8. Hypothesentests

Fragestellungen (1)

- Eine Herstellerfirma umweltfreundlicher Energiesparlampen behauptet, dass die Haltbarkeit ihrer Lampen im Durchschnitt $10\,000$ Stunden beträgt.

  Frage: Ist diese Behauptung gerechtfertigt?

  Oder anders formuliert:
  Stimmt der Erwartungswert der Zufallsvariable $X =$ „Lebensdauer“ mit den angegebenen $10\,000$ Stunden überein?

Fragestellungen (2)

- Eine Firma produziert Nägel. Der Sollwert der Nagellänge einer bestimmten Sorte beträgt 17 cm.

  Frage: Stimmt der Erwartungswert μ der Zufallsvariable $X =$ „Nagellänge“ mit der Solllänge überein?

Statistische Hypothesen

- Eine **statistische Hypothese** ist eine Behauptung über Eigenschaften einer Zufallsvariable.
- Beispiele:
  - Der Erwartungswert μ der Zufallsvariable $X =$ „Nagellänge“ beträgt 17 cm.
  - Die Standardabweichung σ der Zufallsvariable $X =$ „Nagellänge“ beträgt 1.5 cm.

Testen von Hypothesen

- Das Interesse ist es, zu überprüfen, ob bestimmte Behauptungen über Eigenschaften (Parameter) einer Zufallsvariable zutreffen oder nicht.
- Im Gegensatz zum Schätzen liegt hier ein *Entscheidungsproblem* vor.

Bsp. Nagellänge

- Der Sollwert der Länge einer bestimmten Sorte Nägel beträgt $17\,$cm.
- Frage: Stimmt der Erwartungswert μ (Parameter) der Zufallsvariable $X =$ „Nagellänge“ mit der Solllänge überein?
- Statistische Hypothese: „Die zu erwartende Länge μ der Nägel beträgt $17\,$cm.“

Hypothesentests (1)

- Hypothesentests dienen dem statistischen Nachweis von Unterschieden oder Effekten.
- Es sind „objektive“ Entscheidungsverfahren oder „Entscheidungstools“, die auf Basis von Stichproben und Wahrscheinlichkeitsmodellen Entscheidungen herbeiführen.
- Über die Wahrscheinlichkeitsmodelle müssen Annahmen getroffen werden.

Bsp. Hypothesentests

- Frage: Stimmt der Erwartungswert μ der Zufallsvariable $X =$ „Nagellänge“ mit der Solllänge von 17 cm überein?
- Statistische Hypothese: „Die zu erwartende Länge μ der Nägel beträgt 17 cm.“
- Annahme: Es sei bekannt, dass die Länge X der Nägel normalverteilt ist mit $E(X) = \mu$ und $\sigma(X) = \sigma = 1.50$ cm.

Hypothesentests (2)

- Statistische Hypothesen sind in der Regel nur indirekt prüfbar.
- Idee: Man stellt zunächst eine Hypothese auf und verwirft sie genau dann, wenn sich eine Beobachtung einstellt, die bei Gültigkeit der Hypothese unwahrscheinlich ist.

Hypothesen und Testproblem

- **Nullhypothese (H_0):** Formulierung der Gleichheit (keine Abweichung, kein Effekt, keine Wirkung …)
- **Alternativhypothese (H_1):** Formulierung eines Unterschieds (Abweichung, Effekt, Wirkung …); Nullhypothese <u>nicht</u> richtig
- statistisches Testproblem für den Parameter θ, θ_0 vorgegeben:

$$H_0 : \theta = \theta_0 \quad \text{gegen} \quad H_1 : \theta \neq \theta_0$$

Bsp. Statistisches Testproblem

- Frage: Stimmt der Erwartungswert μ der Zufallsvariable $X =$ „Nagellänge“ mit der Solllänge von 17 cm überein?
- Nullhypothese (H_0): „Die zu erwartende Länge der Nägel beträgt 17 cm.“
- Alternativhypothese (H_1): „Die zu erwartende Länge der Nägel beträgt nicht 17 cm.“
- Das statistische Testproblem lautet:

$$H_0 : \mu = 17 \quad \text{gegen} \quad H_1 : \mu \neq 17$$

Aufbau von Hypothesentests

- Intention eines Hypothesentests ist es, die Nullhypothese zu überprüfen:

$$H_0 : \text{« was zu widerlegen ist »}$$
$$H_1 : \text{« was man gerne zeigen möchte »}$$

- Wie wird entschieden, ob H_0 abgelehnt werden kann?

Prüfgrößen

- Die **Prüfgröße** oder **Testfunktion** ist die Vorschrift (Formel), nach der aus einer gegebenen Stichprobe eine Zahl (Wert der Prüfgröße) errechnet wird.
- Der Test besteht nun darin, dass je nach dem Wert der Prüfgröße **für oder gegen die Nullhypothese** entschieden wird.

Bsp. Prüfgröße Nagellänge

- Als Prüfgröße bietet sich $\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ an.
- Es wird eine zufällige Stichprobe aus der laufenden Produktion entnommen. Folgende Längen (in Zentimetern) werden an $n = 10$ Nägeln gemessen:

$$19.2, 17.4, 18.5, 16.2, 18.9, 16.7, 17.1, 17.0, 16.9, 17.6$$

- Der aus der Stichprobe berechnete Wert der Prüfgröße $\overline{X}$ beträgt $\bar{x} = 17.60\,\text{cm}$.
- Ist das nun „ähnlich“ zu $17\,\text{cm}$? Wie ist die Prüfgröße $\overline{X}$ verteilt?

Prüfverteilungen

- **Prüfverteilung:** Verteilung der Prüfgröße
- Voraussetzung: Verteilung der Prüfgröße unter H_0 ist bekannt

- Verteilung der Prüfgröße unter H_0:

$$\overline{X} \sim N(\mu_0, \tfrac{\sigma}{\sqrt{n}})$$

bzw.

$$Z = \frac{\overline{X} - \mu_0}{\sigma}\sqrt{n} \sim N(0,1)\,,$$

mit $\mu_0 = 17\,\text{cm}$, $\sigma = 1.5\,\text{cm}$ und $n = 10$.

Die Verteilung der standardisierten Prüfgröße Z unter H_0 ist gerade die Standardnormalverteilung.

Bsp. Prüfverteilung Nagellänge (2)

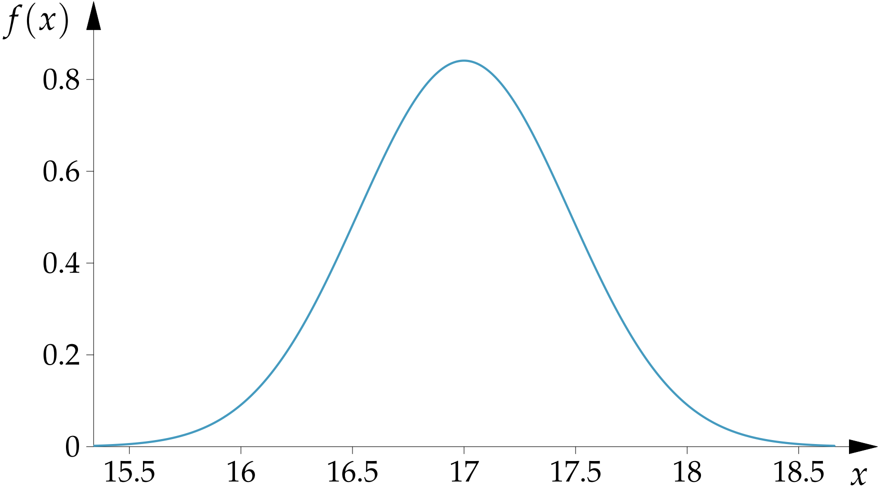

Abb.: Dichtefunktion der Prüfgröße $\overline{X} \sim N(17, \frac{1.5}{\sqrt{10}})$

Bsp. Stand. Prüfverteilung Nagellänge

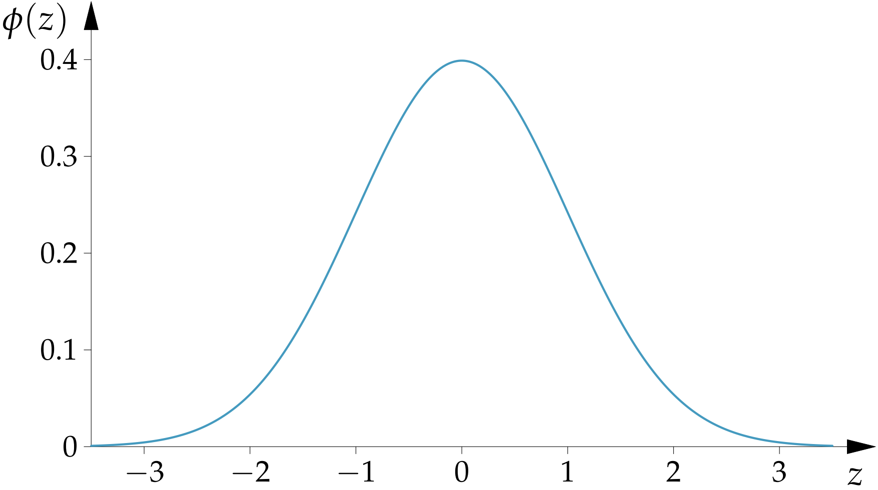

Abb.: Dichtefunktion der stand. Prüfgröße $Z \sim N(0, 1)$

Bsp. Prüfgrößenwert Nagellänge

- Der aus der Stichprobe berechnete Wert der Prüfgröße $\overline{X}$ beträgt $\bar{x} = 17.60\,\text{cm}$.
- Als Wert der standardisierten Prüfgröße Z erhält man

$$z = \frac{\bar{x} - \mu_0}{\sigma}\sqrt{n} = \frac{17.6 - 17}{1.5}\sqrt{10} = 1.26\,.$$

Bsp. Prüfverteilung u. Prüfgrößenwert

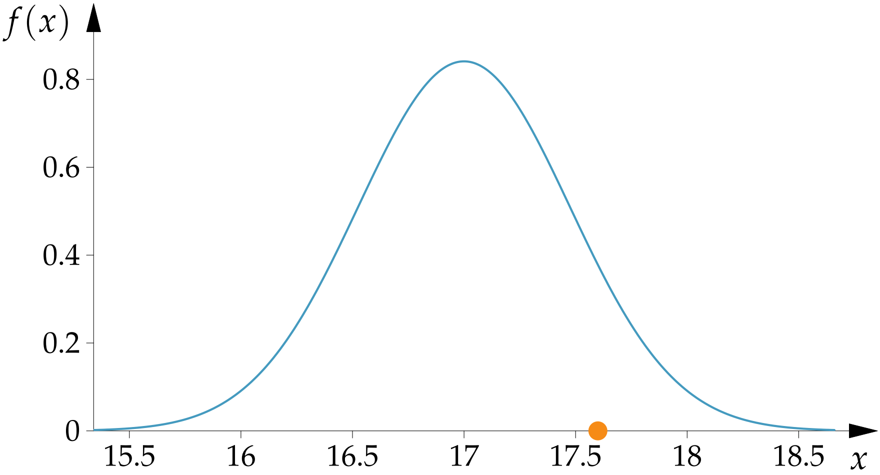

Abb.: Dichte der Prüfgröße $\overline{X}$ mit dem beobachteten Prüfgrößenwert 17.6 (orange)

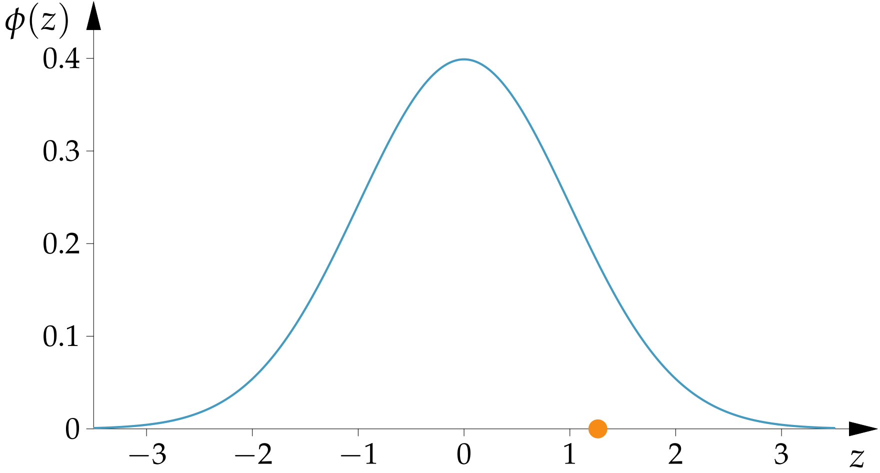

Abb.: Dichte der standardisierten Prüfgröße Z mit dem beobachteten Prüfgrößenwert 1.26 (orange)

Signifikanzniveau α

- Das **Signifikanzniveau $\boldsymbol{\alpha}$** stellt die Entscheidungsgrenze zwischen Nullhypothese und Alternative dar und legt die maximal zulässige Irrtumswahrscheinlichkeit fest.
- Der **kritische Wert $\boldsymbol{c}$** unterteilt den Definitionsbereich der Prüfgröße in einen sogenannten *Annahme*- und *Ablehnbereich*.
- Der kritische Wert ergibt sich aus dem festgelegten α und der Verteilung der Prüfgröße. Der Ablehnbereich wird auch als *kritischer Bereich* bezeichnet.

- Testproblem: $H_0 : \theta = \theta_0$ gegen $H_1 : \theta \neq \theta_0$

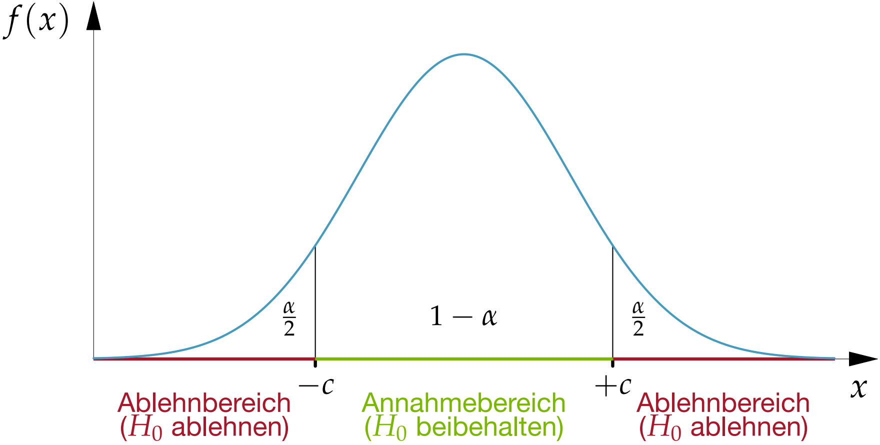

Abb.: Prüfverteilung mit Ablehnbereich

$C = (-\infty, -c) \cup (+c, +\infty)$ (rot)

Bsp. Kritischer Wert Nagellänge (1)

- Der kritische Wert c ergibt sich hier als das $(1 - \alpha/2)$-Quantil $z_{1-\alpha/2}$ der Standardnormalverteilung (Prüfverteilung).
- Für $\alpha = 0.05$ ist das $(1 - \alpha/2)$-Quantil der $N(0,1)$-verteilung $z_{0.975} = 1.96$.

Bsp. Ablehnbereich Nagellänge (1)

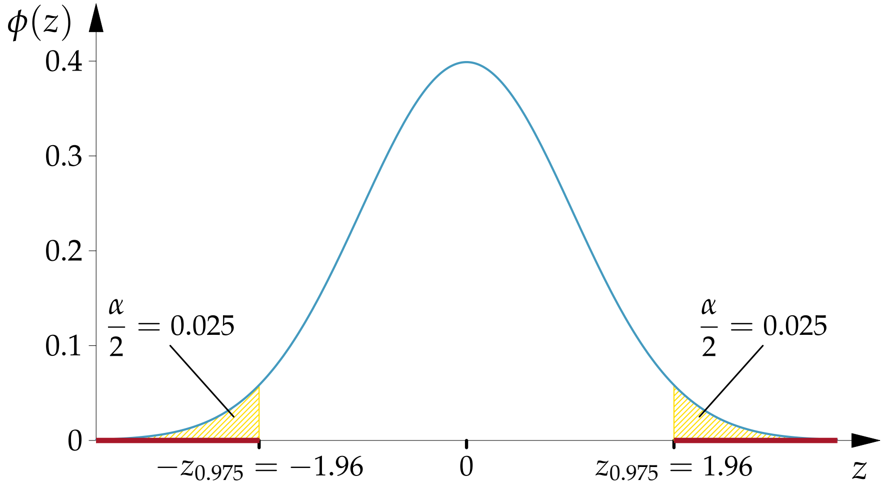

Abb.: Dichte der stand. Prüfgröße Z mit dem Ablehnbereich $C = (-\infty, -1.96) \cup (1.96, \infty)$ (rot) für $\alpha = 0.05$

Bsp. Kritischer Wert Nagellänge (2)

- Die $2.5\,\%$- und $97.5\,\%$-Quantile der $N(17, \frac{1.5}{\sqrt{10}})$-verteilung sind

$$x_{0.025} = 17 - z_{0.975} \cdot \frac{1.5}{\sqrt{10}} = 16.07$$

bzw.

$$x_{0.975} = 17 + z_{0.975} \cdot \frac{1.5}{\sqrt{10}} = 17.93\,.$$

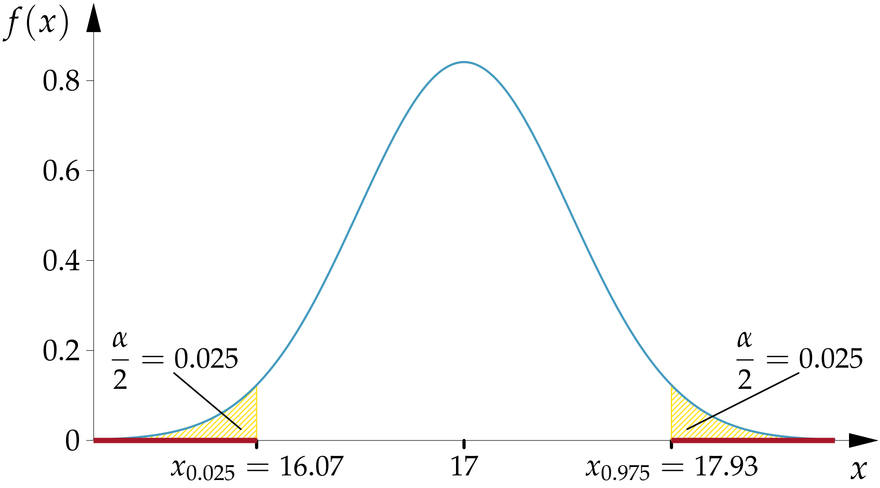

Abb.: Dichte der Prüfgröße $\overline{X}$ mit dem Ablehnbereich
$C = (-\infty, 16.07) \cup (17.93, \infty)$ (rot) für $\alpha = 0.05$

Testentscheidung

- Basierend auf der Prüfgröße fällt die Entscheidung für H_1 im Testproblem

$$H_0 : \theta = \theta_0 \quad \text{gegen} \quad H_1 : \theta \neq \theta_0 \,,$$

falls der Wert der Prüfgröße in den Ablehnbereich fällt.

Bsp. Testentscheidung Nagellänge (1)

- Damit muss davon ausgegangen werden, dass sich der Prozess nicht mehr unter statistischer Kontrolle befindet, falls für den beobachteten Wert z der Prüfgröße Z gilt:

$$|z| > 1.96$$

- Da für die Stichprobe Z den Wert 1.26 angenommen hat, kann H_0 nicht verworfen werden.
- Bei einem Signifikanzniveau von $5\,\%$ kann statistisch nicht nachgewiesen werden, dass die erwartete Nagellänge von der Solllänge von $17\,$cm abweicht.

Bsp. Testentscheidung Nagellänge (2)

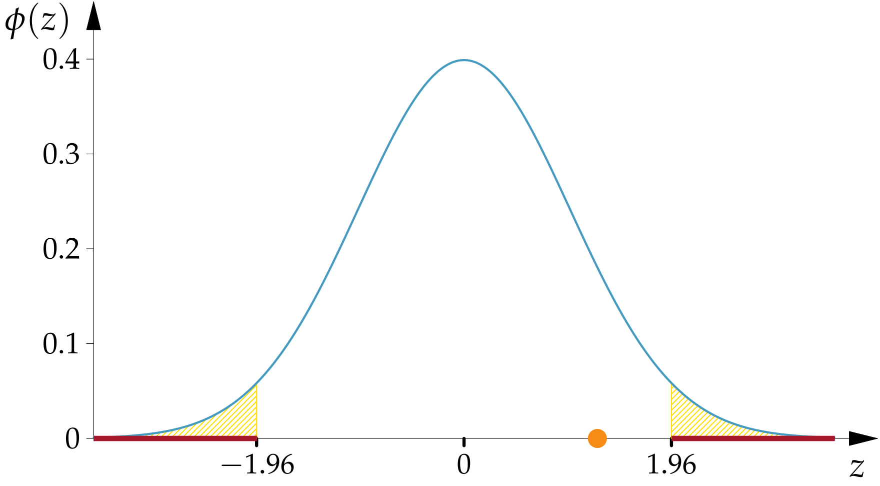

Abb.: Dichte der stand. Prüfgröße Z mit dem beobachteten Prüfgrößenwert 1.26 (orange) und dem Ablehnbereich (rot)

Bsp. Testentscheidung Nagellänge (3)

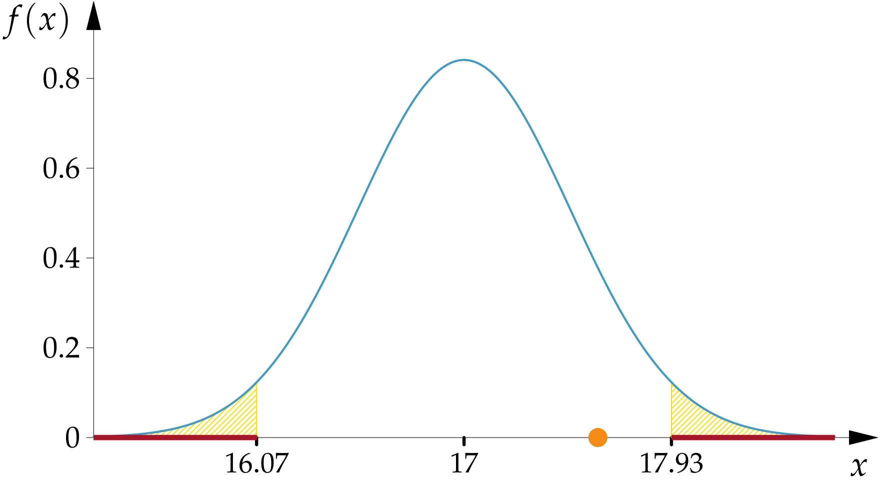

Abb.: Dichte der Prüfgröße $\overline{X}$ mit dem beobachteten Prüfgrößenwert 17.6 (orange) und dem Ablehnbereich (rot)

Kleine Zusammenfassung

- Ein statistisches Testproblem besteht aus einer Nullhypothese H_0 und einer Alternative H_1, die sich gegenseitig ausschließen und Aussagenüber bestimmte Parameter des interessierenden Merkmals in der Grundgesamtheit beinhalten.
- Ein statistischer Test basiert auf einer geeignet gewählten Prüfgröße und liefert eine formale Entscheidungsregel, die aufgrund einer Stichprobe darüber entscheidet, ob eher H_0 oder H_1 für die Grundgesamtheit zutrifft.