Definition (Grundgesamtheit)**.** Die **Grundgesamtheit** ist die Menge aller relevanten statistischen Einheiten; eine Teilmenge davon bezeichnen wir auch als **Teilgesamtheit.**

Definition (Stichprobe)**.** Die tatsächlich untersuchte Teilgesamtheit heißt **Stichprobe.**

Definition (Merkmalstypen)**.** Wir unterscheiden zwischen drei **Merkmalstypen**:

- Kann ein Merkmal nur endlich oder abzählbar viele Ausprägungen annehmen, dann sprechen wir von einem **diskreten Merkmal.**
- Kann ein Merkmal alle Werte eines Intervalls $I \subset \mathbb{R}$ annehmen, dann sprechen wir von einem **stetigen Merkmal.**
- Wir sprechen von einem **quasi-stetigen Merkmal**, wenn eine Größe nur diskret erhoben werden kann, die Abstufungen aber so fein sind, dass sie sich wie ein stetiges Merkmal behandeln lässt.

Definition (Rohdaten)**.** Als **Rohdaten** oder **Urliste** bezeichnen wir die Menge aller beobachteten Merkmalsausprägungen.

Definition (Skala, Skalenniveaus)**.**

1) Wir unterscheiden die folgenden **Skalenniveaus.**
   - Die Ausprägungen **nominalskalierter Merkmale** sind Namen, Klassen oder Kategorien; es gibt keine natürliche Ordnung der Ausprägungen.
   - **Ordinalskalierte Merkmale** hingegen enthalten Nominal-Informationen und können geordnet werden, es gibt aber kein Abstandsmaß.
   - Existiert zu den Ordinal-Informationen noch ein Abstandsmaß, aber kein *natürlicher* Nullpunkt, so liegt ein **intervallskaliertes Merkmal** vor. Da es allenfalls einen *willkürlichen* Nullpunkt gibt, lassen sich Verhältnisse nicht interpretieren. Die Ausprägungen sind quantitativ mittels reeller Zahlen darstellbar.
   - Das höchstskalierte Merkmal ist ein **verhältnisskaliertes Merkmal:** Zu den Eigenschaften eines intervallskalierten Merkmals kommt ein *natürlicher* (absoluter) Nullpunkt hinzu, sodass sich auch Verhältnisse sinnvoll interpretieren lassen.
2) Die Ordinalskala und die Verhältnisskala sind sogenannte **metrische Skalen** und fallen unter den Oberbegriff der **Kardinalskala.**
3) Nominalskalierte oder ordinalskalierte Merkmale nennen wir auch **qualitative Merkmale**, sofern es nur endlich viele Ausprägungen gibt und eine solche Ausprägung kein Ausmaß darstellt und keine Intensität angibt. Sonst spricht man von **quantitativen Merkmalen.**

Definition (Häufigkeit)**.**

1) Die Anzahl derjenigen Beobachtungen x_j, die mit $a_i \in \{a_1,a_2,\ldots,a_\ell\}$ übereinstimmen, heißt **absolute Häufigkeit** von a_i.
2) Der Anteil $f(a_i)$ an Beobachtungen x_j, die mit $a_i \in \{a_1,a_2,\ldots,a_\ell\}$ übereinstimmen, bezogen auf die Rohdaten, heißt **relative Häufigkeit** von a_i.

Definition (Empirische Verteilungsfunktion)**.** Sei $x \in \mathbb{R}$ und $I := \{i \in \mathbb{N} \mid a_i \leq x\}$. Die **empirische Verteilungsfunktion** $\hat{F}(x)$ ist definiert als

$$\hat{F}(x) := \sum_{i\in I} f(a_i)\,.$$

Definition (Lagemaß)**.** Für $n \in \mathbb{N}$ sei $L = \{x_1,\ldots,x_n\} \subset \mathbb{R}$ die Urliste.

1) Eine Funktion $\lambda:\ \mathbb{R}^n \longrightarrow \mathbb{R}$ heißt **additives Lagemaß**, falls für alle $a \in \mathbb{R}$

$$\lambda(x_1+a,\ldots,x_n+a) = \lambda(x_1,\ldots,x_n)+a.$$

2) Eine Funktion $\lambda : \mathbb{R}^n \longrightarrow \mathbb{R}$ heißt **multiplikatives Lagemaß**, falls für alle $a \in \mathbb{R}$

$$\lambda(x_1 \cdot a, \ldots, x_n \cdot a) = \lambda(x_1, \ldots, x_n) \cdot a.$$

Diese Eigenschaft heißt **Translationsäquivarianz.**

Definition (Arithmetisches Mittel)**.** Sei $L = \{x_1, \ldots, x_n\} \subset \mathbb{R}$ die Urliste. Das **arithmetische Mittel** $\overline{x}$ der Rohdaten ist

$$\overline{x} := \frac{1}{n} \cdot \sum_{i=1}^{n} x_i.$$

Definition (Median)**.** Für $n \in \mathbb{N}$ sei $L = \{x_1, \ldots, x_n\} \subset \mathbb{R}$ die Urliste. Eine Zahl $x_{\text{med}} \in \mathbb{R}$ heißt **Median** der Urliste L, falls

1) $x \leq x_{\text{med}}$ für mindestens die Hälfte der $x \in L$ gilt
2) und $x \geq x_{\text{med}}$ für mindestens die Hälfte der $x \in L$ gilt.

Definition (Quantil)**.** Für $n \in \mathbb{N}$ sei $L = \{x_1, \ldots, x_n\} \subset \mathbb{R}$ die Urliste und $p \in [0,1] \subset \mathbb{R}$. Eine Zahl $x_p \in \mathbb{R}$ heißt p-**Quantil**, falls

1) $x \leq x_p$ für mindestens einen Anteil p der $x \in L$ gilt
2) und $x \geq x_p$ für mindestens einen Anteil $1-p$ der $x \in L$ gilt.

Definition (Modus)**.** Als **Modus** bezeichnen wir die Ausprägung, welche am häufigsten auftritt. Wir schreiben auch x_{mod}.

Definition (Geometrisches Mittel)**.** Das **geometrische Mittel** $\overline{\xi}_{\text{geom}}$ der Faktoren $\xi_1, \ldots, \xi_n$ ist

$$\overline{\xi}_{\text{geom}} := \sqrt[n]{\prod_{i=1}^{n} \xi_i} = \sqrt[n]{\xi_1 \cdot \xi_2 \cdot \ldots \cdot \xi_n}.$$

Definition (Korrelationskoeffizient)**.** Der **Korrelationskoeffizient** r_{XY} der Wertepaare $(x_1,y_1), \ldots, (x_n,y_n) \in \mathbb{R}^2$ ist bestimmt durch

$$r_{XY} := \frac{\sum\limits_{i=1}^{n} (x_i - \overline{x}) \cdot (y_i - \overline{y})}{\sqrt{\sum\limits_{i=1}^{n} (x_i - \overline{x})^2} \cdot \sqrt{\sum\limits_{i=1}^{n} (y_i - \overline{y})^2}}.$$

Definition (Korrelation)**.** Zwei Merkmale X,Y heißen **unkorreliert,** falls $r_{XY} = 0$. Andernfalls heißen sie **korreliert.**

Definition (Varianz, Standardabweichung)**.** Sei X ein Merkmal mit beobachteten Ausprägungen $x_1, \ldots, x_n$.

1) Die Zahl

$$s_X^2 = \hat{\sigma}_X^2 := \frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - \overline{x})^2$$

heißt **empirische Varianz** des Merkmals X.

2) Die Zahl

$$\tilde{s}_X^2 := \frac{1}{n-1} \cdot \sum_{i=1}^{n} (x_i - \overline{x})^2$$

heißt **Stichprobenvarianz** des Merkmals X.

3) Die Zahl

$$s_X = \hat{\sigma}_X := \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - \overline{x})^2}$$

heißt **empirische Standardabweichung** des Merkmals X.

Definition (Empirische Kovarianz)**.** Seien $(x_1,y_1), \ldots, (x_n,y_n) \in \mathbb{R}^2$ die gemeinsam beobachteten Ausprägungen zweier Merkmale X,Y. Dann heißt die Zahl

$$s_{XY} := \frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - \overline{x}) \cdot (y_i - \overline{y})$$

empirische Kovarianz der Merkmale X und Y.

Bemerkung. Der Korrelationskoeffizient r_{XY} ist das Verhältnis der gemeinsamen Streuung zur Gesamtstreuung der beobachteten Ausprägungen der Merkmale X und Y um den arithmetischen Mittelwert $(\overline{x},\overline{y})$:

$$r_{XY} = \frac{s_{XY}}{s_X \cdot s_Y} .$$

Satz. Sind zwei Merkmale X,Y unabhängig, so sind sie auch unkorreliert. Die Umkehrung gilt im Allgemeinen nicht.

Bemerkung. Sind $(x_1,y_1), \ldots, (x_n,y_n) \in \mathbb{R}^2$ die gemeinsam beobachteten Ausprägungen der Merkmale X und Y, so ist der Korrelationskoeffizient

- gleich -1, falls ein gegensinniger linearer Zusammenhang zwischen X und Y besteht,
- gleich 1, falls ein gleichsinniger linearer Zusammenhang zwischen X und Y besteht, und
- gleich 0, falls kein (linearer) Zusammenhang zwischen X und Y besteht.

Modell (Regressionsmodell)**.** Seien X und Y zwei Merkmale. Eine Beziehung

$$Y = m(X) + \epsilon \quad \text{mit} \quad \epsilon \sim \mathrm{N}\big(0;\, \sigma^2\big)$$

heißt **Regressionsmodell.** Die zufällige Größe ϵ heißt **Fehlerterm** oder **Störgröße.** Ist $m(X) = \alpha + \beta \cdot X$ eine lineare Funktion, so sprechen wir von einem (klassischen) **linearen Regressionsmodell.** Es hat die (allgemeine) Form

$$Y = \alpha + \beta \cdot X + \epsilon \quad \text{mit } \alpha,\beta \in \mathbb{R} \text{ und } \epsilon \sim \mathrm{N}\big(0;\, \sigma^2\big) ,$$

wobei die Varianz $\sigma^2 \in \mathbb{R}$ in vielen Fällen nicht bekannt ist.

Satz. Es seien die Ausprägungen $(x_1,y_1), \ldots, (x_n,y_n) \in \mathbb{R}^2$ zweier metrisch skalierter Merkmale X und Y gemeinsam beobachtet worden. Unter der Annahme des linearen Regressionsmodells $Y = \alpha + \beta \cdot X + \epsilon$, wobei $\alpha,\beta \in \mathbb{R}$, ist die mit der Methode der kleinsten Quadrate ermittelte Regressionsgerade gegeben durch $\hat{y} = \hat{\alpha} + \hat{\beta} \cdot x$ mit den Kleinste-Quadrate-Schätzern

$$\hat{\beta} = \frac{s_{XY}}{s_X^2} \qquad \text{und} \qquad \hat{\alpha} = \overline{y} - \hat{\beta} \cdot \overline{x}.$$

Definition (Residuum)**.** Die Differenzen

$$\hat{\epsilon}_i := y_i - \hat{y}_i = y_i - \hat{\alpha} - \hat{\beta} \cdot x_i \quad \text{mit} \quad i \in \{1,\ldots,n\}$$

heißen **Residuen.**

Definition (Streuungen)**.** Wir definieren

1) die **Gesamtstreuung** SQT (sum of **sq**ared **t**otals) bzw. TSS (**t**otal **s**um of **s**quares):

$$\mathrm{TSS} = \mathrm{SQT} := \sum_{i=1}^{n} (y_i - \overline{y})^2 ,$$

2) die **Residualstreuung** SQR (sum of **sq**ared **r**esiduals) bzw. RSS (**r**esidual **s**um of **s**quares):

$$\mathrm{RSS} = \mathrm{SQR} := \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} \epsilon_i^2,$$

3) die **erklärte Streuung** SQE (sum of **sq**ares **e**xplained) bzw. ESS (**e**xplained **s**um of **s**quares):

$$\mathrm{ESS} = \mathrm{SQE} := \sum_{i=1}^{n} (\hat{y}_i - \overline{y})^2 .$$

Satz (Streuungszerlegung)**.** Es gilt: SQT = SQE + SQR bzw. TSS = ESS + RSS.

Definition (Determinationskoeffizient)**.** Die Zahl

$$R^2 := \frac{\mathrm{SQE}}{\mathrm{SQT}} = \frac{\mathrm{ESS}}{\mathrm{TSS}} \in [0,\,1] \subset \mathbb{R}$$

heißt **Determinationskoeffizient** des linearen Regressionsmodells.

Satz. Es gilt $R^2 = r_{XY}^2$.

Definition (Stichprobenvariable)**.** Die Zufallsvariablen $X_1,\ldots,X_n$ nennen wir auch **Stichprobenvariablen**, falls sie folgenden Bedingungen genügen:

1) $X_1, \ldots, X_n$ sind unabhängig voneinander.

2) $X_1, \ldots, X_n$ beschreiben n Wiederholungen ein und des selben Zufallsexperiments mit Ergebnismenge Ω, das wir mathematisch durch die Zufallsvariable $X : \Omega \longrightarrow \mathbb{R}$ modellieren.

Definition (Punktschätzer)**.** Seien $X_1, \ldots, X_n$ Stichprobenvariablen und seien $x_1, \ldots, x_n$ Realisierungen der Stichprobenvariablen $X_1, \ldots, X_n$. Unter einem **Punktschätzer** eines Parameters $\theta \in \Theta \subseteq \mathbb{R}$ verstehen wir eine Funktion T der Form

$$T : \mathbb{R}^n \longrightarrow \mathbb{R}, \quad (x_1, \ldots, x_n) \mapsto T(x_1, \ldots, x_n)\,.$$

Den Funktionswert $\hat{\theta} = T(x_1, \ldots, x_n) \in \mathbb{R}$ nennen wir **Schätzwert.**

Bemerkung.

1) Das arithmetische Mittel ist ein Punktschätzer des Erwartungswerts.

2) Die empirische Varianz und die Stichprobenvarianz sind Punktschätzer der Varianz.

Definition (Unverzerrtheit, Erwartungstreue)**.** Einen Punktschätzer T nennen wir **erwartungstreu** oder **unverzerrt,** wenn der Erwartungswert des Punktschätzers gleich θ ist unter der Voraussetzung, dass der wahre Parameterwert $\theta \in \Theta$ zugrunde liegt. In diesem Fall schreiben wir auch kurz $\mathrm{E}_\theta(T) = \theta$. Andernfalls nennen wir den Punktschätzer T **verzerrt.**

Definition (Nullhypothese, Alternative, Testproblem)**.**

1) Die **Nullhypothese** H_0 ist eine Annahme über die Wahrscheinlichkeitsverteilung einer oder mehrerer Zufallsvariablen.

2) Die **Alternative** bezeichnet eine Menge von alternativen Annahmen bezüglich der Nullhypothese.

3) Das **Testproblem** besteht darin, zwischen Nullhypothese und Alternative zu entscheiden.

Bemerkung.

1) Als Nullhypothese wählen wir stets die zu widerlegende Annahme.

2) Ein Verwerfen der Nullhypothese kommt dem gewünschten Beweis der Alternative gleich.

3) Aussagen über Nullhypothese und Alternative sind Entscheidungen des Testproblems und damit immer *Aussagen über die Grundgesamtheit* und nicht über die Stichprobe.

Bemerkung.

1) Wir gehen beim Testen davon aus, dass zufällig erhobene Daten in der Stichprobe das für den untersuchten Sachverhalt typische Verhalten zeigen: Ereignisse mit einer geringen Wahrscheinlichkeit finden wir selten in der Stichprobe, während sich Ereignisse mit einer hohen Wahrscheinlichkeit häufig in der Stichprobe finden.

2) Wir bestimmen nun den Verwerfungsbereich so, dass es unwahrscheinlich ist, dass die Prüfgröße im Verwerfungsbereich liegt, wenn die Nullhypothese wahr ist. Nur so lässt sich eine möglichst verlässliche Aussage über die Nullhypothese und die Alternative treffen und damit das Testproblem entscheiden. Dennoch kann eine auf dieser Grundlage getroffene Entscheidung unter Umständen falsch sein.

Definition (Fehler 1. Art, Fehler 2. Art)**.** Wir unterscheiden zwei Fehler beim Testen.

1) Wir sprechen von einem **Fehler 1. Art,** wenn wir die Nullhypothese ablehnen, obwohl sie richtig ist.

2) Wir sprechen von einem **Fehler 2. Art,** wenn wir die Nullhypothese nicht ablehnen, obwohl sie falsch ist.

Bemerkung.

1) Wir bestimmen den Verwerfungsbereich so, dass die *bedingte Wahrscheinlichkeit* für einen Fehler 1. Art unter der Nullhypothese kleiner ist als eine vorgegebene Zahl $\alpha \in [0, 1] \subset \mathbb{R}$, also

$$P(X \in \mathbb{R} \setminus [c, d] \mid H_0) = P(X < c \mid H_0) + P(X > d \mid H_0) \overset{!}{<} \alpha.$$

Die Schranke α nennen wir **Signifikanzniveau** des statistischen Hypothesentests.

2) Einen Fehler erster Art können wir überhaupt nur dann begehen, wenn wir die Nullhypothese fälschlicherweise verwerfen.

3) Die Wahrscheinlichkeit, einen Fehler zweiter Art zu begehen, können wir erst dann bestimmen, wenn wir die wahre Wahrscheinlichkeitsverteilung kennen, welche den Daten zugrunde liegt.

4) Einen Fehler zweiter Art können wir überhaupt nur dann begehen, wenn wir die Nullhypothese fälschlicherweise nicht verwerfen.

5) Können wir die Nullhypothese zum Signifikanzniveau α *nicht verwerfen,* bedeutet dies, dass uns nicht genügend Daten vorliegen, die der Nullhypothese widersprechen. Wir haben damit die Nullhypothese weder bewiesen noch

widerlegt; der Hypothesentest liefert uns in diesem Fall *keine Aussage* über Nullhypothese und Alternative.

Bemerkung. (Binomialtest)**.** Voraussetzung für den **Binomialest:**

1) die Stichprobenvariablen $X_1, \ldots, X_n \sim \text{Binom}(1;\, p)$ sind *unabhängig*,
2) die Grundwahrscheinlichkeit p ist unbekannt,
3) das Signifikanzniveau α ist festgelegt.

Nur wennn alle drei Voraussetzungen erfüllt sind, können wir mithilfe des Binomialtests die folgenden Testprobleme entscheiden.

- TP-1: Nullhypothese $H_0 : \ p = p_0$ gegen die Alternative $H_1 : \ p \neq p_0$,
- TP-2: Nullhypothese $H_0 : \ p \leq p_0$ gegen die Alternative $H_1 : \ p > p_0$,
- TP-3: Nullhypothese $H_0 : \ p \geq p_0$ gegen die Alternative $H_1 : \ p < p_0$.

Die Prüfgröße ist

$$X := \sum_{i=1}^{n} X_i = X_1 + \ldots + X_n.$$

Unter H_0 ist $X \sim \text{Binom}(n;\, p_0)$. Wir bezeichnen mit x die beobachtete Realisierung von X. Im Testproblem

- TP-1 spricht $x < c$ oder $x > d$ nicht für die Nullhypothese $H_0 : \ p = p_0$.
- TP-2 spricht $x > d$ nicht für die Nullhypothese $H_0 : \ p \leq p_0$.
- TP-3 spricht $x < c$ nicht für die Nullhypothese $H_0 : \ p \geq p_0$.

Wir bestimmen mit der $\text{Binom}(n;\, p_0)$-Verteilungsfunktion die Grenzen c und d: Im Testproblem

- TP-1 ist $c \in \{0, \ldots, n\}$ diejenige ganze Zahl, sodass

$$P(X < c \mid H_0) < \frac{\alpha}{2} \quad \text{und} \quad P(X \leq c \mid H_0) \geq \frac{\alpha}{2}\,,$$

  und $d \in \{0, \ldots, n\}$ ist diejenige ganze Zahl, sodass

$$P(X > d \mid H_0) < \frac{\alpha}{2} \quad \text{und} \quad P(X \geq d \mid H_0) \geq \frac{\alpha}{2}\,.$$

  Beobachten wir $X < c$ oder $x > d$, dann lehnen wir $H_0 : \ p = p_0$ ab.

- TP-2 ist $d \in \{0, \ldots, n\}$ diejenige ganze Zahl, sodass

$$P(X > d \mid H_0) < \alpha \quad \text{und} \quad P(X \geq d \mid H_0) \geq \alpha.$$

  Beobachten wir $X > d$, dann lehnen wir $H_0 : \ p \leq p_0$ ab.

- TP-3 ist $c \in \{0, \ldots, n\}$ diejenige ganze Zahl, sodass

$$P(X < c \mid H_0) < \alpha \quad \text{und} \quad P(X \leq c \mid H_0) \geq \alpha.$$

  Beobachten wir $X < c$, dann lehnen wir $H_0 : \ p \geq p_0$ ab.

Modell (Gauß-Test)**.** Voraussetzung für den **Gauß-Test** ist, dass

- die Stichprobenvariablen $X_1, \ldots, X_n \sim \text{N}\big(\mu;\, \sigma^2\big)$ *unabhängig* sind,
- die Varianz σ^2 *bekannt* ist,
- der Erwartungswert μ *unbekannt* ist und
- wir ein Signifikanzniveau $\alpha \in [0, 1] \subset \mathbb{R}$ gewählt haben.

Nur wenn alle vier Voraussetzungen erfüllt sind, können wir die folgenden Testprobleme mithilfe eines Gauß-Tests entscheiden.

- TP-1: Nullhypothese $H_0 : \ \mu = \mu_0$ gegen die Alternative $H_1 : \ \mu \neq \mu_0$,
- TP-2: Nullhypothese $H_0 : \ \mu \leq \mu_0$ gegen die Alternative $H_1 : \ \mu > \mu_0$,
- TP-3: Nullhypothese $H_0 : \ \mu \geq \mu_0$ gegen die Alternative $H_1 : \ \mu < \mu_0$.

Wir überlegen uns, welche beobachteten Realisierungen $\overline{x}$ der Prüfgröße

$$\overline{X} := \frac{1}{n} \cdot \sum_{i=1}^{n} X_i$$

nicht für die Nullhypothese sprechen: Im Testproblem

- TP-1 spricht eine zu große absolute Abweichung des arithmetischen Stichprobenmittels $\overline{x}$ von μ_0, also $|\overline{x}-\mu_0| > c_1$ mit $c_1 > 0$, *nicht* für die Nullhypothese $H_0 : \ \mu = \mu_0$.
- TP-2 spricht ein zu großes arithmetisches Stichprobenmittel $\overline{x}$, also $\overline{x}-\mu_0 > c_2$ mit $c_2 > 0$, *nicht* für die Nullhypothese $H_0 : \ \mu \leq \mu_0$.
- TP-3 spricht ein zu kleines arithmetisches Stichprobenmittel $\overline{x}$, also $\overline{x}-\mu_0 < c_3$ mit $c_3 < 0$, *nicht* für die Nullhypothese $H_0 : \ \mu \geq \mu_0$.

Unter H_0 ist $\overline{X} \sim \mathrm{N}\big(\mu;\ {}^{\sigma^2}\!/_{n}\big)$. Bezeichnet z_α das α-Quantil der Standardnormalverteilung, so ergibt sich im Testproblem

- TP-1: $c_1 = \dfrac{\sigma}{\sqrt{n}} \cdot z_{1-\alpha/2}$. Beobachten wir $|\overline{x}-\mu_0| > c_1$, dann lehnen wir $H_0 : \ \mu = \mu_0$ ab.
- TP-2: $c_2 = \dfrac{\sigma}{\sqrt{n}} \cdot z_{1-\alpha}$. Beobachten wir $\overline{x}-\mu_0 > c_2$, dann lehnen wir $H_0 : \ \mu \leq \mu_0$ ab.
- TP-3: $c_3 = \dfrac{\sigma}{\sqrt{n}} \cdot z_{\alpha}$. Beobachten wir $\overline{x}-\mu_0 < c_3$, dann lehnen wir $H_0 : \ \mu \geq \mu_0$ ab.

Bemerkung (t-Test)**.** Voraussetzung für den t-Test ist, dass

- die Stichprobenvariablen $X_1, \ldots, X_n \sim \mathrm{N}\big(\mu;\ \sigma^2\big)$ *unabhängig* sind,
- der Erwartungswert μ *unbekannt* ist,
- die Varianz σ^2 *unbekannt* ist und
- wir ein Signifikanzniveau $\alpha \in [0,1] \subset \mathbb{R}$ gewählt haben.

Nur wenn alle vier Voraussetzungen erfüllt sind, können wir die folgenden Testprobleme mithilfe eines t-Tests entscheiden.

- TP-1: Nullhypothese $H_0 : \ \mu = \mu_0$ gegen die Alternative $H_1 : \ \mu \neq \mu_0$,
- TP-2: Nullhypothese $H_0 : \ \mu \leq \mu_0$ gegen die Alternative $H_1 : \ \mu > \mu_0$,
- TP-3: Nullhypothese $H_0 : \ \mu \geq \mu_0$ gegen die Alternative $H_1 : \ \mu < \mu_0$.

Das arithmetische Mittel

$$\overline{X} := \frac{1}{n} \cdot \sum_{i=1}^{n} X_i$$

dient uns als Prüfgröße. Unter H_0 ist

$$T := \sqrt{n} \cdot \frac{\overline{X}-\mu_0}{\tilde{s}_X} \sim \mathrm{t}_{n-1}.$$

Wir überlegen uns, welche beobachteten Realisierungen $\overline{x}$ der Prüfgröße $\overline{X}$ *nicht* für die Nullhypothese sprechen: Im Testproblem

- TP-1 spricht eine zu große absolute Abweichung des arithmetischen Stichprobenmittels $\overline{x}$ von μ_0, also $|\overline{x}-\mu_0| > c_1$ mit $c_1 > 0$, *nicht* für die Nullhypothese $H_0 : \ \mu = \mu_0$.
- TP-2 spricht ein zu großes arithmetisches Stichprobenmittel $\overline{x}$, also $\overline{x}-\mu_0 > c_2$ mit $c_2 > 0$, *nicht* für die Nullhypothese $H_0 : \ \mu \leq \mu_0$.
- TP-3 spricht ein zu kleines arithmetisches Stichprobenmittel $\overline{x}$, also $\overline{x}-\mu_0 < c_3$ mit $c_3 < 0$, *nicht* für die Nullhypothese $H_0 : \ \mu \geq \mu_0$.

Wir können mit den Quantilen der t-Verteilung arbeiten: Bezeichnet $t_{n-1;\,\alpha}$ das α-Quantil der t-Verteilung mit $n-1$ Freiheitsgraden, so ergibt sich im Testproblem

- TP-1: $c_1 = \dfrac{\tilde{s}_X}{\sqrt{n}} \cdot t_{n-1;\,1-\alpha/2}$. Beobachten wir $|\overline{x}-\mu_0| > c_1$, dann lehnen wir $H_0 : \ \mu = \mu_0$ ab.
- TP-2: $c_2 = \dfrac{\tilde{s}_X}{\sqrt{n}} \cdot t_{n-1;\,1-\alpha}$. Beobachten wir $\overline{x}-\mu_0 > c_2$, dann lehnen wir $H_0 : \ \mu \leq \mu_0$ ab.
- TP-3: $c_3 = \dfrac{\tilde{s}_X}{\sqrt{n}} \cdot t_{n-1;\,\alpha}$. Beobachten wir $\overline{x}-\mu_0 < c_3$, dann lehnen wir $H_0 : \ \mu \geq \mu_0$ ab.