3 Grundlagen statistischer Tests (Kap. 8 IS)

3.1 Beispiel zum Hypothesentest

Beispiel:

- Betrachtet wird eine Abfüllanlage für Mineralwasser mit dem Sollgewicht $\mu_0 = 1000g$ und bekannter Standardabweichung des Abfüllgewichts von $\sigma = 3g$.
- Beobachtet wird die einfache Zufallsstichprobe (iid-Stichprobe)

$$x_1 = 1002g,\, x_2 = 1007g,\, x_3 = 997g,\, x_4 = 1014g$$

vom Umfang $n = 4$ mit dem Mittelwert

$$\bar{x} = \frac{1}{4} \sum_{i=1}^{4} x_i = 1005g.$$

Interessierende Fragestellung:

- Wird die **Hypothese** $\mu = \mu_0 = 1000g$ durch die Beobachtung von $\bar{x} = 1005g$ erschüttert?
- Weicht also $\bar{x}$ statistisch signifikant von μ_0 ab?

Stochastisches Modell:

- Die Zufallsvariable Y beschreibe das Abfüllgewicht der Anlage in Gramm, wobei $Y \sim N(\mu, \sigma^2)$ mit unbekannten $\mu \in \mathbb{R}$ und bekannter Varianz $\sigma^2 > 0$ gelte.
- Es liegt eine einfache Zufallsstichprobe (iid-Stichprobe) vom Umfang n zu Y vor, d. h. $X_i \overset{\text{i.i.d.}}{\sim} N(\mu, \sigma^2)$ für $i = 1, \ldots, n$.

Implikationen und Testdurchführung:

- In diesem Modell ist das zufällige Stichprobenmittel $\bar{X} \overset{\text{def}}{=} \frac{1}{n} \sum_{i=1}^{n} X_i$ normalverteilt,

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right).$$

- Um eine Entscheidung bzgl. der Hypothese zu fällen, wird im Folgenden die Zufallsvariable

$$T(X) = T(X_1, \ldots, X_n) \overset{\text{def}}{=} \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n}$$

betrachtet, die auch als **Prüfgröße**, **Testfunktion** oder **Teststatistik** bezeichnet wird.

- Diese Zufallsvariable ist normalverteilt und es gilt

$$T(X) = \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n} \sim N\left(\frac{\mu - \mu_0}{\sigma} \sqrt{n}, 1\right).$$

- Falls die Hypothese $\mu = \mu_0$ richtig wäre, dann wäre die Prüfgröße $T(X)$ standardnormalverteilt,

$$T(X) \overset{\mu=\mu_0}{\sim} N(0,1)$$

und es gilt

$$P\left(T(X) \in [u_{\frac{\alpha}{2}}, u_{1-\frac{\alpha}{2}}]\right) = 1-\alpha \qquad \text{und} \qquad P\left(T(X) \notin [u_{\frac{\alpha}{2}}, u_{1-\frac{\alpha}{2}}]\right) = \alpha.$$

Dabei bezeichnet $0 < \alpha < 1$ eine vorgegebene (in der Regel) kleine **Wahrscheinlichkeit**; übliche Werte für α sind $1\%, 5\%, 10\%$. Es bezeichnet $u_{\frac{\alpha}{2}}$ das $\alpha/2$-Quantil und $u_{1-\frac{\alpha}{2}}$ das $(1-\alpha/2)$-Quantil der Standardnormalverteilung, wobei $u_{\frac{\alpha}{2}} = -u_{1-\frac{\alpha}{2}}$ gilt.

- Falls $\mu > \mu_0$ wäre (Hypothese falsch), dann sind große Werte von $T(X)$ wahrscheinlicher als im Fall $\mu = \mu_0$.

- Falls $\mu < \mu_0$ wäre (Hypothese auch falsch), dann sind kleine Werte von $T(X)$ wahrscheinlicher als im Fall $\mu = \mu_0$.

- Im Beispiel ergibt sich mit $\bar{x} = 1005g$, $\mu_0 = 1000g$, $\sigma = 3g$ und $n = 4$ für die Prüfgröße $T(X)$ die Realisation

$$T(x) = T(x_1, \ldots, x_n) = \frac{1005 - 1000}{3}\sqrt{4} = \frac{10}{3}.$$

Ist $\alpha = 0.01$ gegeben, dann ist $u_{\frac{\alpha}{2}} = u_{0.005} = -2.58$ und $u_{1-\frac{\alpha}{2}} = u_{0.995} = 2.58$. Da $T(x) = \frac{10}{3} > 2.58$, wird die Hypothese $\mu = \mu_0$ in diesem Beispiel auf dem **Signifikanzniveau** $\alpha = 0.01$ verworfen.

3.2 Allgemeine Teststruktur und Grundbegriffe

- **Nullhypothese** H_0 und **Gegenhypothese** (auch Alternativhypothese) H_1: Zerlegung des Parameterraums Γ in zwei disjunkte nichtleere Teilmengen Γ_0 und Γ_1, d. h. $\Gamma_0 \cap \Gamma_1 = \emptyset$, $\Gamma_0 \neq \emptyset$ und $\Gamma_1 \neq \emptyset$. Oft vollständige Zerlegung, so dass $\Gamma = \Gamma_0 \cup \Gamma_1$.
  Testproblem lautet somit:

$$H_0: \ \gamma \in \Gamma_0 \qquad \text{vs.} \qquad H_1: \ \gamma \in \Gamma_1$$

- **Test** und **Entscheidungsfunktion** δ: Entscheidungsfunktion δ: $\mathcal{X} \to \{d_0, d_1\}$ wird Test genannt; $\mathcal{X}$ ... Stichprobenraum, d_0 ... H_0 wird nicht verworfen, d_1 ... H_0 wird verworfen

- **Prüfgröße** (auch Testfunktion, Teststatistik) $T(X)$: Stichprobenfunktion $T(X) = T(X_1, \ldots, X_n)$, die als Hilfsmittel zur Entscheidungsfindung dient, Realisationen: $T(x) = T(x_1, \ldots, x_n) \in \mathbb{R}$

- **Annahmebereich** A_δ und **Ablehnungsbereich** (auch Ablehnbereich, kritischer Bereich, Verwerfungsbereich) K_δ:

$$A_\delta \overset{\text{def}}{=} \{x \in \mathcal{X} \mid \delta(x) = d_0\} \qquad \text{und} \qquad K_\delta \overset{\text{def}}{=} \{x \in \mathcal{X} \mid \delta(x) = d_1\}$$

  Es gilt $\mathcal{X} = A_\delta \cup K_\delta$, $A_\delta \cap K_\delta = \emptyset$.
  Alternative Definition: $A_\delta^* \overset{\text{def}}{=} \{T(x) \in \mathbb{R} \mid \delta(x) = d_0\}$ und $K_\delta^* \overset{\text{def}}{=} \{T(x) \in \mathbb{R} \mid \delta(x) = d_1\}$

- **Fehler 1. Art** und **Fehler 2. Art**:
  - Der Fehler, dass H_0 abgelehnt wird (Entscheidung d_1), obwohl H_0 richtig ist ($\gamma \in \Gamma_0$), heißt Fehler 1. Art.
  - Der Fehler, dass H_0 nicht abgelehnt wird (Entscheidung d_0), obwohl H_0 falsch ist ($\gamma \in \Gamma_1$), heißt Fehler 2. Art.
  - Die Wahrscheinlichkeit, den Fehler 1. Art zu begehen, ist

$$P_I(\delta, \gamma) \overset{\text{def}}{=} P_\gamma(\delta(X) = d_1) = P_\gamma(X \in K_\delta) = P_\gamma(T(X) \in K_\delta^*) \qquad \text{für} \qquad \gamma \in \Gamma_0$$

    und heißt **Fehlerwahrscheinlichkeit 1. Art**.
  - Die Wahrscheinlichkeit, den Fehler 2. Art zu begehen, ist

$$P_{II}(\delta, \gamma) \overset{\text{def}}{=} P_\gamma(\delta(X) = d_0) = P_\gamma(X \in A_\delta) = P_\gamma(T(X) \in A_\delta^*) \qquad \text{für} \qquad \gamma \in \Gamma_1$$

    und heißt **Fehlerwahrscheinlichkeit 2. Art**.
  - Beide Fehlerwahrscheinlichkeiten sind im Allgemeinen nicht konstant und variieren mit n.
  - $P_I(\delta, \gamma)$ verringert sich durch Verkleinerung von K_δ (Vergrößerung von A_δ) und $P_{II}(\delta, \gamma)$ verringert sich durch Verkleinerung von A_δ (Vergrößerung von K_δ). In der Regel führt daher Verringerung von $P_I(\delta, \gamma)$ zur Erhöhung von $P_{II}(\delta, \gamma)$ und umgekehrt.
- **Signifikanzniveau** α: vorgegebene Oberschranke für Fehlerwahrscheinlichkeit 1. Art, $0 < \alpha < 1$, übliche Werte: 1%, 5%, 10%,

- **Gütefunktion** und **Operationscharakteristik** (auch OC-Kurve):
  - Die Funktion
$$G_\delta(\gamma) \stackrel{\text{def}}{=} P_\gamma(\delta(X) = d_1), \qquad \gamma \in \Gamma$$
    heißt Gütefunktion (*power*) eines Tests $\delta : \mathcal{X} \to \{d_0, d_1\}$ und gibt zu jedem γ die Wahrscheinlichkeit an, H_0 zu verwerfen.
  - Für $\gamma \in \Gamma_0$ ist $G_\delta(\gamma)$ die Fehlerwahrscheinlichkeit 1. Art. Für $\gamma \in \Gamma_1$ ist $G_\delta(\gamma)$ die Wahrscheinlichkeit für die korrekte Entscheidung H_0 zu verwerfen.
  - Die Funktion
$$L_\delta(\gamma) \stackrel{\text{def}}{=} 1 - G_\delta(\gamma) = P_\gamma(\delta(X) = d_0), \qquad \gamma \in \Gamma$$
    heißt Operationscharakteristik eines Tests $\delta : \mathcal{X} \to \{d_0, d_1\}$ und gibt zu jedem γ die Wahrscheinlichkeit an, H_0 nicht zu verwerfen.
  - Für $\gamma \in \Gamma_0$ ist $L_\delta(\gamma)$ die Wahrscheinlichkeit für die korrekte Entscheidung H_0 nicht zu verwerfen. Für $\gamma \in \Gamma_1$ ist $L_\delta(\gamma)$ die Fehlerwahrscheinlichkeit 2. Art.
  - Beide Funktionen hängen auch vom Stichprobenumfang n ab.

3.3 Tests zum Niveau α

Es sei $0 < \alpha < 1$.

- Ein Test $\delta : \mathcal{X} \to \{d_0, d_1\}$ mit der Eigenschaft

  $$P_I(\delta,\gamma) = P_\gamma(\delta(X) = d_1) = P_\gamma(X \in K_\delta) = P_\gamma(T(X) \in K_\delta^*) \leq \alpha \qquad \text{für alle} \qquad \gamma \in \Gamma_0$$

  heißt **Test zum Niveau** α, d. h. die Gütefunktion $G_\delta(\gamma)$ überschreitet im gesamten Bereich $\gamma \in \Gamma_0$ nie den Wert α.

- Ein Test zum Niveau α heißt **unverfälscht**, falls

  $$G_\delta(\gamma) \geq \alpha \qquad \text{für alle} \qquad \gamma \in \Gamma_1.$$

- Ein Test δ zum Niveau α heißt **gleichmäßig bester Test zum Niveau** α, falls für jeden anderen Test δ' zum Niveau α gilt, dass

  $$P_{II}(\delta,\gamma) = P_\gamma(\delta(X) = d_0) \leq P_{II}(\delta',\gamma) = P_\gamma(\delta'(X) = d_0) \qquad \text{für alle} \qquad \gamma \in \Gamma_1.$$

  D. h. für die Operationscharakteristiken gilt $L_\delta(\gamma) \leq L_{\delta'}(\gamma)$ im gesamten Bereich $\gamma \in \Gamma_1$.

4 Tests für die Parameter der Normalverteilung

Im Folgenden werden der **Gauß-Test**, der t-**Test** und ein **Varianz-Test** vorgestellt.

Für alle drei Tests müssen folgende Voraussetzungen erfüllt sein:

- Es sei $Y \sim N(\mu, \sigma^2)$ mit $\mu \in \mathbb{R}$ und Varianz $\sigma^2 > 0$.
- Es liegt eine einfache Zufallsstichprobe (iid-Stichprobe) vom Umfang n zu Y vor, d. h. $X_i \overset{\text{i.i.d.}}{\sim} N(\mu, \sigma^2)$ für $i = 1, \ldots, n$.

Der Gauß-Test ist ein Test über den Erwartungswert μ, falls die Varianz σ^2 bekannt ist.

Der t-Test ist ein Test über den Erwartungswert μ, falls die Varianz σ^2 unbekannt ist und geschätzt werden muss.

Der vorgestellte Varianz-Test ist ein Test über die Varianz σ^2 der Normalverteilung, falls der Erwartungswert μ der Normalverteilung unbekannt ist und geschätzt werden muss.

4.1 Gauß-Test

Zweck: Test über den Parameter μ einer Normalverteilung bei bekannter Varianz σ^2

Voraussetzung: $X_i \overset{\text{i.i.d.}}{\sim} N(\mu, \sigma^2)$ für $i = 1, \ldots, n$

Gegeben: Signifikanzniveau $\alpha \in]0, 1[$, Varianz $\sigma^2 > 0$, $\mu_0 \in \mathbb{R}$

4.1.1 Zweiseitiger Gauß-Test

Das in Abschnitt 3.1 diskutierte Beispiel ist ein zweiseitiger Gauß-Test.

Null- und Gegenhypothese: $H_0: \ \mu = \mu_0$ vs. $H_1: \ \mu \neq \mu_0$

Prüfgröße: $T(X) \overset{\text{def}}{=} \frac{\bar{X}-\mu_0}{\sigma}\sqrt{n} \sim N\left(\frac{\mu-\mu_0}{\sigma}\sqrt{n}, 1\right)$

Testverteilung: $T(X) \overset{\mu=\mu_0}{\sim} N(0, 1)$

Die Testverteilung ist die Verteilung der Prüfgröße, falls $\mu = \mu_0$ und somit beim zweiseitigen Test die Hypothese H_0 richtig ist. Die Testverteilung ist beim Gauß-Test die Standardnormalverteilung, die auch Gauß-Verteilung heißt, und dem Test den Namen gibt.

Ablehnungsbereich: $K^*_\delta =]-\infty, -u_{1-\frac{\alpha}{2}}[\, \cup \,]u_{1-\frac{\alpha}{2}}, \infty[$

Im Gauß-Test ist $u_{1-\frac{\alpha}{2}}$ das $(1-\frac{\alpha}{2})$-Quantil der Standardnormalverteilung. Da diese Verteilung eine Dichtefunktion besitzt, die symmetrisch um Null ist, gilt $u_{\frac{\alpha}{2}} = -u_{1-\frac{\alpha}{2}}$. Beim zweiseitigen Gauß-Test gilt $P(T(X) \in K^*_\delta) = \alpha$.

Testentscheidung:

$$\delta(x) = \begin{cases} d_0, & \text{falls} \quad T(x) \notin K^*_\delta \\ d_1, & \text{falls} \quad T(x) \in K^*_\delta \end{cases}$$

Dabei ist $x \overset{\text{def}}{=} (x_1, \ldots, x_n)$ eine Realisation des Stichprobenvektors $X \overset{\text{def}}{=} (X_1, \ldots, X_n)$ und $T(x)$ eine Realisation der Prüfgröße $T(X)$.

4.1.2 Einseitige Gauß-Tests

Bei einseitigen und zweiseitigen Gauß-Tests stimmen Prüfgröße, Testverteilung und das Treffen der Testentscheidung überein, nur Hypothesenpaare und Ablehnungsbereiche sind verschieden.

Null- und Gegenhypothese: $H_0: \ \mu = \mu_0$ oder $\mu \geq \mu_0$ vs. $H_1: \ \mu < \mu_0$
Ablehnungsbereich: $K_\delta^* =]-\infty, -u_{1-\alpha}[$
Dabei ist $u_{1-\alpha}$ das $(1-\alpha)$-Quantil der Standardnormalverteilung. Da diese Verteilung eine Dichtefunktion besitzt, die symmetrisch um Null ist, gilt $u_\alpha = -u_{1-\alpha}$.

Null- und Gegenhypothese: $H_0: \ \mu = \mu_0$ oder $\mu \leq \mu_0$ vs. $H_1: \ \mu > \mu_0$
Ablehnungsbereich: $K_\delta^* =]u_{1-\alpha}, \infty[$

Bei ein- oder zweiseitigen Gauß-Tests mit **einfacher Nullhypothese** $\mu = \mu_0$ gilt

$$P_\mu(T(X) \in K_\delta^*) = \alpha \qquad \text{für} \qquad \mu = \mu_0.$$

Liegt beim einseitigen Gauß-Test eine **zusammengesetzte Nullhypothese** $\mu \geq \mu_0$ vor, dann gilt

$$P_\mu(T(X) \in K_\delta^*) \leq \alpha \qquad \text{für alle} \qquad \mu \geq \mu_0,$$

liegt eine **zusammengesetzte Nullhypothese** der Form $\mu \leq \mu_0$ vor, dann gilt

$$P_\mu(T(X) \in K_\delta^*) \leq \alpha \qquad \text{für alle} \qquad \mu \leq \mu_0.$$

4.1.3 Beispiel für DAX-Renditen

- Zufallsvariable Y beschreibe die stetige Jahresrendite des DAX
- Annahme des Standardmodells für stetige Renditen: $Y \sim N(\mu, \sigma^2)$ mit $\mu \in \mathbb{R}$ und $\sigma^2 > 0$
- Erwartungswert μ unbekannt; Volatilität (p. a.) mit $\sigma = 25\%$ bekannt
- einfache Zufallsstichprobe vom Umfang $n = 25$ Jahre zu Y: $X_t \overset{\text{i.i.d.}}{\sim} N(\mu, \sigma^2)$, $t = 1, \ldots, n$
- beobachtete durchschnittliche Jahresrendite des DAX über 25 Jahre: $\bar{x} = 10\%$

Testfragestellung: Ist $\bar{x}$ signifikant ($\alpha = 5\%$) von Null verschieden?
Hypothesen: $H_0 : \mu = \mu_0$ vs. $H_1 : \mu \neq \mu_0$ mit $\mu_0 = 0$
Realisation der Prüfgröße: $T(x) = \frac{\bar{x}-\mu_0}{\sigma}\sqrt{n} = \frac{0.1-0}{0.25}\sqrt{25} = 2$
Ablehnungsbereich: Es ist $u_{1-\frac{\alpha}{2}} = u_{0.975} = 1.96$ und damit $K_\delta^* =]-\infty, -1.96[\cup]1.96, \infty[$.
Testentscheidung: H_0 wird verworfen, da $T(x) \in K_\delta^*$.

Testfragestellung: Ist $\bar{x}$ signifikant ($\alpha = 5\%$) größer als 4%?
Hypothesen: $H_0 : \mu = \mu_0$ vs. $H_1 : \mu > \mu_0$ mit $\mu_0 = 0.04$
Realisation der Prüfgröße: $T(x) = \frac{\bar{x}-\mu_0}{\sigma}\sqrt{n} = \frac{0.1-0.04}{0.25}\sqrt{25} = 1.2$
Ablehnungsbereich: Es ist $u_{1-\alpha} = u_{0.95} = 1.65$ und damit $K_\delta^* =]1.65, \infty[$.
Testentscheidung: H_0 wird nicht verworfen, da $T(x) \notin K_\delta^*$.

4.1.4 Eigenschaften des Gauß-Tests

Es wird der einseitige Gauß-Test mit dem folgenden Hypothesenpaar betrachtet:

$$H_0: \ \mu \leq \mu_0 \qquad \text{vs.} \qquad H_1: \ \mu > \mu_0 \qquad \text{mit gegebenen} \qquad \mu_0 \in \mathbb{R}.$$

Die **Gütefunktion** dieses Gauß-Tests ist für alle $\mu \in \mathbb{R}$ gegeben durch

$$\begin{aligned}
G_\delta(\mu) &\overset{\text{def}}{=} P_\mu(\delta(X) = d_1) \\
&= P_\mu(T(X) \in K_\delta^*) \\
&= P_\mu\left(\frac{\bar{X} - \mu_0}{\sigma}\sqrt{n} > u_{1-\alpha}\right) \\
&= P_\mu\left(\frac{\bar{X}}{\sigma}\sqrt{n} > u_{1-\alpha} + \frac{\mu_0}{\sigma}\sqrt{n}\right) \\
&= P_\mu\left(\frac{\bar{X} - \mu}{\sigma}\sqrt{n} > u_{1-\alpha} - \frac{\mu - \mu_0}{\sigma}\sqrt{n}\right) \\
&= 1 - P_\mu\left(\frac{\bar{X} - \mu}{\sigma}\sqrt{n} \leq u_{1-\alpha} - \frac{\mu - \mu_0}{\sigma}\sqrt{n}\right) \\
&= 1 - \Phi\left(u_{1-\alpha} - \frac{\mu - \mu_0}{\sigma}\sqrt{n}\right),
\end{aligned}$$

wobei Φ die Verteilungsfunktion der Standardnormalverteilung bezeichnet.

Es gilt

$$G_\delta(\mu) < G_\delta(\mu_0) \qquad \text{für alle} \qquad \mu < \mu_0$$

und

$$\max_{\mu \leq \mu_0} G_\delta(\mu) = G_\delta(\mu_0) = \alpha,$$

so dass der Test das **Niveau** α hat. Wegen

$$G_\delta(\mu) = 1 - \Phi\left(u_{1-\alpha} - \frac{\mu - \mu_0}{\sigma}\sqrt{n}\right) > \alpha \qquad \text{für alle} \qquad \mu > \mu_0$$

ist der Test **unverfälscht**.

Die **Operationscharakteristik** dieses Gauß-Tests ist für alle $\mu \in \mathbb{R}$ gegeben durch

$$L_\delta(\mu) \stackrel{\text{def}}{=} 1 - G_\delta(\mu) = \Phi\left(u_{1-\alpha} - \frac{\mu - \mu_0}{\sigma}\sqrt{n}\right).$$

4.2 t-Test

Zweck: Test über den Parameter μ einer Normalverteilung bei **unbekannter** Varianz σ^2

Voraussetzung: $X_i \overset{\text{i.i.d.}}{\sim} N(\mu, \sigma^2)$ für $i = 1, \ldots, n$

Gegeben: Signifikanzniveau $\alpha \in]0, 1[$, $\mu_0 \in \mathbb{R}$

4.2.1 Zweiseitiger t-Test

Null- und Gegenhypothese: $H_0 : \mu = \mu_0$ vs. $H_1 : \mu \neq \mu_0$

Prüfgröße: $T(X) \overset{\text{def}}{=} \frac{\bar{X}-\mu_0}{S^*(X)}\sqrt{n} = \frac{\bar{X}-\mu_0}{S(X)}\sqrt{n-1}$

Testverteilung: $T(X) \overset{\mu=\mu_0}{\sim} t(n-1)$

Die Testverteilung ist die Verteilung der Prüfgröße, falls $\mu = \mu_0$ und somit beim zweiseitigen Test die Hypothese H_0 richtig ist. Die Testverteilung ist beim t-Test die t-Verteilung mit $n-1$ Freiheitsgraden.

Ablehnungsbereich: $K_\delta^* =]-\infty, -t(n-1)_{1-\frac{\alpha}{2}}[\cup]t(n-1)_{1-\frac{\alpha}{2}}, \infty[$

Im t-Test ist $t(n-1)_{1-\frac{\alpha}{2}}$ das $(1-\frac{\alpha}{2})$-Quantil der t-Verteilung mit $n-1$ Freiheitsgraden. Da diese Verteilung eine Dichtefunktion besitzt, die symmetrisch um Null ist, gilt
$t(n-1)_{\frac{\alpha}{2}} = -t(n-1)_{1-\frac{\alpha}{2}}$.

Testentscheidung:

$$\delta(x) = \begin{cases} d_0, & \text{falls} \quad T(x) \notin K_\delta^* \\ d_1, & \text{falls} \quad T(x) \in K_\delta^* \end{cases}$$

Dabei ist $x \stackrel{\text{def}}{=} (x_1, \ldots, x_n)$ eine Realisation des Stichprobenvektors $X \stackrel{\text{def}}{=} (X_1, \ldots, X_n)$ und $T(x)$ eine Realisation der Prüfgröße $T(X)$.

4.2.2 Einseitige t-Tests

Bei einseitigen und zweiseitigen t-Tests stimmen Prüfgröße, Testverteilung und das Treffen der Testentscheidung überein, nur Hypothesenpaare und Ablehnungsbereiche sind verschieden.

Null- und Gegenhypothese: $H_0: \mu = \mu_0$ oder $\mu \geq \mu_0$ vs. $H_1: \mu < \mu_0$
Ablehnungsbereich: $K_\delta^* =]-\infty, -t(n-1)_{1-\alpha}[$
Dabei ist $t(n-1)_{1-\alpha}$ das $(1-\alpha)$-Quantil der t-Verteilung mit $n-1$ Freiheitsgraden. Da diese Verteilung eine Dichtefunktion besitzt, die symmetrisch um Null ist, gilt
$t(n-1)_\alpha = -t(n-1)_{1-\alpha}$.

Null- und Gegenhypothese: $H_0: \mu = \mu_0$ oder $\mu \leq \mu_0$ vs. $H_1: \mu > \mu_0$
Ablehnungsbereich: $K_\delta^* =]t(n-1)_{1-\alpha}, \infty[$

Hinweis: Für große Stichprobenumfänge n kann die t-Verteilung mit $n-1$ Freiheitsgraden durch die Standardnormalverteilung approximiert werden, d. h. $t(n-1)_\alpha \approx u_\alpha$, $t(n-1)_{1-\alpha} \approx u_{1-\alpha}$ und $t(n-1)_{1-\frac{\alpha}{2}} \approx u_{1-\frac{\alpha}{2}}$ (z. B. für $n-1 \geq 30$).

4.2.3 Beispiel für DAX-Renditen

- Zufallsvariable Y beschreibe die stetige Jahresrendite des DAX
- Annahme des Standardmodells für stetige Renditen: $Y \sim N(\mu, \sigma^2)$ mit unbekannten Parametern $\mu \in \mathbb{R}$ und $\sigma^2 > 0$
- einfache Zufallsstichprobe vom Umfang $n = 25$ Jahre zu Y: $X_t \overset{\text{i.i.d.}}{\sim} N(\mu, \sigma^2)$, $t = 1, \ldots, n$
- Schätzwerte: $\bar{x} = 10\%$, $S^*(x) = 25\%$

Testfragestellung: Ist $\bar{x}$ signifikant ($\alpha = 5\%$) von Null verschieden?
Hypothesen: $H_0: \ \mu = \mu_0$ vs. $H_1: \ \mu \neq \mu_0$ mit $\mu_0 = 0$
Realisation der Prüfgröße: $T(x) = \frac{\bar{x}-\mu_0}{S^*(x)}\sqrt{n} = \frac{0.1-0}{0.25}\sqrt{25} = 2$
Ablehnungsbereich: Mit $t(24)_{0.975} = 2.064$ folgt $K_\delta^* =]-\infty, -2.064[\cup]2.064, \infty[$.
Testentscheidung: H_0 wird nicht verworfen, da $T(x) \notin K_\delta^*$.

Testfragestellung: Ist $\bar{x}$ signifikant ($\alpha = 5\%$) größer als 4%?
Hypothesen: $H_0: \ \mu = \mu_0$ vs. $H_1: \ \mu > \mu_0$ mit $\mu_0 = 0.04$
Realisation der Prüfgröße: $T(x) = \frac{\bar{x}-\mu_0}{S^*(x)}\sqrt{n} = \frac{0.1-0.04}{0.25}\sqrt{25} = 1.2$
Ablehnungsbereich: Mit $t(24)_{0.95} = 1.711$ folgt $K_\delta^* =]1.711, \infty[$.
Testentscheidung: H_0 wird nicht verworfen, da $T(x) \notin K_\delta^*$.

4.3 Test über die Varianz der Normalverteilung

Zweck: Test über die Varianz σ^2 einer Normalverteilung bei **unbekanntem** Erwartungswert μ

Voraussetzung: $X_i \overset{\text{i.i.d.}}{\sim} N(\mu, \sigma^2)$ für $i = 1, \dots, n$

Gegeben: Signifikanzniveau $\alpha \in]0, 1[$, $\sigma_0^2 > 0$

4.3.1 Zweiseitiger Varianz-Test

Null- und Gegenhypothese: $H_0 : \sigma^2 = \sigma_0^2$ vs. $H_1 : \sigma^2 \neq \sigma_0^2$

Prüfgröße: $T(X) \overset{\text{def}}{=} \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma_0^2}$

Testverteilung: $T(X) \overset{\sigma^2 = \sigma_0^2}{\sim} \chi^2(n-1)$

Die Testverteilung ist die Verteilung der Prüfgröße, falls $\sigma^2 = \sigma_0^2$ und somit beim zweiseitigen Test die Hypothese H_0 richtig ist. Die Testverteilung ist beim Varianz-Test die χ^2-Verteilung mit $n-1$ Freiheitsgraden.

Ablehnungsbereich: $K_\delta^* = [0, \chi^2(n-1)_{\frac{\alpha}{2}}[\cup]\chi^2(n-1)_{1-\frac{\alpha}{2}}, \infty[$

Im Varianz-Test ist $\chi^2(n-1)_{\frac{\alpha}{2}}$ das $(\frac{\alpha}{2})$-Quantil und $\chi^2(n-1)_{1-\frac{\alpha}{2}}$ das $(1-\frac{\alpha}{2})$-Quantil der χ^2-Verteilung mit $n-1$ Freiheitsgraden.

Testentscheidung:

$$\delta(x) = \begin{cases} d_0, & \text{falls} \quad T(x) \notin K_\delta^* \\ d_1, & \text{falls} \quad T(x) \in K_\delta^* \end{cases}$$

Dabei ist $x \stackrel{\text{def}}{=} (x_1, \ldots, x_n)$ eine Realisation des Stichprobenvektors $X \stackrel{\text{def}}{=} (X_1, \ldots, X_n)$ und $T(x)$ eine Realisation der Prüfgröße $T(X)$.

4.3.2 Einseitige Varianz-Tests

Bei einseitigen und zweiseitigen Varianz-Tests stimmen Prüfgröße, Testverteilung und das Treffen der Testentscheidung überein, nur Hypothesenpaare und Ablehnungsbereiche sind verschieden.

Null- und Gegenhypothese: $H_0 : \sigma^2 = \sigma_0^2$ oder $\sigma^2 \geq \sigma_0^2$ vs. $H_1 : \sigma^2 < \sigma_0^2$
Ablehnungsbereich: $K_\delta^* = [0, \chi^2(n-1)_\alpha[$
Dabei ist $\chi^2(n-1)_\alpha$ das α-Quantil der χ^2-Verteilung mit $n-1$ Freiheitsgraden.

Null- und Gegenhypothese: $H_0 : \sigma^2 = \sigma_0^2$ oder $\sigma^2 \leq \sigma_0^2$ vs. $H_1 : \sigma^2 > \sigma_0^2$
Ablehnungsbereich: $K_\delta^* =]\chi^2(n-1)_{1-\alpha}, \infty[$
Dabei ist $\chi^2(n-1)_{1-\alpha}$ das $(1-\alpha)$-Quantil der χ^2-Verteilung mit $n-1$ Freiheitsgraden.

Bemerkung: Ist der Erwartungswert $\mu \in \mathbb{R}$ der Normalverteilung bekannt, ergeben sich ein- und zweiseitige Varianz-Tests zu den oben angegebenen Hypothesenpaaren, falls die Prüfgröße

$$T(X) = \frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma_0^2}$$

verwendet wird. Als Testverteilung ergibt sich dann für $\sigma^2 = \sigma_0^2$ die χ^2-Verteilung mit n Freiheitsgraden. In den jeweiligen Ablehnungsbereichen müssen dann die entsprechenden Quantile der χ^2-Verteilung mit n Freiheitsgraden eingesetzt werden.

Hinweis: Für große Stichprobenumfänge kann die χ^2-Verteilung mit ν Freiheitsgraden durch die $N(\nu, 2\nu)$-Verteilung approximiert werden. Für das α-Quantil gilt dann beispielsweise
$\chi^2(\nu)_\alpha \approx \nu + u_\alpha \cdot \sqrt{2\nu}$ für $\nu \geq 30$.

5 Approximative Tests

Im Folgenden werden zwei **approximative Gauß-Tests** und ein **Test über eine Wahrscheinlichkeit** vorgestellt.

Nachteil: Tests basieren auf Zentralem Grenzwertsatz der Statistik $\Longrightarrow$ Approximationsfehler für endlichen Stichprobenumfang n; Approximationsfehler wird mit zunehmendem n zwar kleiner, verschwindet aber nur asymptotisch ($n \to \infty$)

Vorteil: keine normalverteilte Grundgesamtheit als Voraussetzung erforderlich

Voraussetzungen:

- Die Verteilung von Y hat den Erwartungswert $\mu \stackrel{\text{def}}{=} E(Y)$ und die Varianz $0 < \sigma^2 \stackrel{\text{def}}{=} Var(Y) < \infty$.
- Es liegt eine einfache Zufallsstichprobe (iid-Stichprobe) vom Umfang n zu Y vor, d. h. die X_i sind unabhängig und identisch verteilt mit Erwartungswert $\mu = E(X_i)$ und Varianz $\sigma^2 = Var(X_i)$ für $i = 1, \ldots, n$.

Die approximativen Gauß-Tests sind Tests über den Erwartungswert μ bei bekannter oder unbekannter Varianz σ^2. Der Test über eine Wahrscheinlichkeit ist ein Test über den Parameter p einer Bernoulli-Verteilung.

5.1 Approximativer Gauß-Test bei bekannter Varianz

Zweck: Test über den Erwartungswert μ einer Grundgesamtheit bei bekannter Varianz σ^2

Voraussetzung: Die X_i sind unabhängig und identisch verteilt (iid) mit Erwartungswert $\mu = E(X_i)$ und Varianz $\sigma^2 = Var(X_i)$ für $i = 1, \ldots, n$.

Gegeben: Signifikanzniveau $\alpha \in]0, 1[$, Varianz $\sigma^2 > 0$, $\mu_0 \in \mathbb{R}$

Der auf der **Prüfgröße**

$$T(X) \stackrel{\text{def}}{=} \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n} \tag{5}$$

beruhende Gauß-Test aus Abschnitt 4.1 ist approximativ gültig. D. h. falls $\mu = \mu_0$, dann ist die Prüfgröße (5) für große Stichprobenumfänge n approximativ standardnormalverteilt, so dass für die **Testverteilung** gilt:

$$T(X) \stackrel{\mu=\mu_0}{\approx} N(0, 1).$$

Als Faustregel gilt $n \geq 30$. Der Approximationsfehler besteht darin, dass das vorgegebene Signifikanzniveau α nur approximativ eingehalten wird. Hypothesen, Ablehnungsbereiche und Testentscheidungen stimmen mit denen der exakten Gauß-Tests aus Abschnitt 4.1 überein.

5.2 Approximativer Gauß-Test bei unbekannter Varianz

Zweck: Test über den Erwartungswert μ einer Grundgesamtheit bei unbekannter Varianz σ^2

Voraussetzung: Die X_i sind unabhängig und identisch verteilt (iid) mit Erwartungswert $\mu = E(X_i)$ und Varianz $\sigma^2 = Var(X_i)$ für $i = 1, \ldots, n$.

Gegeben: Signifikanzniveau $\alpha \in]0, 1[$, $\mu_0 \in \mathbb{R}$

Ist die Varianz σ^2 unbekannt aber endlich, dann kann σ in Prüfgröße (5) durch den Schätzer $S(X)$ ersetzt werden und man erhält die modifizierte **Prüfgröße**

$$T(X) \stackrel{\text{def}}{=} \frac{\bar{X} - \mu_0}{S(X)} \sqrt{n}.$$

Falls $\mu = \mu_0$, dann ist diese modifizierte Prüfgröße für große Stichprobenumfänge n approximativ standardnormalverteilt, so dass für die **Testverteilung** gilt:

$$T(X) \stackrel{\mu=\mu_0}{\approx} N(0, 1).$$

Als Faustregel gilt $n \geq 30$. Der Approximationsfehler besteht darin, dass das vorgegebene Signifikanzniveau α nur approximativ eingehalten wird. Hypothesen, Ablehnungsbereiche und Testentscheidungen stimmen mit denen der exakten Gauß-Tests aus Abschnitt 4.1 überein.

5.3 Test über eine Wahrscheinlichkeit

Zweck: Test über den Parameter p einer Bernoulli-Verteilung

Voraussetzung: $X_i \overset{\text{i.i.d.}}{\sim} B(1,p)$ mit $0 < p < 1$ für $i = 1, \ldots, n$.

Gegeben: Signifikanzniveau $\alpha \in]0,1[$, $p_0 \in]0,1[$

Null- und Gegenhypothese:

$H_0: \ p = p_0$	vs.	$H_1: \ p \neq p_0$
$H_0: \ p = p_0$ oder $p \geq p_0$	vs.	$H_1: \ p < p_0$
$H_0: \ p = p_0$ oder $p \leq p_0$	vs.	$H_1: \ p > p_0$

Prüfgröße: $T(X) \overset{\text{def}}{=} \frac{\bar{X} - p_0}{\sqrt{p_0(1-p_0)}} \sqrt{n}$

Testverteilung: $T(X) \overset{p=p_0}{\approx} N(0,1)$

Falls $p = p_0$, dann ist die Prüfgröße für große Stichprobenumfänge n approximativ standardnormalverteilt. Als Faustregel gilt $np_0(1-p_0) > 9$. Ablehnungsbereiche und Testentscheidungen stimmen mit denen des exakten Gauß-Tests aus Abschnitt 4.1 überein.