittelwerten.

Graph

χ²-Verteilung – schattiert der kritische Bereich, gestrichelt der Wert der Teststatistik



Unabhängigkeitstest in einer 2×2-Tafel.

Eingabe: `data(20 30; 25 15)` χ²-Test

Ergebnis

$$\chi^2 = 4{,}5$$
$$p \approx 0{,}03389 \leq \alpha = 0{,}05$$
$$V \approx 0{,}2236$$

Lösungsweg

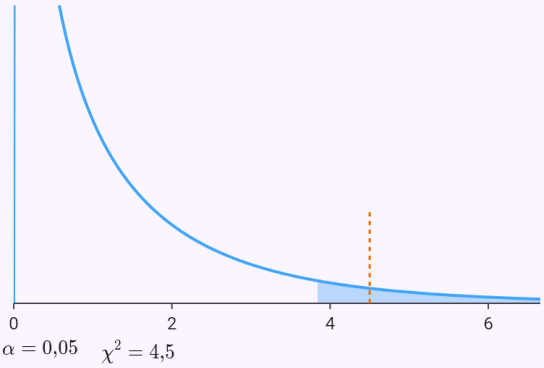
- Nullhypothese: Zeilen- und Spaltenvariable sind unabhängig; Alternative: sie sind abhängig.
- Die erwarteten Häufigkeiten bei Unabhängigkeit $E_{ij} = \frac{n_{i\cdot}\, n_{\cdot j}}{N}$, in Klammern die beobachteten

	1	2	∑
1	25 (20)	25 (30)	50
2	20 (25)	20 (15)	40
∑	45	45	90

- Die Teststatistik – unter H_0 χ²-verteilt mit 1 Freiheitsgraden:
  $\chi^2 = \sum \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = 1 + 1 + 1{,}25 + 1{,}25 = 4{,}5$
- Der kritische Wert zum Niveau α: $\chi^2_{0{,}95}(1) \approx 3{,}841$, der kritische Bereich $\chi^2 \geq 3{,}841$
- Der p-Wert: $p = P(\chi^2 \geq 4{,}5) \approx 0{,}03389$
- Da $p \approx 0{,}03389 \leq \alpha = 0{,}05$, lehnen wir die Nullhypothese ab – das Ergebnis ist statistisch signifikant.
- Der p-Wert ist die Wahrscheinlichkeit eines mindestens so extremen Ergebnisses wie des beobachteten, falls H_0 gilt – nicht die Wahrscheinlichkeit, dass H_0 gilt. H_0 nicht abzulehnen heißt nicht, dass H_0 wahr ist; statistische Signifikanz sagt nichts über Größe oder praktische Bedeutung eines Unterschieds.
- Die Stärke des Zusammenhangs – Cramérs V (0 keiner, 1 vollständig)
  $V = \sqrt{\frac{\chi^2}{N(\min(r,c) - 1)}} = \sqrt{\frac{4{,}5}{90 \cdot 1}} \approx 0{,}2236$
- Der χ²-Test arbeitet mit Häufigkeiten (Anzahlen von Fällen), nicht mit Prozenten oder Mittelwerten.

Graph

χ²-Verteilung – schattiert der kritische Bereich, gestrichelt der Wert der Teststatistik



Zwei Gruppen mit dem Mann-Whitney-Test – kleine Stichproben, exakter p-Wert.

Eingabe: `data(A = 23 25 28 30 26 27; B = 20 22 21 25 19 24 23)`

Rangtest (Mann-Whitney, Wilcoxon, Kruskal-Wallis)

Ergebnis

$$U = 3$$
$$z \approx -2{,}579$$
$$p \approx 0{,}009922 \leq \alpha = 0{,}05$$

Lösungsweg

- Nullhypothese: Beide Gruppen stammen aus derselben Verteilung, Alternative: Die Werte einer Gruppe sind systematisch höher

  $H_0\colon F_\mathrm{A} = F_\mathrm{B}$
  $H_1\colon F_\mathrm{A} \neq F_\mathrm{B}$

- Die Werte beider Gruppen zusammen sortieren und mit ihrem Rang R nummerieren (gleiche Werte erhalten den Mittelwert ihrer Ränge)

x	R	
19	1	B
20	2	B
21	3	B
22	4	B
23	5,5	A
23	5,5	B
24	7	B
25	8,5	A
25	8,5	B
26	10	A
27	11	A
28	12	A
30	13	A

- Gruppengrößen und Rangsummen:
  $n_1 = 6, \quad n_2 = 7, \quad R_1 = 60, \quad R_2 = 31$

- Die Mann-Whitney-Statistik U – die Anzahl der Paare, in denen der Wert der einen Gruppe einen höheren Rang hat als der der anderen; geprüft wird die kleinere

$$U_1 = R_1 - \frac{n_1(n_1+1)}{2} = 60 - \frac{6 \cdot 7}{2} = 39$$
$$U_2 = n_1 n_2 - U_1 = 6 \cdot 7 - 39 = 3$$
$$U = \min(U_1, U_2) = 3$$

- Die Teststatistik – unter H_0 standardnormalverteilt:
  $z = \dfrac{U - \frac{n_1 n_2}{2}}{\sigma_U} = \dfrac{3 - 21}{6{,}981} \approx -2{,}579$

- Der kritische Wert zum Niveau α: $z_{0{,}975} \approx 1{,}96$, der kritische Bereich $|z| \geq 1{,}96$

- Der p-Wert: $p = 2\,P(Z \geq 2{,}579) \approx 0{,}009922$

- Da $p \approx 0{,}009922 \leq \alpha = 0{,}05$, lehnen wir die Nullhypothese ab – das Ergebnis ist statistisch signifikant.

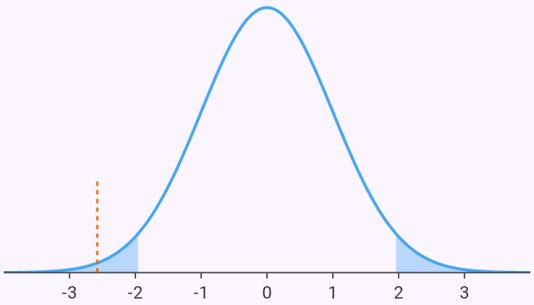
- Der p-Wert ist die Wahrscheinlichkeit eines mindestens so extremen Ergebnisses wie des beobachteten, falls H_0 gilt – nicht die Wahrscheinlichkeit, dass H_0 gilt. H_0 nicht abzulehnen heißt nicht, dass H_0 wahr ist; statistische Signifikanz sagt nichts über Größe oder praktische Bedeutung eines Unterschieds.

- Die Effektstärke r (grob 0,1 klein, 0,3 mittel, 0,5 groß):
  $r = \dfrac{|z|}{\sqrt{N}} = \dfrac{2{,}579}{\sqrt{13}} \approx 0{,}7152$

- Der Mann-Whitney-Test ist ein Rangtest: Er setzt keine Normalverteilung voraus und eignet sich auch für ordinale Daten und Ausreißer. Er vergleicht die ganzen Verteilungen – bei ähnlicher Form ist er ein Test auf einen Unterschied der Mediane.

Graph

Standardnormalverteilung – schattiert der kritische Bereich, gestrichelt der Wert der Teststatistik



$\alpha = 0{,}05 \quad z \approx -2{,}579$

Gepaarte Werte vorher und nachher mit dem Wilcoxon-Test.

Eingabe: `data(pred po; 72 70; 75 73; 68 69; 80 76; 77 74; 74 71; 79 72)`

Rangtest (Mann-Whitney, Wilcoxon, Kruskal-Wallis) · Gepaart

Ergebnis

$$T = 1$$
$$z \approx -2{,}205$$
$$p \approx 0{,}02743 \leq \alpha = 0{,}05$$

Lösungsweg

- Nullhypothese: Der Median der Differenzen ist null (keine Veränderung), Alternative: Er ist es nicht: $H_0\colon \tilde{d} = 0, \quad H_1\colon \tilde{d} \neq 0$
- Die Differenzen der Paare d_i (1. Spalte − 2. Spalte): $2,\ 2,\ -1,\ 4,\ 3,\ 3,\ 7$; Nulldifferenzen werden weggelassen: $n = 7$
- Die von null verschiedenen Differenzen nach ihrem Betrag sortieren und mit dem Rang R nummerieren; das Vorzeichen bleibt beim Rang

d_i	R	
-1	1	$-$
2	$2{,}5$	$+$
2	$2{,}5$	$+$
3	$4{,}5$	$+$
3	$4{,}5$	$+$
4	6	$+$
7	7	$+$

- Die Rangsummen der positiven und der negativen Differenzen
  $T^+ = 27$
  $T^- = 1$
  $T = \min\left(T^+, T^-\right) = 1$
- Die Teststatistik – unter H_0 standardnormalverteilt: $z = \dfrac{T - \frac{n(n+1)}{4}}{\sigma_T} = \dfrac{1 - 14}{5{,}895} \approx -2{,}205$
- Der kritische Wert zum Niveau α: $z_{0{,}975} \approx 1{,}96$, der kritische Bereich $|z| \geq 1{,}96$
- Der p-Wert: $p = 2\,P(Z \geq 2{,}205) \approx 0{,}02743$
- Da $p \approx 0{,}02743 \leq \alpha = 0{,}05$, lehnen wir die Nullhypothese ab – das Ergebnis ist statistisch signifikant.
- Der p-Wert ist die Wahrscheinlichkeit eines mindestens so extremen Ergebnisses wie des beobachteten, falls H_0 gilt – nicht die Wahrscheinlichkeit, dass H_0 gilt. H_0 nicht abzulehnen heißt nicht, dass H_0 wahr ist; statistische Signifikanz sagt nichts über Größe oder praktische Bedeutung eines Unterschieds.
- Die Effektstärke r (grob 0,1 klein, 0,3 mittel, 0,5 groß): $r = \dfrac{|z|}{\sqrt{n}} = \dfrac{2{,}205}{\sqrt{7}} \approx 0{,}8335$
- Der Wilcoxon-Test ist das Rang-Gegenstück zum gepaarten t-Test: Er setzt keine normalverteilten Differenzen voraus und berücksichtigt ihre Vorzeichen und Beträge über die Ränge.

Graph

Standardnormalverteilung – schattiert der kritische Bereich, gestrichelt der Wert der Teststatistik



$\alpha = 0{,}05 \quad z \approx -2{,}205$

Drei Gruppen mit dem Kruskal-Wallis-Test.

Eingabe: `data(A = 12 15 14 11 13; B = 16 18 17 15 19; C = 13 14 12 16 15)`

Rangtest (Mann-Whitney, Wilcoxon, Kruskal-Wallis)

Ergebnis

$$H \approx 8{,}263$$
$$p \approx 0{,}01606 \leq \alpha = 0{,}05$$

Lösungsweg

- Nullhypothese: Alle Gruppen stammen aus derselben Verteilung, Alternative: Mindestens eine unterscheidet sich

  $H_0: F_1 = \cdots = F_3$

  $H_1: \exists\, i, j: F_i \neq F_j$

- Die Werte aller Gruppen zusammen sortieren und mit ihrem Rang R nummerieren (gleiche Werte erhalten den Mittelwert ihrer Ränge)

x	R	
11	1	A
12	2,5	A
12	2,5	C
13	4,5	A
13	4,5	C
14	6,5	A
14	6,5	C
15	9	A
15	9	B
15	9	C
16	11,5	B
16	11,5	C
17	13	B
18	14	B
19	15	B

- Die Rangsumme jeder Gruppe

  $R_\mathrm{A} = 23{,}5 \ (n = 5)$

  $R_\mathrm{B} = 62{,}5 \ (n = 5)$

  $R_\mathrm{C} = 34 \ (n = 5)$

- Die Teststatistik – unter H_0 χ^2-verteilt mit 2 Freiheitsgraden

$$H = \frac{12}{N(N+1)} \sum \frac{R_j^2}{n_j} - 3(N+1) = \frac{12}{15 \cdot 16} \left(\frac{23{,}5^2}{5} \right.$$

$$C = 1 - \frac{\sum (t^3 - t)}{N^3 - N} = 1 - \frac{48}{3360} \approx 0{,}9857$$

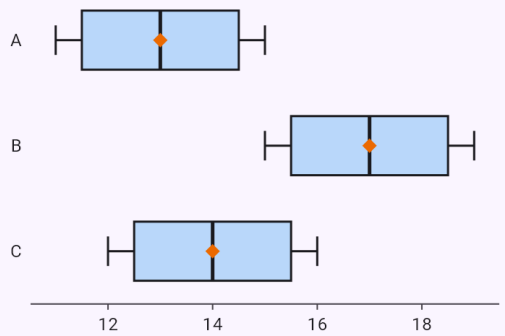
$$H_c = \frac{H}{C} = \frac{8{,}145}{0{,}9857} \approx 8{,}263$$

- Der kritische Wert zum Niveau α: $\chi^2_{0{,}95}(2) \approx 5{,}991$, der kritische Bereich $\chi^2 \geq 5{,}991$
- Der p-Wert: $p = P\left(\chi^2 \geq 8{,}263\right) \approx 0{,}01606$
- Da $p \approx 0{,}01606 \leq \alpha = 0{,}05$, lehnen wir die Nullhypothese ab – das Ergebnis ist statistisch signifikant.

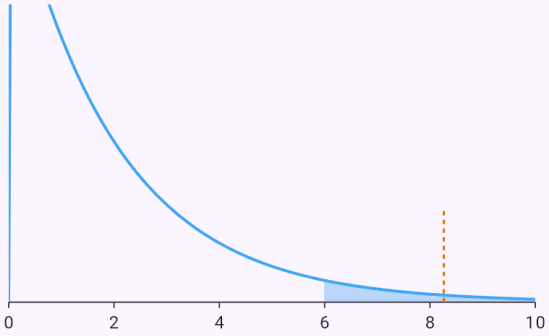
- Der p-Wert ist die Wahrscheinlichkeit eines mindestens so extremen Ergebnisses wie des beobachteten, falls H_0 gilt – nicht die Wahrscheinlichkeit, dass H_0 gilt. H_0 nicht abzulehnen heißt nicht, dass H_0 wahr ist; statistische Signifikanz sagt nichts über Größe oder praktische Bedeutung eines Unterschieds.
- Die Effektstärke ε^2 – der durch die Gruppen erklärte Anteil der Rangvarianz: $\varepsilon^2 = \frac{H}{N-1} = \frac{8{,}263}{14} \approx 0{,}5902$
- Der Kruskal-Wallis-Test ist das Rang-Gegenstück zur einfaktoriellen ANOVA. Er sagt nur, dass sich einige Gruppen unterscheiden – welche, zeigt erst ein Paarvergleich (Mann-Whitney-Test mit Bonferroni-Korrektur).

Graph

Boxplot (Median, Quartile, Ausreißer; ◆ Mittelwert)



χ^2-Verteilung – schattiert der kritische Bereich, gestrichelt der Wert der Teststatistik



$\alpha = 0{,}05$ $\chi^2 \approx 8{,}263$

Die Mittelwerte dreier Gruppen mit der einfaktoriellen ANOVA.

Eingabe: `data(A = 12 15 14 11 13; B = 16 18 17 15 19; C = 13 14 12 16 15)`

ANOVA (Varianzanalyse)

Ergebnis

$$F \approx 8{,}667$$
$$p \approx 0{,}004687 \leq \alpha = 0{,}05$$
$$\eta^2 \approx 0{,}5909$$

Lösungsweg

- Nullhypothese: Die Mittelwerte aller Gruppen sind gleich, Alternative: Mindestens zwei unterscheiden sich

  $H_0: \mu_1 = \cdots = \mu_3$
  $H_1: \exists\, i, j: \mu_i \neq \mu_j$

- Größe, Mittelwert und Standardabweichung der Gruppen und der Gesamtmittelwert $\bar{x} \approx 14{,}67$

	n_j	$\bar{x}_j$	s_j
A	5	13	1,581
B	5	17	1,581
C	5	14	1,581
$\sum$	15	14,67	

- Varianzanalyse – Quadratsumme zwischen den Gruppen (B) und innerhalb (W), Freiheitsgrade und mittlere Quadrate

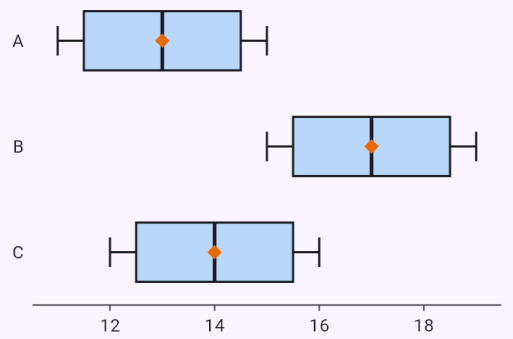
$$SS_B = \sum n_j (\bar{x}_j - \bar{x})^2$$
$$SS_W = \sum \sum (x_{ij} - \bar{x}_j)^2$$

	SS	df	MS	F
B	43,33	2	21,67	8,667
W	30	12	2,5	
$\sum$	73,33	14		

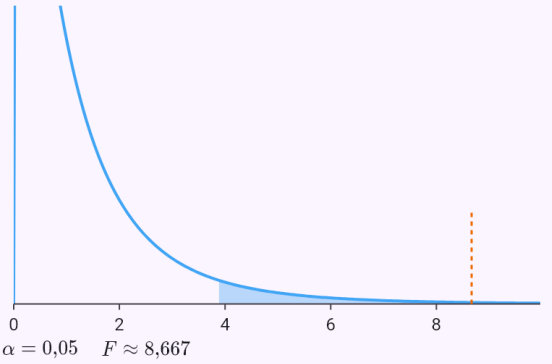
- Die Prüfgröße – unter H_0 F-verteilt mit 2 und 12 Freiheitsgraden: $F = \frac{MS_B}{MS_W} = \frac{21{,}67}{2{,}5} \approx 8{,}667$
- Der kritische Wert zum Niveau α: $F_{0{,}95}(2;\, 12) \approx 3{,}885$, der kritische Bereich $F \geq 3{,}885$
- Der p-Wert: $p = P(F \geq 8{,}667) \approx 0{,}004687$
- Da $p \approx 0{,}004687 \leq \alpha = 0{,}05$, lehnen wir die Nullhypothese ab – das Ergebnis ist statistisch signifikant.
- Die Effektstärke η² – der durch die Unterschiede zwischen den Gruppen erklärte Anteil der Varianz (grob 0,01 klein, 0,06 mittel, 0,14 groß): $\eta^2 = \frac{SS_B}{SS_T} = \frac{43{,}33}{73{,}33} \approx 0{,}5909$
- Der p-Wert ist die Wahrscheinlichkeit eines mindestens so extremen Ergebnisses wie des beobachteten, falls H_0 gilt – nicht die Wahrscheinlichkeit, dass H_0 gilt. H_0 nicht abzulehnen heißt nicht, dass H_0 wahr ist; statistische Signifikanz sagt nichts über Größe oder praktische Bedeutung eines Unterschieds.
- Die ANOVA setzt unabhängige Gruppen, in jeder eine annähernd normale Verteilung und ähnliche Varianzen voraus. Gilt das nicht (kleine Gruppen, Ausreißer, ordinale Daten), nehmen Sie den Kruskal-Wallis-Test.
- Die Ablehnung von H_0 sagt nur, dass sich mindestens zwei Mittelwerte unterscheiden. Welche, zeigt erst ein Paarvergleich (Post-hoc-Tests, z. B. t-Tests mit Bonferroni-Korrektur).

Graph

Boxplot (Median, Quartile, Ausreißer; ◆ Mittelwert)



F-Verteilung – schattiert der kritische Bereich, gestrichelt der Wert der Prüfgröße



$\alpha = 0{,}05$ $F \approx 8{,}667$

Anpassungstest gegen die erwarteten Anteile 1 : 2 : 1.

Eingabe: `data(30 50 20)` χ²-Test · Erwartete Anteile 1:2:1

Ergebnis

$$\chi^2 = 2$$
$$p \approx 0{,}3679 > \alpha = 0{,}05$$

Lösungsweg

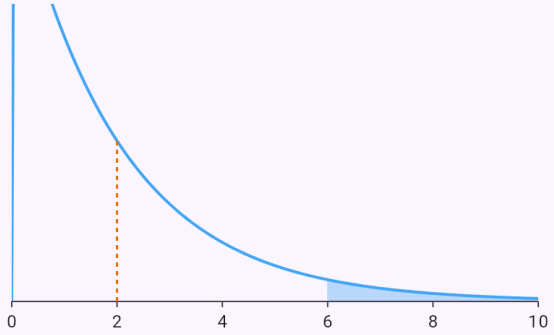
- Nullhypothese: Die Kategorien haben die angegebenen Anteile, Alternative: Mindestens ein Anteil weicht ab: $H_0: p_1 = 0{,}25,\ p_2 = 0{,}5,\ p_3 = 0{,}25$
- Die erwarteten Häufigkeiten nach den angegebenen Anteilen $E_i = N \cdot p_i, \quad N = 100$ und die Beiträge der Kategorien

i	O_i	E_i	$\frac{(O_i-E_i)^2}{E_i}$
1	30	25	1
2	50	50	0
3	20	25	1
$\sum$	100	100	2

- Die Teststatistik – unter H_0 χ²-verteilt mit 2 Freiheitsgraden: $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = 1 + 0 + 1 = 2$
- Der kritische Wert zum Niveau α: $\chi^2_{0{,}95}(2) \approx 5{,}991$, der kritische Bereich $\chi^2 \geq 5{,}991$
- Der p-Wert: $p = P(\chi^2 \geq 2) \approx 0{,}3679$
- Da $p \approx 0{,}3679 > \alpha = 0{,}05$, lehnen wir die Nullhypothese nicht ab – die Daten liefern nicht genug Belege gegen sie.
- Der p-Wert ist die Wahrscheinlichkeit eines mindestens so extremen Ergebnisses wie des beobachteten, falls H_0 gilt – nicht die Wahrscheinlichkeit, dass H_0 gilt. H_0 nicht abzulehnen heißt nicht, dass H_0 wahr ist; statistische Signifikanz sagt nichts über Größe oder praktische Bedeutung eines Unterschieds.
- Der χ²-Test arbeitet mit Häufigkeiten (Anzahlen von Fällen), nicht mit Prozenten oder Mittelwerten.

Graph

χ²-Verteilung – schattiert der kritische Bereich, gestrichelt der Wert der Teststatistik



Quadratische Regression.

Eingabe: `data(x y; 1 2.1; 2 3.9; 3 9.2; 4 15.8; 5 25.1; 6 36.2)`

Korrelation und Regression · Quadratisch

Ergebnis

$$\hat{y} = 1{,}73 - 0{,}9354x + 1{,}116x^2$$

$$R^2 \approx 0{,}9997$$

Lösungsweg

- Anzahl der Werte $n = 6$, sortierte Daten
  $[1;\,2{,}1],\ [2;\,3{,}9],\ [3;\,9{,}2],\ [4;\,15{,}8],\ [5;\,25{,}1],\ [6;\,36{,}2]$
- Summen und Mittelwerte
  $\sum x_i = 21$
  $\sum x_i^2 = 91$
  $\sum x_i^3 = 441$
  $\sum x_i^4 = 2275$
  $\sum y_i = 92{,}3$
  $\sum x_i y_i = 443{,}4$
  $\sum x_i^2 y_i = 2284$
- Die Normalgleichungen der Methode der kleinsten Quadrate für ŷ = a + bx + cx²

$$\begin{aligned} 6a + 21b + 91c &= 92{,}3 \\ 21a + 91b + 441c &= 443{,}4 \\ 91a + 441b + 2275c &= 2284 \end{aligned}$$

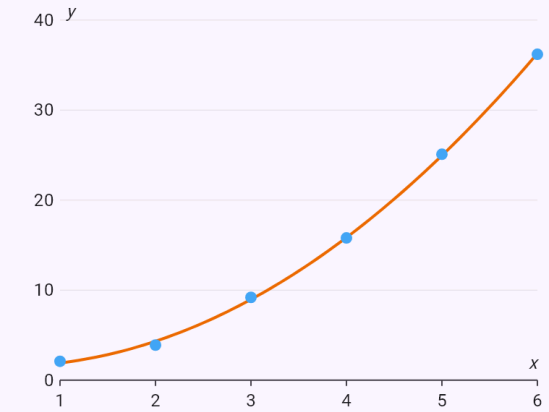
- Die Regressionskurve: $\hat{y} = 1{,}73 - 0{,}9354x + 1{,}116x^2$
- Das Bestimmtheitsmaß – die Kurve erklärt $99{,}97\,\%$ der Varianz von y.

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} = 1 - \frac{0{,}3012}{874{,}5} \approx 0{,}9997$$

- Die Regressionskurve ist nur im Bereich der gemessenen x sinnvoll – außerhalb (Extrapolation) kann die Schätzung sehr ungenau sein.

Graph

Streudiagramm mit der Regressionskurve



Exponentielle Regression durch Linearisierung.

Eingabe: `data(x y; 0 2.0; 1 2.7; 2 3.7; 3 5.0; 4 6.8; 5 9.1)`

Korrelation und Regression · Exponentiell

Ergebnis

$$\hat{y} = 2{,}002\, e^{0{,}3042x}$$

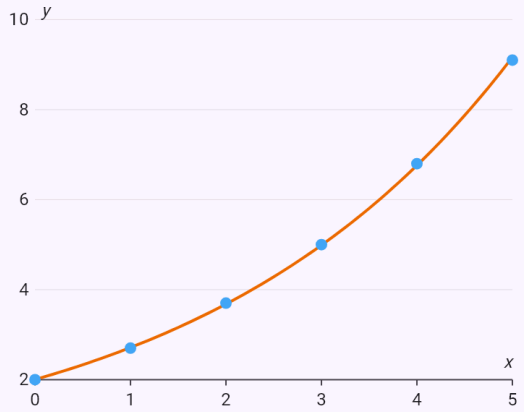
$$R^2 \approx 0{,}9998$$

Lösungsweg

- Anzahl der Werte $n = 6$, sortierte Daten
  $[0;\ 2],\ [1;\ 2{,}7],\ [2;\ 3{,}7],\ [3;\ 5],\ [4;\ 6{,}8],\ [5;\ 9{,}1]$
- Durch Logarithmieren wird das Modell zu einer Geraden in den Variablen (u, v): $y = a\, e^{bx} \Rightarrow \ln y = \ln a + bx$; die transformierten Paare
  $[0;\ 0{,}6931],\ [1;\ 0{,}9933],\ [2;\ 1{,}308],\ [3;\ 1{,}609],\ [4;\ 1{,}91$
- Koeffizienten der Regressionsgeraden nach der Methode der kleinsten Quadrate
  $$b = \frac{\sum (u_i - \bar{u})(v_i - \bar{v})}{\sum (u_i - \bar{u})^2} = \frac{5{,}324}{17{,}5} \approx 0{,}3042$$
  $$\ln a = \overline{\ln y} - b\bar{x} = 1{,}455 - 0{,}3042 \cdot 2{,}5 \approx 0{,}6943$$
- Den Koeffizienten a erhalten wir durch Rücktransformation: $a = e^{0{,}6943} \approx 2{,}002$
- Die Regressionskurve: $\hat{y} = 2{,}002\, e^{0{,}3042x}$
- Das Bestimmtheitsmaß – die Kurve erklärt $99{,}98\,\%$ der Varianz von y.
  $$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} = 1 - \frac{0{,}006563}{35{,}95} \approx 0{,}9998$$
- Die Koeffizienten sind nach der Methode der kleinsten Quadrate für die transformierten Daten (Logarithmen) berechnet, nicht für die ursprünglichen Werte – das Ergebnis kann leicht von einer nichtlinearen Regression in Statistikprogrammen abweichen.
- Die Regressionskurve ist nur im Bereich der gemessenen x sinnvoll – außerhalb (Extrapolation) kann die Schätzung sehr ungenau sein.

Graph

Streudiagramm mit der Regressionskurve



Rechtsseitiger t-Test aus zusammengefassten Werten: n = 25, x̄ = 104, s = 15, μ₀ = 100.

Eingabe: `model(ttest; find=test[one]; n=25[one]; mean=104[one]; sd=15[one]; mu0=100[one];`

`alpha=0.05[one]; alternative=greater; sigma=sample)`

Ergebnis

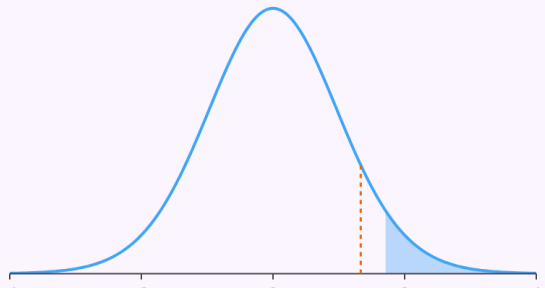
$$t \approx 1{,}333$$
$$p \approx 0{,}09747 > \alpha = 0{,}05$$

Lösungsweg

- Gesucht t; gegeben: $n = 25, \quad \bar{x} = 104, \quad s = 15, \quad \mu_0 = 100, \quad \alpha = 0{,}05$
- Der t-Test prüft, ob der Mittelwert der Grundgesamtheit von einem Wert μ_0 abweicht. Die Teststatistik $t = (\bar{x} - \mu_0)/(s/\sqrt{n})$ ist unter der Nullhypothese t-verteilt mit $n - 1$ Freiheitsgraden; der p-Wert gibt an, wie wahrscheinlich ein solches oder extremeres Ergebnis wäre, wenn H_0 gälte. Bei bekanntem σ ist es ein z-Test. Gleichung des Modells: $t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$
- Nullhypothese und Alternativhypothese: $H_0: \mu = 100, \quad H_1: \mu > 100$
- Der Standardfehler: $SE = \frac{s}{\sqrt{n}} = \frac{15}{\sqrt{25}} = 3$
- Die Teststatistik – unter H_0 t-verteilt mit 24 Freiheitsgraden: $t = \frac{\bar{x} - \mu_0}{SE} = \frac{104 - 100}{3} \approx 1{,}333$
- Der kritische Wert zum Niveau α: $t_{0{,}95}(24) \approx 1{,}711$, der kritische Bereich $t \geq 1{,}711$
- Der p-Wert: $p = P(T \geq 1{,}333) \approx 0{,}09747$
- Da $p \approx 0{,}09747 > \alpha = 0{,}05$, lehnen wir die Nullhypothese nicht ab – die Daten liefern nicht genug Belege gegen sie.
- Der p-Wert ist die Wahrscheinlichkeit eines mindestens so extremen Ergebnisses wie des beobachteten, falls H_0 gilt – nicht die Wahrscheinlichkeit, dass H_0 gilt. H_0 nicht abzulehnen heißt nicht, dass H_0 wahr ist; statistische Signifikanz sagt nichts über Größe oder praktische Bedeutung eines Unterschieds.
- Bei kleinen Stichproben (n < 30) setzt die t-Verteilung annähernd normalverteilte Daten voraus. Bei stark schiefen Daten oder Ausreißern ist ein nichtparametrischer Test besser geeignet.

Graph

t-Verteilung – schattiert der kritische Bereich, gestrichelt der Wert der Teststatistik



$\alpha = 0{,}05 \quad t \approx 1{,}333$

Welch-t-Test zweier Gruppen aus zusammengefassten Werten.

Eingabe: `model(twosample; find=test[one]; n1=30[one]; mean1=52[one]; sd1=8[one];`

```
n2=28[one]; mean2=47[one]; sd2=10[one]; alpha=0.05[one]; variances=welch; alternative=two)
```

Ergebnis

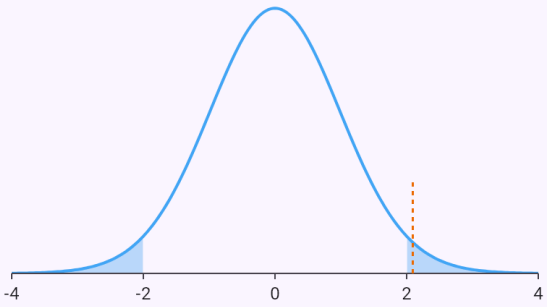
$$t \approx 2{,}093$$
$$p \approx 0{,}04124 \leq \alpha = 0{,}05$$

Lösungsweg

- Gesucht t; gegeben
  $n_1 = 30, \quad \bar{x}_1 = 52, \quad s_1 = 8, \quad n_2 = 28, \quad \bar{x}_2 = 47,$
- Vergleich der Mittelwerte zweier unabhängiger Gruppen. Der Welch-Test setzt keine gleichen Varianzen voraus und wird heute empfohlen; der klassische t-Test mit gepoolter Varianz setzt gleiche Varianzen voraus. Es genügen Umfänge, Mittelwerte und Standardabweichungen beider Gruppen. Gleichung des Modells: $t = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$
- Nullhypothese und Alternativhypothese:
  $H_0 \colon \mu_1 = \mu_2, \quad H_1 \colon \mu_1 \neq \mu_2$
- Der Standardfehler
  $$SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{8^2}{30} + \frac{10^2}{28}} \approx 2{,}388$$
- Die Freiheitsgrade nach Welch und Satterthwaite
  $$\nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}} \approx 51{,}71$$
- Die Teststatistik – unter H_0 t-verteilt mit $51{,}71$ Freiheitsgraden: $t = \dfrac{\bar{x}_1 - \bar{x}_2}{SE} = \dfrac{52 - 47}{2{,}388} \approx 2{,}093$
- Der kritische Wert zum Niveau α: $t_{0{,}975}(51{,}71) \approx 2{,}007$, der kritische Bereich $|t| \geq 2{,}007$
- Der p-Wert: $p = 2\,P(T \geq 2{,}093) \approx 0{,}04124$
- Da $p \approx 0{,}04124 \leq \alpha = 0{,}05$, lehnen wir die Nullhypothese ab – das Ergebnis ist statistisch signifikant.
- Der p-Wert ist die Wahrscheinlichkeit eines mindestens so extremen Ergebnisses wie des beobachteten, falls H_0 gilt – nicht die Wahrscheinlichkeit, dass H_0 gilt. H_0 nicht abzulehnen heißt nicht, dass H_0 wahr ist; statistische Signifikanz sagt nichts über Größe oder praktische Bedeutung eines Unterschieds.
- Der Welch-Test setzt keine gleichen Varianzen der Gruppen voraus; seine Freiheitsgrade ν müssen nicht ganzzahlig sein.

Graph

t-Verteilung – schattiert der kritische Bereich, gestrichelt der Wert der Teststatistik



$\alpha = 0{,}05 \quad t \approx 2{,}093$

Normalverteilung: P(X < 85) für N(100; 15²).

Eingabe: `model(normal; find=below[one]; mu=100[one]; sigma=15[one]; x=85[one])`

Ergebnis

$$P(X < 85) \approx 0{,}1587 = 15{,}87\,\%$$

Lösungsweg

- Gesucht P; gegeben: $\mu = 100, \quad \sigma = 15, \quad x = 85$
- Viele Größen (Körpergröße, Testleistung, Messfehler) schwanken symmetrisch um den Mittelwert μ, je weiter weg, desto seltener – das beschreibt die Normalverteilung (Gauß) N(μ; σ²) mit der Standardabweichung σ. Wahrscheinlichkeiten berechnet man über die Standardisierung z = (x − μ)/σ und die Verteilungsfunktion Φ der Standardnormalverteilung N(0; 1): P(X < x) = Φ(z). Ein Quantil ist der Wert, unter dem der Anteil p liegt. Gleichung des Modells

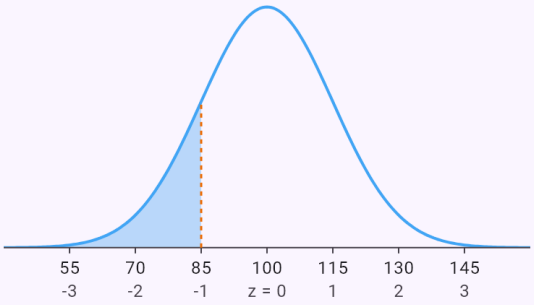
  $$X \sim N(\mu;\, \sigma^2)$$
  $$Z = \frac{X - \mu}{\sigma} \sim N(0;\, 1)$$

- Gesuchte Größe ausdrücken: $z = \frac{x - \mu}{\sigma}$
- Einsetzen: $z = \frac{85 - 100}{15} = -1$
- Einsetzen

  $$P(X < 85) = \Phi(-1) = (1 - \Phi(1)) \approx (1 - 0{,}8413) \approx 0$$

Graph

Dichte der Normalverteilung, die gefärbte Fläche ist die gesuchte Wahrscheinlichkeit



$P \approx 0{,}1587 \quad x = 85$

Der Anteil der Menschen mit 160 bis 185 cm Körpergröße bei μ = 170, σ = 8.

Eingabe: `model(normal; find=between[one]; mu=170[one]; sigma=8[one]; a=160[one];`

```
b=185[one])
```

Ergebnis

$$P(160 < X < 185) \approx 0{,}864 = 86{,}4\,\%$$

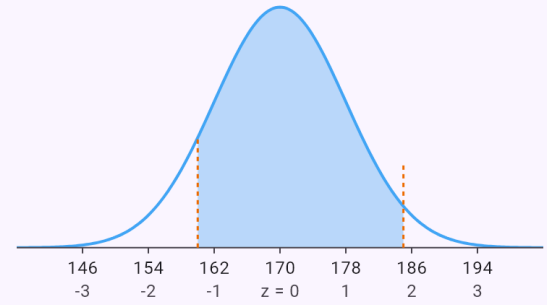
Lösungsweg

- Gesucht P; gegeben:
  $\mu = 170, \quad \sigma = 8, \quad a = 160, \quad b = 185$
- Viele Größen (Körpergröße, Testleistung, Messfehler) schwanken symmetrisch um den Mittelwert μ, je weiter weg, desto seltener – das beschreibt die Normalverteilung (Gauß) N(μ; σ²) mit der Standardabweichung σ. Wahrscheinlichkeiten berechnet man über die Standardisierung z = (x − μ)/σ und die Verteilungsfunktion Φ der Standardnormalverteilung N(0; 1): P(X < x) = Φ(z). Ein Quantil ist der Wert, unter dem der Anteil p liegt. Gleichung des Modells
  $X \sim N\left(\mu;\, \sigma^2\right)$
  $Z = \dfrac{X - \mu}{\sigma} \sim N(0;\, 1)$
- Gesuchte Größe ausdrücken: $z_a = \dfrac{a - \mu}{\sigma}, \quad z_b = \dfrac{b - \mu}{\sigma}$
- Einsetzen:
  $z_a = \dfrac{160 - 170}{8} = -1{,}25, \quad z_b = \dfrac{185 - 170}{8} = 1{,}875$
- Einsetzen

$$P(160 < X < 185) = \Phi(1{,}875) - \Phi(-1{,}25) \approx 0{,}9696$$

Graph

Dichte der Normalverteilung, die gefärbte Fläche ist die gesuchte Wahrscheinlichkeit



$P \approx 0{,}864 \quad a \quad b$

Ein Quantil: der Wert, unter dem 95 % der Population liegen.

Eingabe: `model(normal; find=quantile[one]; mu=100[one]; sigma=15[one]; p=95[pct])`

Ergebnis

$$x \approx 124{,}7$$

Lösungsweg

- Gesucht x; gegeben: $\mu = 100, \quad \sigma = 15, \quad p = 95\,\%$
- Einheiten umrechnen: $p = 95\,\% = 0{,}95$
- Viele Größen (Körpergröße, Testleistung, Messfehler) schwanken symmetrisch um den Mittelwert μ, je weiter weg, desto seltener – das beschreibt die Normalverteilung (Gauß) N(μ; σ²) mit der Standardabweichung σ. Wahrscheinlichkeiten berechnet man über die Standardisierung z = (x − μ)/σ und die Verteilungsfunktion Φ der Standardnormalverteilung N(0; 1): P(X < x) = Φ(z). Ein Quantil ist der Wert, unter dem der Anteil p liegt. Gleichung des Modells

$$X \sim N(\mu;\, \sigma^2)$$

$$Z = \frac{X - \mu}{\sigma} \sim N(0;\, 1)$$

- Gesuchte Größe ausdrücken

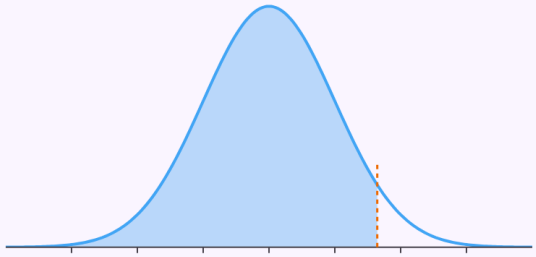
$$P(X < x) = \Phi\left(\frac{x - \mu}{\sigma}\right) = p \Rightarrow x = \mu + z_p \sigma, \quad z_p =$$

- Einsetzen

$$z_{0{,}95} = \Phi^{-1}(0{,}95) \approx 1{,}645, \quad x = 100 + 1{,}645 \cdot 15 \approx 1$$

Graph

Dichte der Normalverteilung, die gefärbte Fläche ist die gesuchte Wahrscheinlichkeit



$p = 0{,}95 \quad x \approx 124{,}7$

Bekannte Einschränkungen

- Die Quartile sind die Mediane der Hälften (das Schulverfahren); Statistikprogramme (Excel, SPSS) nutzen andere Formeln und erhalten bei kleinen Daten leicht andere Werte.
- Paarvergleiche nach ANOVA oder Kruskal-Wallis-Test (post hoc) und mehrfaktorielle ANOVA gibt es noch nicht – vergleichen Sie die Paare mit t-Test oder Mann-Whitney-Test mit Bonferroni-Korrektur; nichtlineare Regressionen werden über Linearisierung berechnet.
- Daten werden nur von Hand, im Gitter oder als CSV-Datei eingegeben – eine Tabelle auf einem Foto erkennt die App noch nicht.

Erzeugt aus der App Mathtrace