

# Hypothesentests

- *Bisher* wichtigstes betrachtetes Anwendungsbeispiel der schließenden Statistik:

## **Punkt- bzw. Intervallschätzung** des unbekanntes Mittelwerts

- Hierzu: Verwendung der
  - 1 theoretischen Information über Verteilung von  $\bar{X}$
  - 2 empirischen Information aus Stichprobenrealisation  $\bar{x}$  von  $\bar{X}$
 zur Konstruktion einer
  - ▶ Punktschätzung (inkl. Beurteilung der Genauigkeit des Schätzers!)
  - ▶ Intervallschätzung, bei der jede Stichprobenziehung mit einer vorgegebenen Chance ein realisiertes (Konfidenz-)Intervall liefert, welches den (wahren) Mittelwert enthält.
- Nächste Anwendung: **Hypothesentests**:
 

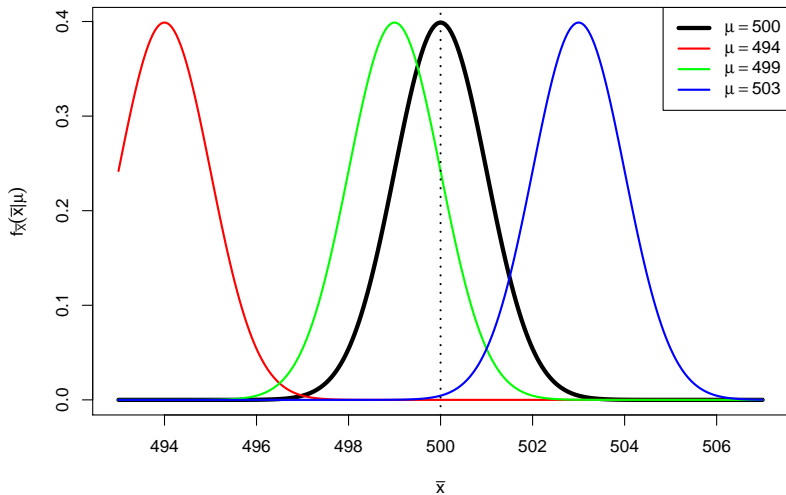
**Entscheidung**, ob die unbekanntes, wahre Verteilung von  $Y$  zu einer vorgegebenen Teilmenge der Verteilungsannahme  $W$  gehört oder nicht.
- Zunächst: Illustration der Vorgehensweise am Beispiel einer Entscheidung über den Mittelwert der Verteilung.

# Einführendes Beispiel

- Interessierende Zufallsvariable  $Y$ :  
Von einer speziellen Abfüllmaschine abgefüllte Inhaltmenge von Müslipackungen mit Soll-Inhalt  $\mu_0 = 500$  (in [g]).
- Verteilungsannahme:  
 $Y \sim N(\mu, 4^2)$  mit unbekanntem Erwartungswert  $\mu = E(Y)$ .
- Es liege eine Realisation  $x_1, \dots, x_{16}$  einer einfachen Stichprobe  $X_1, \dots, X_{16}$  vom Umfang  $n = 16$  zu  $Y$  vor.
- **Ziel:** Verwendung der Stichprobeninformation (über  $\bar{X}$  bzw.  $\bar{x}$ ), um zu **entscheiden**, ob die tatsächliche mittlere Füllmenge (also der wahre, unbekannte Parameter  $\mu$ ) mit dem Soll-Inhalt  $\mu_0 = 500$  übereinstimmt oder nicht.
- Offensichtlich gilt:
  - ▶  $\bar{X}$  schwankt um den wahren Mittelwert  $\mu$ ; selbst wenn  $\mu = 500$  gilt, wird  $\bar{X}$  praktisch nie genau den Wert  $\bar{x} = 500$  annehmen!
  - ▶ Realisationen  $\bar{x}$  „in der Nähe“ von 500 sprechen eher dafür, dass  $\mu = 500$  gilt.
  - ▶ Realisationen  $\bar{x}$  „weit weg“ von 500 sprechen eher dagegen, dass  $\mu = 500$  gilt.
- Also: Entscheidung für Hypothese  $\mu = 500$ , wenn  $\bar{x}$  nahe bei 500, und gegen  $\mu = 500$  (also für  $\mu \neq 500$ ), wenn  $\bar{x}$  weit weg von 500.
- **Aber:** Wo ist die Grenze zwischen „in der Nähe“ und „weit weg“?

# Verteilungen von $\bar{X}$

für verschiedene Parameter  $\mu$  bei  $\sigma = 4$  und  $n = 16$



# Entscheidungsproblem

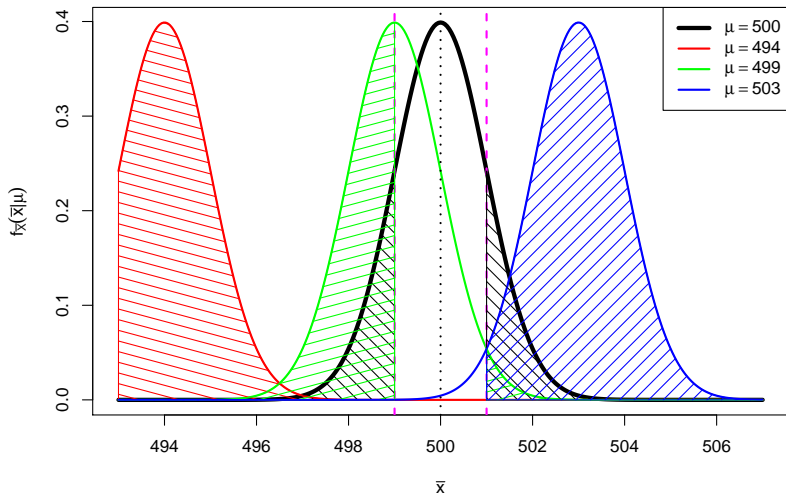
- Fällen einer Entscheidung zwischen  $\mu = 500$  und  $\mu \neq 500$  führt zu *genau einer* der folgenden vier verschiedenen Situationen:

	Tatsächliche Situation: $\mu = 500$	Tatsächliche Situation: $\mu \neq 500$
Entscheidung für $\mu = 500$	richtige Entscheidung	Fehler 2. Art
Entscheidung für $\mu \neq 500$	Fehler 1. Art	richtige Entscheidung

- Wünschenswert:  
Sowohl „Fehler 1. Art“ als auch „Fehler 2. Art“ möglichst selten begehen.
- Aber:** Zielkonflikt vorhanden:  
Je näher Grenze zwischen „in der Nähe“ und „weit weg“ an  $\mu_0 = 500$ , desto
  - ▶ seltener Fehler 2. Art
  - ▶ häufiger Fehler 1. Art
 und umgekehrt für fernere Grenzen zwischen „in der Nähe“ und „weit weg“.

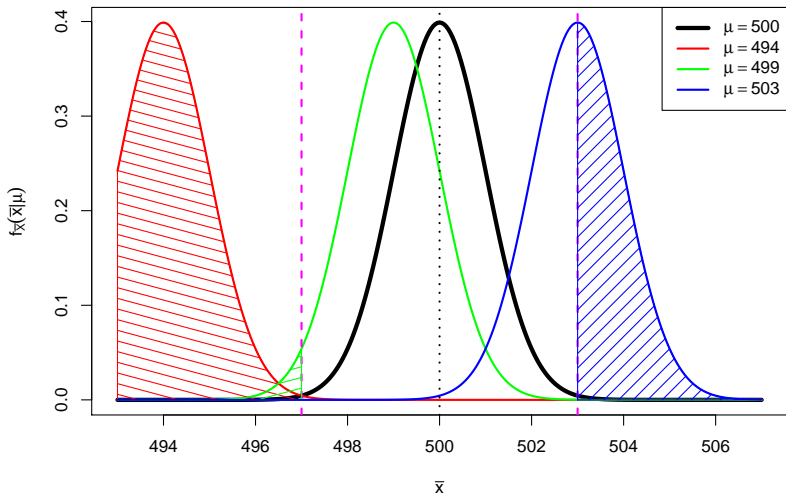
# Beispiel für „nahe“ Grenze

Für  $\mu \neq 500$  (gegen  $\mu = 500$ ) entscheiden, wenn Abstand zwischen  $\bar{x}$  und 500 größer als 1



# Beispiel für „ferne“ Grenze

Für  $\mu \neq 500$  (gegen  $\mu = 500$ ) entscheiden, wenn Abstand zwischen  $\bar{x}$  und 500 größer als 3



- Unmöglich, Wahrscheinlichkeiten der Fehler 1. Art und 2. Art gleichzeitig für alle möglichen Situationen (also alle  $\mu$ ) zu verringern.
- Übliche Vorgehensweise: **Fehler(wahrscheinlichkeit) 1. Art kontrollieren!**
- Also: Vorgabe einer *kleinen* Schranke  $\alpha$  („**Signifikanzniveau**“) für die Wahrscheinlichkeit, mit der man einen Fehler 1. Art begehen darf.
- Festlegung der Grenze zwischen „in der Nähe“ und „weit weg“ so, dass man den Fehler 1. Art nur mit Wahrscheinlichkeit  $\alpha$  begeht, also die Realisation  $\bar{x}$  **bei Gültigkeit von**  $\mu = 500$  nur mit einer Wahrscheinlichkeit von  $\alpha$  jenseits der Grenzen liegt, bis zu denen man sich für  $\mu = 500$  entscheidet!
- Damit liefert aber das Schwankungsintervall für  $\bar{X}$  zur Sicherheitswahrscheinlichkeit  $1 - \alpha$

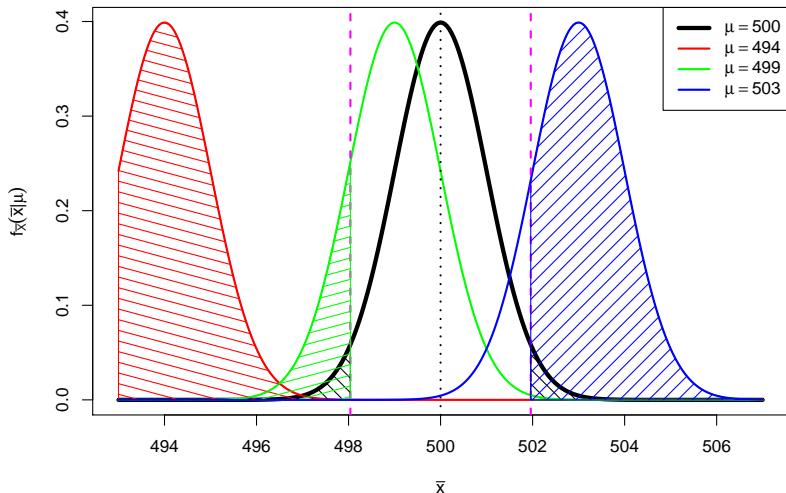
$$\left[ \mu - \frac{\sigma}{\sqrt{n}} \cdot N_{1-\frac{\alpha}{2}}, \mu + \frac{\sigma}{\sqrt{n}} \cdot N_{1-\frac{\alpha}{2}} \right]$$

für  $\mu = \mu_0 = 500$  (!) gerade solche Grenzen, denn es gilt im Fall  $\mu = \mu_0 = 500$

$$P \left\{ \bar{X} \notin \left[ \mu_0 - \frac{\sigma}{\sqrt{n}} \cdot N_{1-\frac{\alpha}{2}}, \mu_0 + \frac{\sigma}{\sqrt{n}} \cdot N_{1-\frac{\alpha}{2}} \right] \right\} = \alpha .$$

# Beispiel für Grenze zum Signifikanzniveau $\alpha = 0.05$

Grenzen aus Schwankungsintervall zur Sicherheitswahrscheinlichkeit  $1 - \alpha = 0.95$





- Bei einem Signifikanzniveau von  $\alpha = 0.05$  entscheidet man sich also **für**  $\mu = \mu_0 = 500$  genau dann, wenn die Realisation  $\bar{x}$  von  $\bar{X}$  im Intervall

$$\left[ 500 - \frac{4}{\sqrt{16}} \cdot N_{0.975}, 500 + \frac{4}{\sqrt{16}} \cdot N_{0.975} \right] = [498.04, 501.96] ,$$

dem sog. **Annahmebereich** des Hypothesentests, liegt.

- Entsprechend fällt die Entscheidung für  $\mu \neq 500$  (bzw. **gegen**  $\mu = 500$ ) aus, wenn die Realisation  $\bar{x}$  von  $\bar{X}$  in der Menge

$$(-\infty, 498.04) \cup (501.96, \infty) ,$$

dem sog. **Ablehnungsbereich** oder **kritischen Bereich** des Hypothesentests, liegt.

- Durch Angabe eines dieser Bereiche ist die Entscheidungsregel offensichtlich schon vollständig spezifiziert!
- Statt Entscheidungsregel auf Grundlage der Realisation  $\bar{x}$  von  $\bar{X}$  (unter Verwendung der Eigenschaft  $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$ ) üblicher:

**Äquivalente** Entscheidungsregel auf Basis der sog. **Testgröße** oder

**Teststatistik**  $N := \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n}$  unter Verwendung der Eigenschaft

$$\frac{\bar{X} - \mu_0}{\sigma} \sqrt{n} \sim N(0, 1) \quad \text{falls} \quad \mu = \mu_0 .$$

- Die Verteilungseigenschaft von  $N = \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n}$  für  $\mu = \mu_0$  aus Folie 98 erhält man aus der allgemeineren Verteilungsaussage

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

die man wiederum aus der Verteilung  $\bar{X} \sim N \left( \mu, \frac{\sigma^2}{n} \right)$  durch Anwendung der aus der Wahrscheinlichkeitsrechnung bekannten Transformationsregeln ableiten kann. Damit:  $F_N(x) = P\{N \leq x\} = \Phi \left( x - \frac{\mu - \mu_0}{\sigma} \sqrt{n} \right)$

- Man rechnet außerdem leicht nach:

$$\bar{X} \in \left[ \mu_0 - \frac{\sigma}{\sqrt{n}} \cdot N_{1-\frac{\alpha}{2}}, \mu_0 + \frac{\sigma}{\sqrt{n}} \cdot N_{1-\frac{\alpha}{2}} \right] \Leftrightarrow \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n} \in \left[ -N_{1-\frac{\alpha}{2}}, N_{1-\frac{\alpha}{2}} \right]$$

- Als Annahmehereich  $A$  für die Testgröße  $N = \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n}$  erhält man also  $\left[ -N_{1-\frac{\alpha}{2}}, N_{1-\frac{\alpha}{2}} \right]$ , als kritischen Bereich  $K$  entsprechend

$$K = \mathbb{R} \setminus A = \left( -\infty, -N_{1-\frac{\alpha}{2}} \right) \cup \left( N_{1-\frac{\alpha}{2}}, \infty \right)$$

und damit eine Formulierung der Entscheidungsregel auf Grundlage von  $N$ .

- In Abhängigkeit des tatsächlichen Erwartungswerts  $\mu$  von  $Y$  kann so die Wahrscheinlichkeit für die Ablehnung der Hypothese  $\mu = \mu_0$  berechnet werden:

$$\begin{aligned} P\{N \in K\} &= P\{N \in (-\infty, -N_{1-\frac{\alpha}{2}}) \cup (N_{1-\frac{\alpha}{2}}, \infty)\} \\ &= P\{N < -N_{1-\frac{\alpha}{2}}\} + P\{N > N_{1-\frac{\alpha}{2}}\} \\ &= \Phi\left(-N_{1-\frac{\alpha}{2}} - \frac{\mu - \mu_0}{\sigma} \sqrt{n}\right) + 1 - \Phi\left(N_{1-\frac{\alpha}{2}} - \frac{\mu - \mu_0}{\sigma} \sqrt{n}\right) \end{aligned}$$

- Im Beispiel erhält man damit die folgenden Wahrscheinlichkeiten für Annahme bzw. Ablehnung der Hypothese  $\mu = \mu_0 = 500$  zu den betrachteten Szenarien (also unterschiedlichen wahren Parametern  $\mu$ ):

	Wahrscheinlichkeit der Annahme von $\mu = 500$ $P\{N \in A\}$	Wahrscheinlichkeit der Ablehnung von $\mu = 500$ $P\{N \in K\}$
$\mu = 500$	<b>0.95</b>	0.05
$\mu = 494$	0	<b>1</b>
$\mu = 499$	0.8299	<b>0.1701</b>
$\mu = 503$	0.1492	<b>0.8508</b>

*(Fettgedruckte Wahrscheinlichkeiten entsprechen korrekter Entscheidung.)*

# Grundbegriffe: Hypothesen

- Bekannt: Hypothesentests sind Entscheidungsregeln für die Fragestellung „Liegt die (unbekannte) Verteilung  $Q_Y$  von  $Y$  in einer bestimmten **Teilmenge** der Verteilungsannahme  $W$  oder nicht?“
- Zur präzisen Formulierung der Fragestellung: Angabe der interessierenden Teilmenge  $W_0$  von Verteilungen mit  $\emptyset \neq W_0 \subsetneq W$
- Man nennt dann die Hypothese  $Q_Y \in W_0$  auch **Nullhypothese** und schreibt  $H_0 : Q_Y \in W_0$ . Die Verletzung der Nullhypothese entspricht dem Eintreten der sog. **Gegenhypothese** oder **Alternative**  $Q_Y \in W_1 := W \setminus W_0$ ; man schreibt auch  $H_1 : Q_Y \in W_1$ .
- Formulierung *prinzipiell* immer in zwei Varianten möglich, da  $W_0$  und  $W_1$  vertauscht werden können. Welche der beiden Varianten gewählt wird, ist allerdings wegen der Asymmetrie in den Wahrscheinlichkeiten für Fehler 1. und 2. Art **nicht unerheblich** (später mehr!).
- Eine Hypothese heißt **einfach**, wenn die zugehörige Teilmenge von  $W$  einelementig ist, **zusammengesetzt** sonst.
- Im Beispiel:  $W = \{N(\mu, 4^2) \mid \mu \in \mathbb{R}\}$ ,  $W_0 = \{N(500, 4^2)\}$ .  
 $H_0$  ist also einfach,  $H_1$  zusammengesetzt.

# Hypothesen bei parametrischen Verteilungsannahmen

- Ist  $W$  eine parametrische Verteilungsannahme mit Parameterraum  $\Theta$ , so existiert offensichtlich immer auch eine (äquivalente) Darstellung von  $H_0$  und  $H_1$  in der Gestalt

$$H_0 : \theta \in \Theta_0 \quad \text{gegen} \quad H_1 : \theta \in \Theta_1 := \Theta \setminus \Theta_0$$

für eine Teilmenge  $\Theta_0$  des Parameterraums  $\Theta$  mit  $\emptyset \neq \Theta_0 \subsetneq \Theta$ .

- Im Beispiel:  $W = \{N(\mu, 4^2) \mid \mu \in \Theta = \mathbb{R}\}$ ,  $\Theta_0 = \{500\}$
- Hypothesenformulierung damit z.B. in der folgenden Form möglich:

$$H_0 : \mu = \mu_0 = 500 \quad \text{gegen} \quad H_1 : \mu \neq \mu_0 = 500$$

- Hypothesentests bei parametrischer Verteilungsannahme heißen auch **parametrische (Hypothesen-)Tests**.
- Parametrische Tests heißen (für  $\Theta \subseteq \mathbb{R}$ ) **zweiseitig**, wenn  $\Theta_1$  **links und rechts** von  $\Theta_0$  liegt, **einseitig** sonst (Im Beispiel: zweiseitiger Test).
- Einseitige Tests heißen **linksseitig**, wenn  $\Theta_1$  links von  $\Theta_0$  liegt, **rechtsseitig** sonst.

# Teststatistik und Ablehnungsbereich

- Nach Präzisierung der Fragestellung in den Hypothesen benötigt man nun eine geeignete **Entscheidungsregel**, die *im Prinzip* jeder möglichen Stichprobenrealisation (aus dem Stichprobenraum  $\mathcal{X}$ ) eine Entscheidung **entweder** für  $H_0$  **oder** für  $H_1$  zuordnet.
- In der Praxis Entscheidung (fast) immer in 3 Schritten:
  - ① „Zusammenfassung“ der für die Entscheidungsfindung relevanten Stichprobeninformation mit einer geeigneten Stichprobenfunktion, der sog. **Teststatistik** oder **Testgröße**  $T$ .
  - ② Angabe eines **Ablehnungsbereichs** bzw. **kritischen Bereichs**  $K$ , in den die Teststatistik **bei Gültigkeit von  $H_0$**  nur mit einer typischerweise kleinen Wahrscheinlichkeit (durch eine obere Grenze  $\alpha$  beschränkt) fallen darf.
  - ③ Entscheidung **gegen  $H_0$**  bzw. für  $H_1$ , falls realisierter Wert der Teststatistik in den **Ablehnungsbereich** bzw. **kritischen Bereich**  $K$  fällt, also  $T \in K$  gilt (für  $H_0$ , falls  $T \notin K$ ).
- Konstruktion des kritischen Bereichs  $K$  in Schritt ② gerade so, dass **Wahrscheinlichkeit für Fehler 1. Art beschränkt** bleibt durch ein vorgegebenes **Signifikanzniveau** (auch „Irrtumswahrscheinlichkeit“)  $\alpha$ .
- Konstruktion meist so, dass Niveau  $\alpha$  gerade eben eingehalten wird (also **kleinste** obere Schranke für die Fehlerwahrscheinlichkeit 1. Art ist).

- Für Konstruktion des kritischen Bereichs wesentlich:  
**Analyse der Verteilung der Teststatistik, insbesondere falls  $H_0$  gilt!**
- Im Beispiel:**

① Teststatistik:  $N = \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n}$

Verteilung:  $N \sim N\left(\frac{\mu - \mu_0}{\sigma} \sqrt{n}, 1\right)$ , also insbesondere

$N \sim N(0, 1)$  falls  $H_0$  (also  $\mu = \mu_0$ ) gilt.

② Kritischer Bereich:  $K = \left(-\infty, -N_{1-\frac{\alpha}{2}}\right) \cup \left(N_{1-\frac{\alpha}{2}}, \infty\right)$

Wahrscheinlichkeit der Ablehnung von  $H_0$  (abhängig vom Parameter  $\mu$ ):

$$P\{N \in K\} = \Phi\left(-N_{1-\frac{\alpha}{2}} - \frac{\mu - \mu_0}{\sigma} \sqrt{n}\right) + 1 - \Phi\left(N_{1-\frac{\alpha}{2}} - \frac{\mu - \mu_0}{\sigma} \sqrt{n}\right)$$

- Die Zuordnung  $G : \Theta \rightarrow \mathbb{R}$ ;  $G(\theta) = P\{T \in K\}$  heißt (allgemein) auch **Gütefunktion** des Tests. Im Beispiel also:

$$G(\mu) = \Phi\left(-N_{1-\frac{\alpha}{2}} - \frac{\mu - \mu_0}{\sigma} \sqrt{n}\right) + 1 - \Phi\left(N_{1-\frac{\alpha}{2}} - \frac{\mu - \mu_0}{\sigma} \sqrt{n}\right)$$

- Mit der Gütefunktion können also offensichtlich
  - ▶ Fehlerwahrscheinlichkeiten 1. Art (für  $\theta \in \Theta_0$ ) direkt durch  $G(\theta)$  und
  - ▶ Fehlerwahrscheinlichkeiten 2. Art (für  $\theta \in \Theta_1$ ) durch  $1 - G(\theta)$

berechnet werden!

- Berechnung der Eintrittswahrscheinlichkeiten  $EW$  mit Gütefunktion  $G(\theta)$ :

	Tatsächliche Situation: $\theta \in \Theta_0$ ( $H_0$ wahr)	Tatsächliche Situation: $\theta \notin \Theta_0$ ( $H_0$ falsch)
Entscheidung für $H_0$ ( $\theta \in \Theta_0$ )	richtige Entscheidung $EW : 1 - G(\theta)$	Fehler 2. Art $EW : 1 - G(\theta)$
Entscheidung gegen $H_0$ ( $\theta \notin \Theta_0$ )	Fehler 1. Art $EW : G(\theta)$	richtige Entscheidung $EW : G(\theta)$

- Welche Teststatistik geeignet ist und wie die Teststatistik dann verteilt ist, hängt nicht nur von der Problemformulierung (Hypothesen), sondern oft auch von der Verteilungsannahme ab!
- Test aus Beispiel zum Beispiel *exakt* anwendbar, falls  $Y \sim N(\mu, \sigma^2)$  **mit bekannter Varianz**, und *approximativ* anwendbar, wenn  $Y$  beliebig verteilt ist **mit bekannter Varianz** (Güte der Näherung abhängig von  $n$  sowie Verteilung von  $Y$ !)
- Test aus Beispiel heißt auch „zweiseitiger Gauß-Test für den Mittelwert einer Zufallsvariablen mit bekannter Varianz“.



# Zweiseitiger Gauß-Test

für den Mittelwert einer Zufallsvariablen mit bekannter Varianz

## Anwendung

- als **exakter** Test, falls  $Y$  normalverteilt und  $\text{Var}(Y) = \sigma^2$  bekannt,
- als **approximativer** Test, falls  $Y$  beliebig verteilt mit bekannter Varianz  $\sigma^2$ .

„Testrezept“ des **zweiseitigen Tests**:

- Hypothesen:  $H_0 : \mu = \mu_0$  gegen  $H_1 : \mu \neq \mu_0$  für ein vorgegebenes  $\mu_0 \in \mathbb{R}$ .
- Teststatistik:

$$N := \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n} \text{ mit } N \sim N(0, 1) \text{ (bzw. } N \overset{\bullet}{\sim} N(0, 1)), \text{ falls } H_0 \text{ gilt } (\mu = \mu_0).$$

- Kritischer Bereich zum Signifikanzniveau  $\alpha$ :

$$K = \left(-\infty, -N_{1-\frac{\alpha}{2}}\right) \cup \left(N_{1-\frac{\alpha}{2}}, \infty\right)$$

- Berechnung der realisierten Teststatistik  $N$
- Entscheidung:  $H_0$  ablehnen  $\Leftrightarrow N \in K$ .

## Beispiel: Qualitätskontrolle (Länge von Stahlstiften)

- Untersuchungsgegenstand: Weicht die mittlere Länge der von einer bestimmten Maschine produzierten Stahlstifte von der Solllänge  $\mu_0 = 10$  (in [cm]) ab, so dass die Produktion gestoppt werden muss?
- Annahmen: Für Länge  $Y$  der produzierten Stahlstifte gilt:  $Y \sim N(\mu, 0.4^2)$
- Stichprobeninformation: Realisation einer einfachen Stichprobe vom Umfang  $n = 64$  zu  $Y$  liefert Stichprobenmittel  $\bar{x} = 9.7$ .
- Gewünschtes Signifikanzniveau (max. Fehlerwahrscheinlichkeit 1. Art):  $\alpha = 0.05$

Geeigneter Test:

### (Exakter) Gauß-Test für den Mittelwert bei bekannter Varianz

- 1 Hypothesen:  $H_0 : \mu = \mu_0 = 10$  gegen  $H_1 : \mu \neq \mu_0 = 10$
- 2 Teststatistik:  $N = \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n} \sim N(0, 1)$ , falls  $H_0$  gilt ( $\mu = \mu_0$ )
- 3 Kritischer Bereich zum Niveau  $\alpha = 0.05$ :  
 $K = (-\infty, -N_{0.975}) \cup (N_{0.975}, \infty) = (-\infty, -1.96) \cup (1.96, \infty)$
- 4 Realisierter Wert der Teststatistik:  $N = \frac{9.7 - 10}{0.4} \sqrt{64} = -6$
- 5 Entscheidung:  $N \in K \rightsquigarrow H_0$  wird abgelehnt und die Produktion gestoppt.

# Einseitige Gauß-Tests

## Wahl der Hypothesen

- Bei zweiseitigem Test: Hypothesentest zu

$$H_0 : \mu \neq \mu_0 \quad \text{gegen} \quad H_1 : \mu = \mu_0$$

zwar konstruierbar, aber ohne praktische Bedeutung.

- Neben zweiseitigem Test zwei (symmetrische) einseitige Varianten:

$$H_0 : \mu \leq \mu_0 \quad \text{gegen} \quad H_1 : \mu > \mu_0$$

$$H_0 : \mu \geq \mu_0 \quad \text{gegen} \quad H_1 : \mu < \mu_0$$

- Konstruktion der Tests beschränkt Wahrscheinlichkeit,  $H_0$  **fälschlicherweise abzulehnen**. Entscheidung zwischen beiden Varianten daher wie folgt:

$H_0$  : **Nullhypothese** ist in der Regel die Aussage, die von vornherein als glaubwürdig gilt und die man beibehält, wenn das Stichprobenergebnis bei Gültigkeit von  $H_0$  nicht sehr untypisch bzw. überraschend ist.

$H_1$  : **Gegenhypothese** ist in der Regel die Aussage, die man statistisch absichern möchte und für deren Akzeptanz man hohe Evidenz fordert.

Die Entscheidung für  $H_1$  hat typischerweise erhebliche Konsequenzen, so dass man das Risiko einer fälschlichen Ablehnung von  $H_0$  zugunsten von  $H_1$  kontrollieren will.

- Auch für einseitige Tests fasst Teststatistik

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

die empirische Information über den Erwartungswert  $\mu$  geeignet zusammen.

- Allerdings gilt nun offensichtlich
  - ▶ im Falle des **rechtsseitigen** Tests von

$$H_0 : \mu \leq \mu_0 \quad \text{gegen} \quad H_1 : \mu > \mu_0 ,$$

dass **große (insbesondere positive)** Realisationen von  $N$  gegen  $H_0$  und für  $H_1$  sprechen, sowie

- ▶ im Falle des **linksseitigen** Tests von

$$H_0 : \mu \geq \mu_0 \quad \text{gegen} \quad H_1 : \mu < \mu_0 ,$$

dass **kleine (insbesondere negative)** Realisationen von  $N$  gegen  $H_0$  und für  $H_1$  sprechen.

- Noch nötig zur Konstruktion der Tests:  
Geeignetes Verfahren zur Wahl der **kritischen Bereiche** so, dass Wahrscheinlichkeit für Fehler 1. Art durch vorgegebenes Signifikanzniveau  $\alpha$  beschränkt bleibt.

## Kritischer Bereich (rechtsseitiger Test)

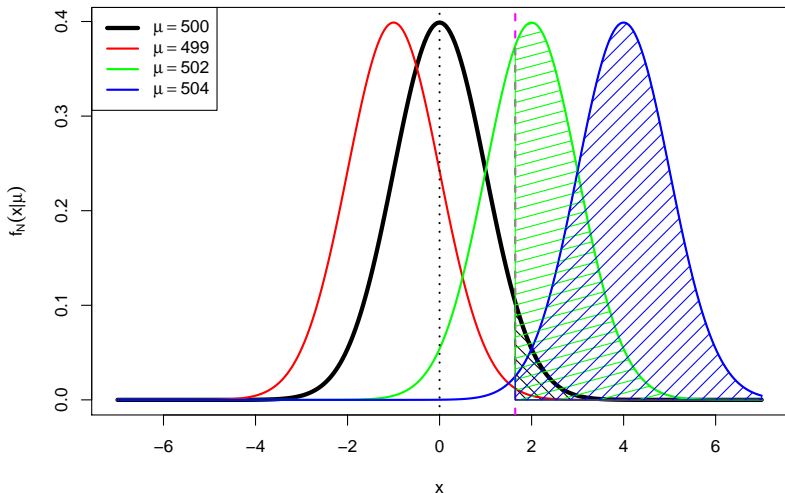
- Für **rechtsseitigen** Test muss also zur Konstruktion des kritischen Bereichs ein **kritischer Wert** bestimmt werden, den die Teststatistik  $N$  im Fall der Gültigkeit von  $H_0$  **maximal** mit einer Wahrscheinlichkeit von  $\alpha$  überschreitet.
- Gesucht ist also ein Wert  $k_\alpha$  mit  $P\{N \in (k_\alpha, \infty)\} \leq \alpha$  **für alle**  $\mu \leq \mu_0$ .
- Offensichtlich wird  $P\{N \in (k_\alpha, \infty)\}$  mit wachsendem  $\mu$  größer, es genügt also, die Einhaltung der Bedingung  $P\{N \in (k_\alpha, \infty)\} \leq \alpha$  für das **größtmögliche**  $\mu$  mit der Eigenschaft  $\mu \leq \mu_0$ , also  $\mu = \mu_0$ , zu gewährleisten.
- Um die Fehlerwahrscheinlichkeit 2. Art unter Einhaltung der Bedingung an die Fehlerwahrscheinlichkeit 1. Art möglichst klein zu halten, wird  $k_\alpha$  gerade so gewählt, dass  $P\{N \in (k_\alpha, \infty)\} = \alpha$  für  $\mu = \mu_0$  gilt.
- Wegen  $N \sim N(0, 1)$  für  $\mu = \mu_0$  erhält man hieraus

$$\begin{aligned}
 & P\{N \in (k_\alpha, \infty)\} = \alpha \\
 \Leftrightarrow & \quad 1 - P\{N \in (-\infty, k_\alpha)\} = \alpha \\
 \Leftrightarrow & \quad \Phi(k_\alpha) = 1 - \alpha \\
 \Leftrightarrow & \quad k_\alpha = N_{1-\alpha}
 \end{aligned}$$

und damit insgesamt den kritischen Bereich  $K = (N_{1-\alpha}, \infty)$  für den rechtsseitigen Test.

# Beispiel für Verteilungen von $N$

Rechtsseitiger Test ( $\mu_0 = 500$ ) zum Signifikanzniveau  $\alpha = 0.05$



# Rechtsseitiger Gauß-Test

für den Mittelwert einer Zufallsvariablen mit bekannter Varianz

## Anwendung

- als **exakter** Test, falls  $Y$  normalverteilt und  $\text{Var}(Y) = \sigma^2$  bekannt,
- als **approximativer** Test, falls  $Y$  beliebig verteilt mit bekannter Varianz  $\sigma^2$ .

„Testrezept“ des **rechtsseitigen Tests**:

- 1 Hypothesen:  $H_0 : \mu \leq \mu_0$  gegen  $H_1 : \mu > \mu_0$  für ein vorgegebenes  $\mu_0 \in \mathbb{R}$ .
- 2 Teststatistik:

$$N := \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n} \text{ mit } N \sim N(0, 1) \text{ (} N \overset{\bullet}{\sim} N(0, 1)\text{)}, \text{ falls } H_0 \text{ gilt (mit } \mu = \mu_0\text{)}.$$

- 3 Kritischer Bereich zum Signifikanzniveau  $\alpha$ :

$$K = (N_{1-\alpha}, \infty)$$

- 4 Berechnung der realisierten Teststatistik  $N$
- 5 Entscheidung:  $H_0$  ablehnen  $\Leftrightarrow N \in K$ .

## Kritischer Bereich (linksseitiger Test)

- Für **linksseitigen** Test muss zur Konstruktion des kritischen Bereichs ein **kritischer Wert** bestimmt werden, den die Teststatistik  $N$  im Fall der Gültigkeit von  $H_0$  **maximal** mit einer Wahrscheinlichkeit von  $\alpha$  unterschreitet.
- Gesucht ist also ein Wert  $k_\alpha$  mit  $P\{N \in (-\infty, k_\alpha)\} \leq \alpha$  **für alle**  $\mu \geq \mu_0$ .
- Offensichtlich wird  $P\{N \in (-\infty, k_\alpha)\}$  mit fallendem  $\mu$  größer, es genügt also, die Einhaltung der Bedingung  $P\{N \in (-\infty, k_\alpha)\} \leq \alpha$  für das **kleinstmögliche**  $\mu$  mit  $\mu \geq \mu_0$ , also  $\mu = \mu_0$ , zu gewährleisten.
- Um die Fehlerwahrscheinlichkeit 2. Art unter Einhaltung der Bedingung an die Fehlerwahrscheinlichkeit 1. Art möglichst klein zu halten, wird  $k_\alpha$  gerade so gewählt, dass  $P\{N \in (-\infty, k_\alpha)\} = \alpha$  für  $\mu = \mu_0$  gilt.
- Wegen  $N \sim N(0, 1)$  für  $\mu = \mu_0$  erhält man hieraus

$$P\{N \in (-\infty, k_\alpha)\} = \alpha$$

$$\Leftrightarrow \Phi(k_\alpha) = \alpha$$

$$\Leftrightarrow k_\alpha = N_\alpha$$

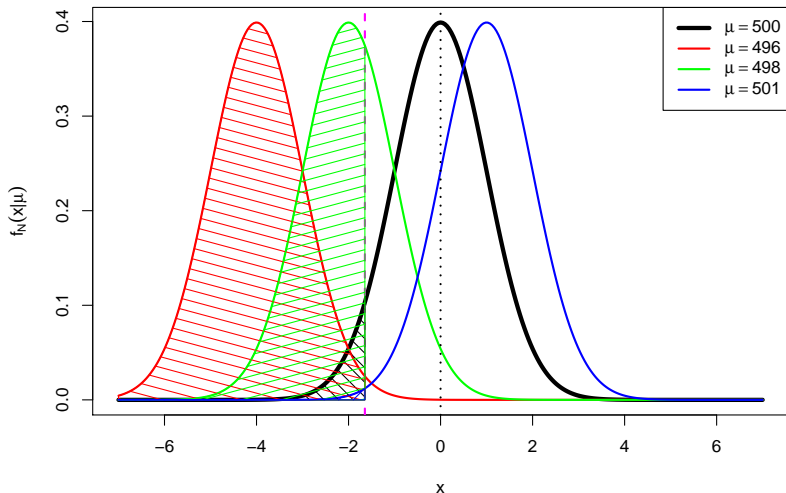
$$\Leftrightarrow k_\alpha = -N_{1-\alpha}$$

und damit insgesamt den kritischen Bereich  $K = (-\infty, -N_{1-\alpha})$  für den linksseitigen Test.



# Beispiel für Verteilungen von $N$

Linksseitiger Test ( $\mu_0 = 500$ ) zum Signifikanzniveau  $\alpha = 0.05$



# Linksseitiger Gauß-Test

für den Mittelwert einer Zufallsvariablen mit bekannter Varianz

## Anwendung

- als **exakter** Test, falls  $Y$  normalverteilt und  $\text{Var}(Y) = \sigma^2$  bekannt,
- als **approximativer** Test, falls  $Y$  beliebig verteilt mit bekannter Varianz  $\sigma^2$ .

„Testrezept“ des **linksseitigen Tests**:

- Hypothesen:  $H_0 : \mu \geq \mu_0$  gegen  $H_1 : \mu < \mu_0$  für ein vorgegebenes  $\mu_0 \in \mathbb{R}$ .
- Teststatistik:

$$N := \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n} \text{ mit } N \sim N(0, 1) \text{ (} N \overset{\bullet}{\sim} N(0, 1)\text{)}, \text{ falls } H_0 \text{ gilt (mit } \mu = \mu_0\text{)}.$$

- Kritischer Bereich zum Signifikanzniveau  $\alpha$ :

$$K = (-\infty, -N_{1-\alpha})$$

- Berechnung der realisierten Teststatistik  $N$
- Entscheidung:  $H_0$  ablehnen  $\Leftrightarrow N \in K$ .

# Gütefunktionen einseitiger Gauß-Tests

- Gütefunktion allgemein:  $G(\theta) = P\{T \in K\}$
- Für **rechtsseitigen** Gauß-Test:
  - ▶  $G(\mu) = P\{N \in (N_{1-\alpha}, \infty)\}$
  - ▶ Mit  $N \sim N\left(\frac{\mu - \mu_0}{\sigma} \sqrt{n}, 1\right)$  erhält man

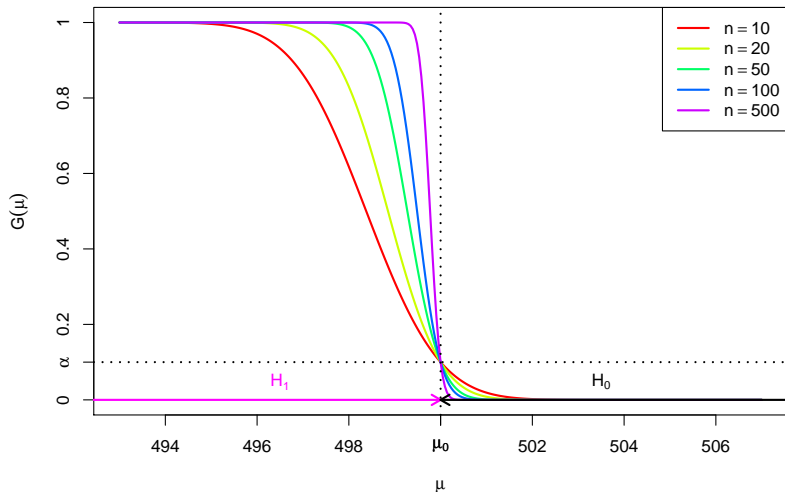
$$\begin{aligned}
 P\{N \in (N_{1-\alpha}, \infty)\} &= 1 - P\{N \leq N_{1-\alpha}\} \\
 &= 1 - \Phi\left(N_{1-\alpha} - \frac{\mu - \mu_0}{\sigma} \sqrt{n}\right) \\
 &= \Phi\left(\frac{\mu - \mu_0}{\sigma} \sqrt{n} - N_{1-\alpha}\right)
 \end{aligned}$$

- Für **linksseitigen** Gauß-Test:
  - ▶  $G(\mu) = P\{N \in (-\infty, -N_{1-\alpha})\}$
  - ▶ Mit  $N \sim N\left(\frac{\mu - \mu_0}{\sigma} \sqrt{n}, 1\right)$  erhält man hier

$$\begin{aligned}
 P\{N \in (-\infty, -N_{1-\alpha})\} &= P\{N < -N_{1-\alpha}\} \\
 &= \Phi\left(-N_{1-\alpha} - \frac{\mu - \mu_0}{\sigma} \sqrt{n}\right)
 \end{aligned}$$

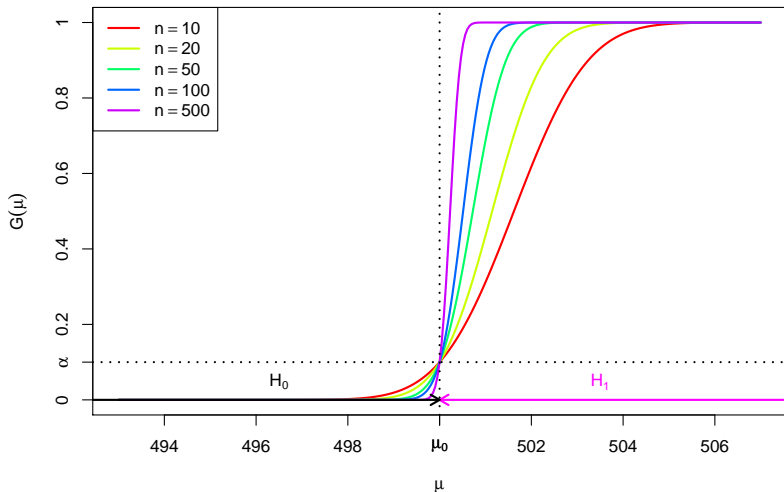
# Beispiel für Gütefunktionen

Linksseitiger Test ( $\mu_0 = 500$ ) zum Signifikanzniveau  $\alpha = 0.10$



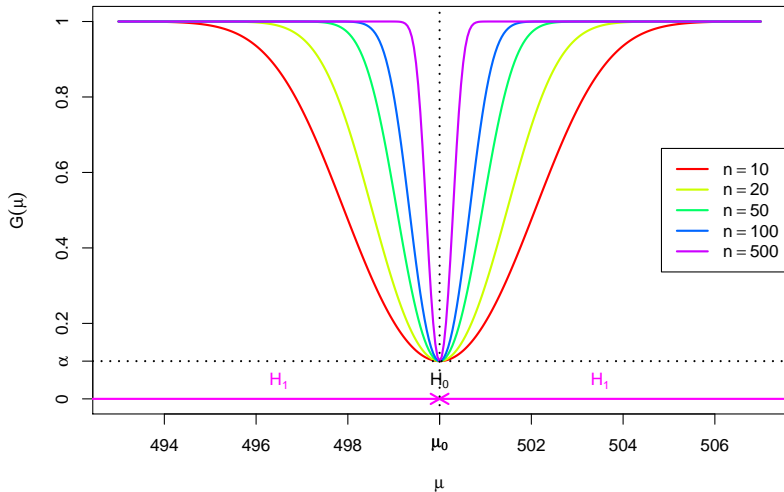
# Beispiel für Gütefunktionen

Rechtsseitiger Test ( $\mu_0 = 500$ ) zum Signifikanzniveau  $\alpha = 0.10$



# Beispiel für Gütefunktionen

Zweiseitiger Test ( $\mu_0 = 500$ ) zum Signifikanzniveau  $\alpha = 0.10$



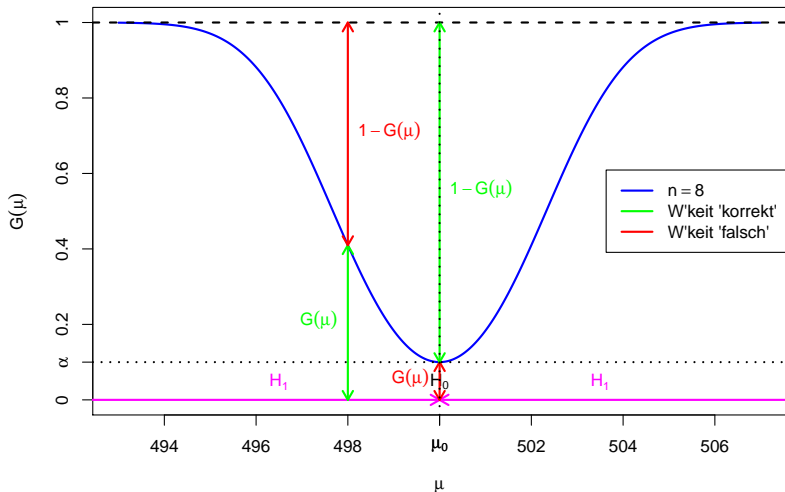
# Gütefunktion und Fehlerwahrscheinlichkeiten

für Gauß-Tests auf den Mittelwert bei bekannter Varianz

- Entscheidungsregel (nicht nur) bei Gauß-Tests stets:  $H_0$  ablehnen  $\Leftrightarrow N \in K$
- Gütefunktion  $G(\mu)$  gibt also für Gauß-Tests auf den Mittelwert bei bekannter Varianz zu jedem möglichen wahren Mittelwert  $\mu$  die Wahrscheinlichkeit an, eine Stichprobenrealisation zu erhalten, die zu einer Entscheidung **gegen**  $H_0$  führt.
- Dies kann — abhängig davon, ob für  $\mu$   $H_0$  oder  $H_1$  zutreffend ist — also die Wahrscheinlichkeit einer falschen bzw. richtigen Entscheidung sein (vgl. Folie 105).
- Gängige Abkürzung
  - ▶ für Fehlerwahrscheinlichkeiten 1. Art:  $\alpha(\mu)$  für  $\mu \in \Theta_0$ ,
  - ▶ für Fehlerwahrscheinlichkeiten 2. Art:  $\beta(\mu)$  für  $\mu \in \Theta_1$ .
- Für  $\mu \in \Theta_0$  (also bei Gültigkeit der Nullhypothese für  $\mu$ ) gilt also:
  - ▶ Fehlerwahrscheinlichkeit 1. Art:  $\alpha(\mu) = G(\mu)$
  - ▶ Wahrscheinlichkeit richtiger Entscheidung:  $1 - G(\mu)$
- Für  $\mu \in \Theta_1$  (also bei Verletzung der Nullhypothese für  $\mu$ ) erhält man:
  - ▶ Fehlerwahrscheinlichkeit 2. Art:  $\beta(\mu) = 1 - G(\mu)$
  - ▶ Wahrscheinlichkeit richtiger Entscheidung:  $G(\mu)$

# Gütefunktion und Fehlerwahrscheinlichkeiten

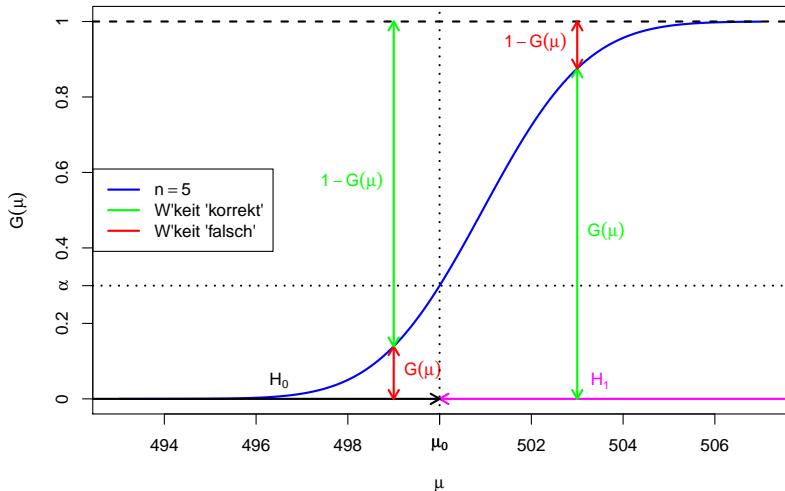
Zweiseitiger Test ( $\mu_0 = 500$ ) zum Signifikanzniveau  $\alpha = 0.10$





# Gütefunktion und Fehlerwahrscheinlichkeiten

Rechtsseitiger Test ( $\mu_0 = 500$ ) zum Signifikanzniveau  $\alpha = 0.30$



# Gütefunktion und Fehlerwahrscheinlichkeiten

Linksseitiger Test ( $\mu_0 = 500$ ) zum Signifikanzniveau  $\alpha = 0.30$

