

Hypothesentests für Erwartungswert und Median

Statistik (Biol./Pharm./HST) – FS 2014



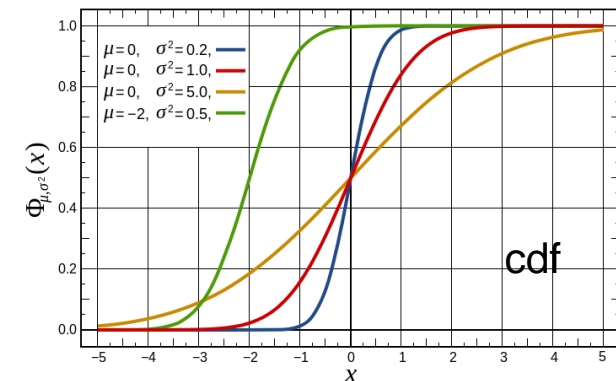
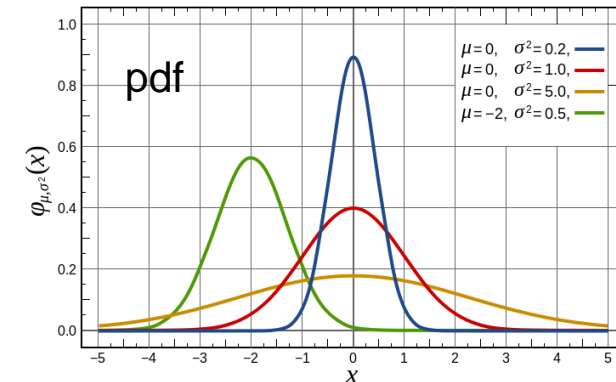
Normalverteilung

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

‘X ist normalverteilt mit Erwartungswert μ und Varianz σ^2 ’

- pdf: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$

- cdf: ???



Zentraler Grenzwertsatz (ZGS)

- Ann: $X_1, \dots, X_n \sim F$ iid; $E(X_i) = \mu$, $Var(X_i) = \sigma_X^2$

$$\bar{X}_n \approx N(\mu, \underbrace{\frac{\sigma_X^2}{n}}_{\text{aus GGZ}})$$

↓
neu

oder äquivalent mit $S_n = X_1 + \dots + X_n$:

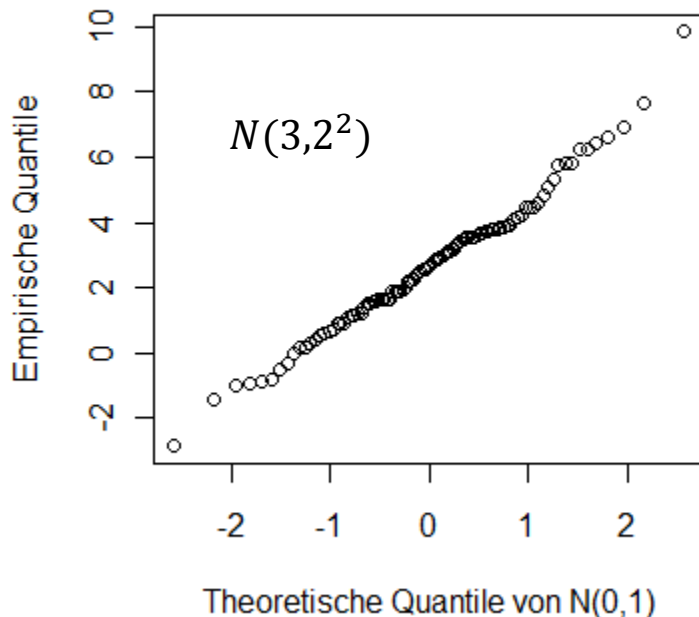
$$S_n \approx N(n\mu, n\sigma_X^2)$$

Wie prüft man, ob eine Normalverteilung vorliegt?

- Histogramm der Daten mit pdf vergleichen
Schwierig kleine Abweichungen zu erkennen
- Einfacher: QQ-Plot – Theoretische Quantile gegen Empirische Quantile

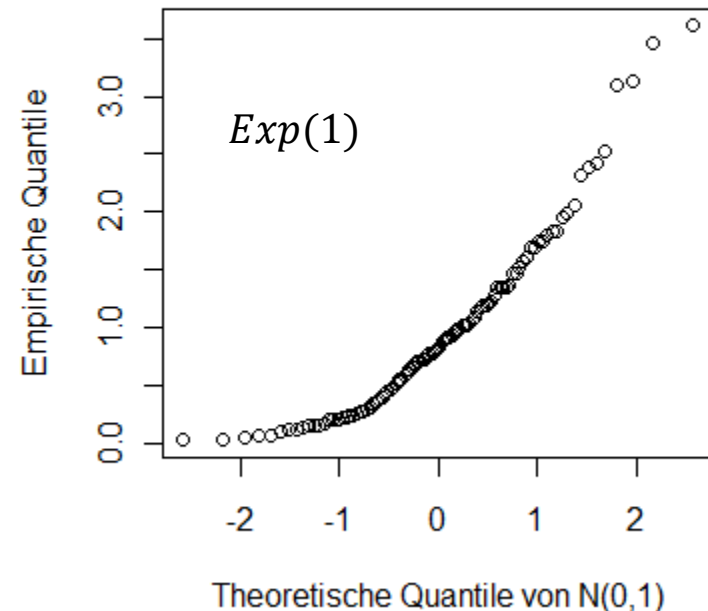
Gerade:

Normalverteilung OK
Normal Q-Q Plot



Krümmung:

Keine Normalverteilung
Normal Q-Q Plot



Reaktionszeit

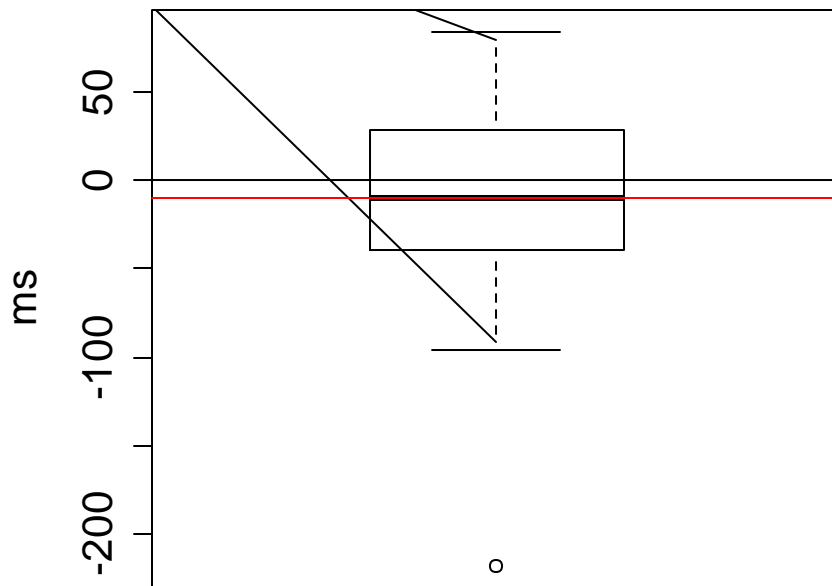


- Reagiert ein Rechtshänder mit rechts schneller als mit links ? (analog für Linkshänder)
- Experiment:
 - **Population:** Alle StudentInnen der VL
 - **Stichprobe:** 50 zufällige StudentInnen angeschrieben
 - Reaktionszeittest online:
<http://www.bbc.co.uk/science/humanbody/sleep/sheep/>
 - Mit jeder Hand einmal ausprobieren
 - Dann mit jeder Hand nocheinmal
 - **Reihenfolge randomisiert** (Geburtstag)
 - **Robust:** Bei beiden Messungen das beste und schlechteste Resultat streichen, dann mitteln und an mich senden
 - Berechne “Nebenhand – Haupthand”
 - Z.B.: Haupthand – 227 ms, Nebenhand – 248 ms
Differenz = 248 ms – 227 ms = 21 ms
 - **Anreiz:** Verlose Gutschein unter Teilnehmern: Rücklauf erhöhen

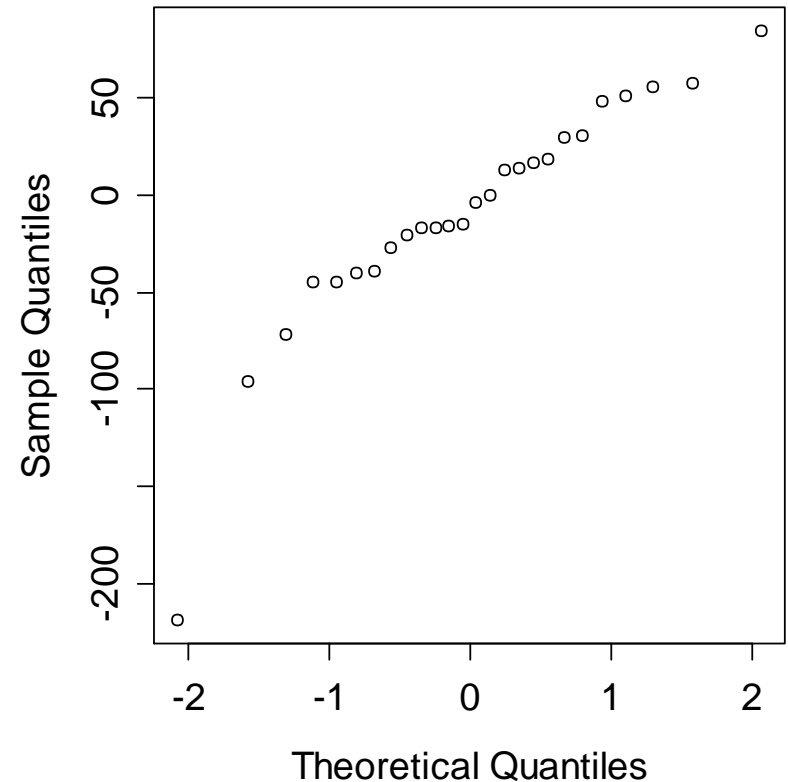
Ergebnis

- 50 StudentInnen angeschrieben
- Rücklauf: 26
- **Nebenhand ist im Mittel 10 ms schneller**

Zeit NH - Zeit HH



Normal Q-Q Plot



Stichprobe vs. Population

- In der Stichprobe war die Nebenhand im Mittel um 10 ms schneller
- Bedeutet das, dass die Nebenhand **auch in der ganzen Bevölkerung** im Mittel schneller ist?
- Antwort darauf:
z-Test, t-Test, Wilcoxon-Test, Vorzeichen-Test

z-Test: σ_X bekannt

1. **Modell:** X_i ist eine kontinuierliche Messgrösse;

$X_1, \dots, X_n \text{ iid } \mathcal{N}(\mu, \sigma_X^2), \sigma_X \text{ bekannt}$

2. **Nullhypothese:** $H_0 : \mu = \mu_0,$

Alternative: $H_A : \mu \neq \mu_0$ (oder “<” oder “>”)

3. **Teststatistik:**

$$Z = \frac{(\bar{X}_n - \mu_0)}{\sigma_{\bar{X}_n}} = \frac{\sqrt{n}(\bar{X}_n - \mu_0)}{\sigma_X} = \frac{\text{beobachtet} - \text{erwartet}}{\text{Standardfehler}}.$$

Verteilung der Teststatistik unter H_0 : $Z \sim \mathcal{N}(0, 1)$

4. **Signifikanzniveau:** α

5. **Verwerfungsbereich für die Teststatistik:**

$$K = (-\infty, -\Phi^{-1}(1 - \frac{\alpha}{2})] \cup [\Phi^{-1}(1 - \frac{\alpha}{2}), \infty) \text{ bei } H_A : \mu \neq \mu_0,$$

$$K = (-\infty, -\Phi^{-1}(1 - \alpha)] \text{ bei } H_A : \mu < \mu_0,$$

$$K = [\Phi^{-1}(1 - \alpha), \infty) \text{ bei } H_A : \mu > \mu_0.$$

6. **Testentscheid:** Überprüfe, ob der beobachtete Wert der Teststatistik im Verwerfungsbereich liegt.

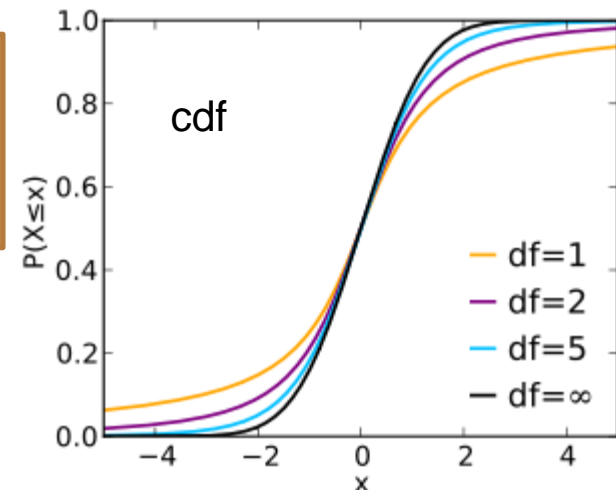
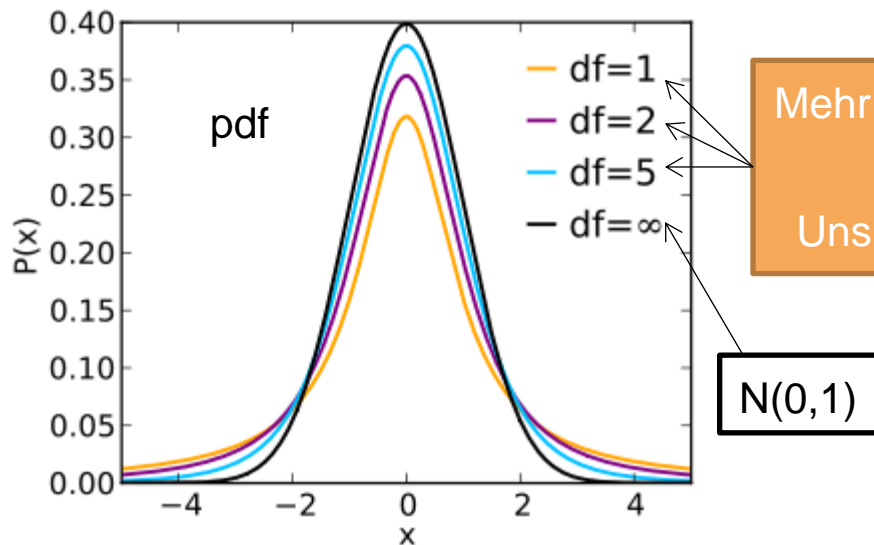
Problem in Praxis: σ_X ist unbekannt !

- Schätze Varianz: $\hat{\sigma}_X^2 = \frac{1}{n-1} \sum (X_i - \bar{X}_n)$
- Neue Teststatistik: $T = \frac{\bar{X}_n - \mu_0}{\frac{\hat{\sigma}_X}{\sqrt{n}}}$
- Verteilung von T, falls H_0 stimmt: $T \sim t_{n-1}$

“Student’s” t-Verteilung – Zoo Teil 3



- Annahme: $X_1, X_2, \dots, X_n \sim N(\mu, \sigma_X^2)$ und unabhängig
- $\hat{\sigma}_X^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$ ist geschätzte Varianz
- $T = (\bar{X}_n - \mu) / (\frac{\hat{\sigma}_X}{\sqrt{n}})$ folgt einer ‘t-Verteilung mit n Freiheitsgraden’, $T \sim t_n$
- Werte sind tabelliert oder im Computer verfügbar
- Falls $n = \infty$: $t_n = N(0,1)$



t-Test: σ_X unbekannt

R: t.test

Macht: power.t.test

1. **Modell:** X_i ist eine kontinuierliche Messgrösse;
 $X_1, \dots, X_n \text{ iid } \mathcal{N}(\mu, \sigma_X^2)$, σ_X wird durch $\widehat{\sigma}_X$ geschätzt
2. **Nullhypothese:** $H_0 : \mu = \mu_0$,
Alternative: $H_A : \mu \neq \mu_0$ (oder “<” oder “>”)
3. **Teststatistik:**

$$T = \frac{(\bar{X}_n - \mu_0)}{\widehat{\sigma}_{\bar{X}_n}} = \frac{\sqrt{n}(\bar{X}_n - \mu_0)}{\widehat{\sigma}_X} = \frac{\text{beobachtet} - \text{erwartet}}{\text{geschätzter Standardfehler}}.$$

Verteilung der Teststatistik unter H_0 : $T \sim t_{n-1}$

4. **Signifikanzniveau:** α
5. **Verwerfungsbereich für die Teststatistik:**

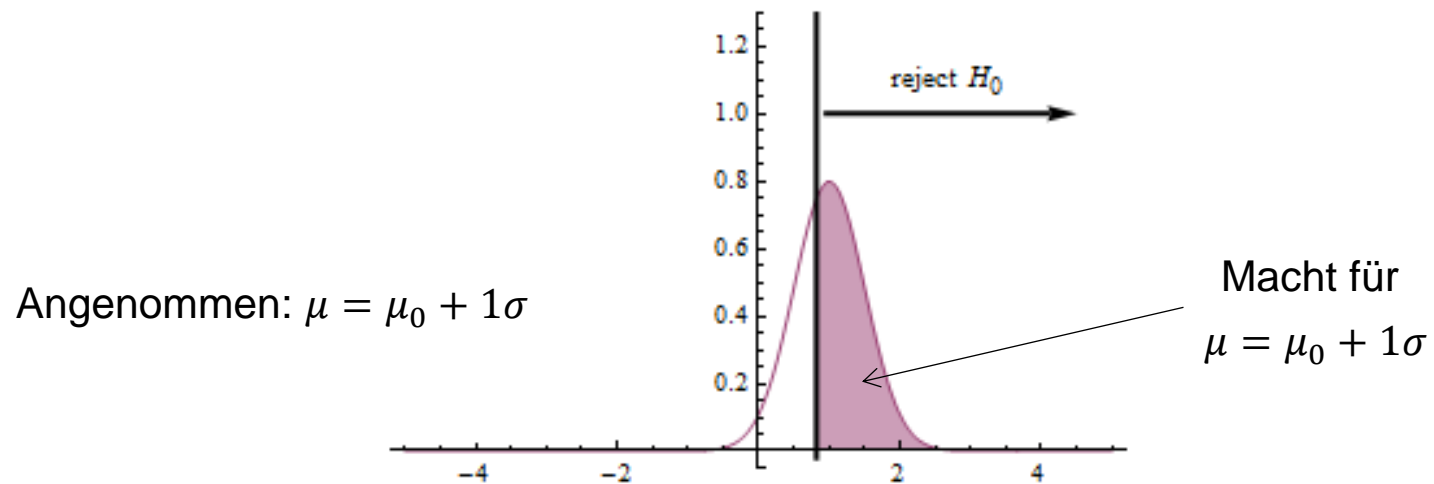
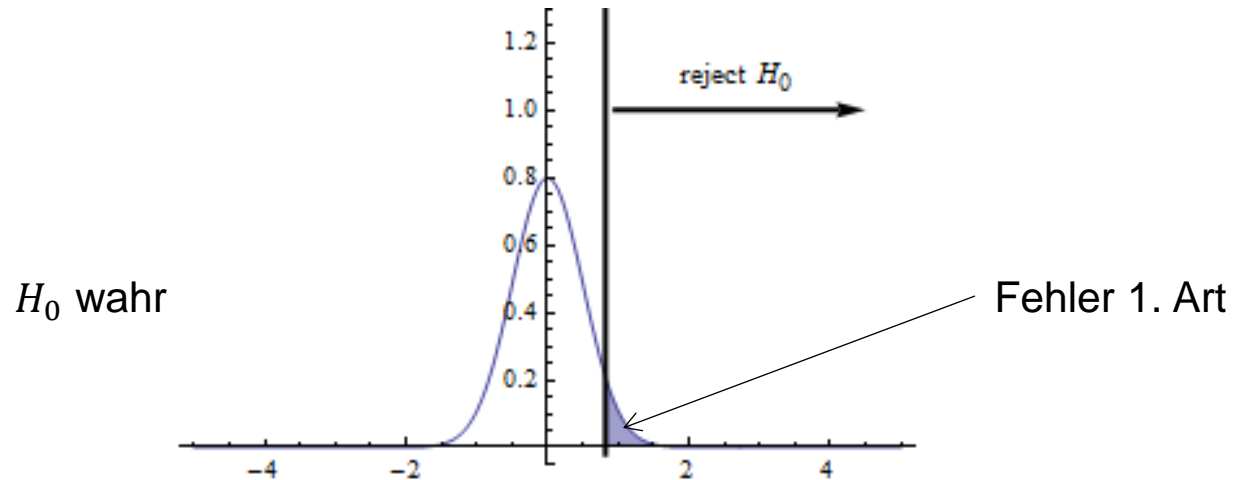
$$K = (-\infty, -t_{n-1; 1-\frac{\alpha}{2}}] \cup [t_{n-1; 1-\frac{\alpha}{2}}, \infty) \text{ bei } H_A : \mu \neq \mu_0,$$

$$K = (-\infty, -t_{n-1; 1-\alpha}] \text{ bei } H_A : \mu < \mu_0,$$

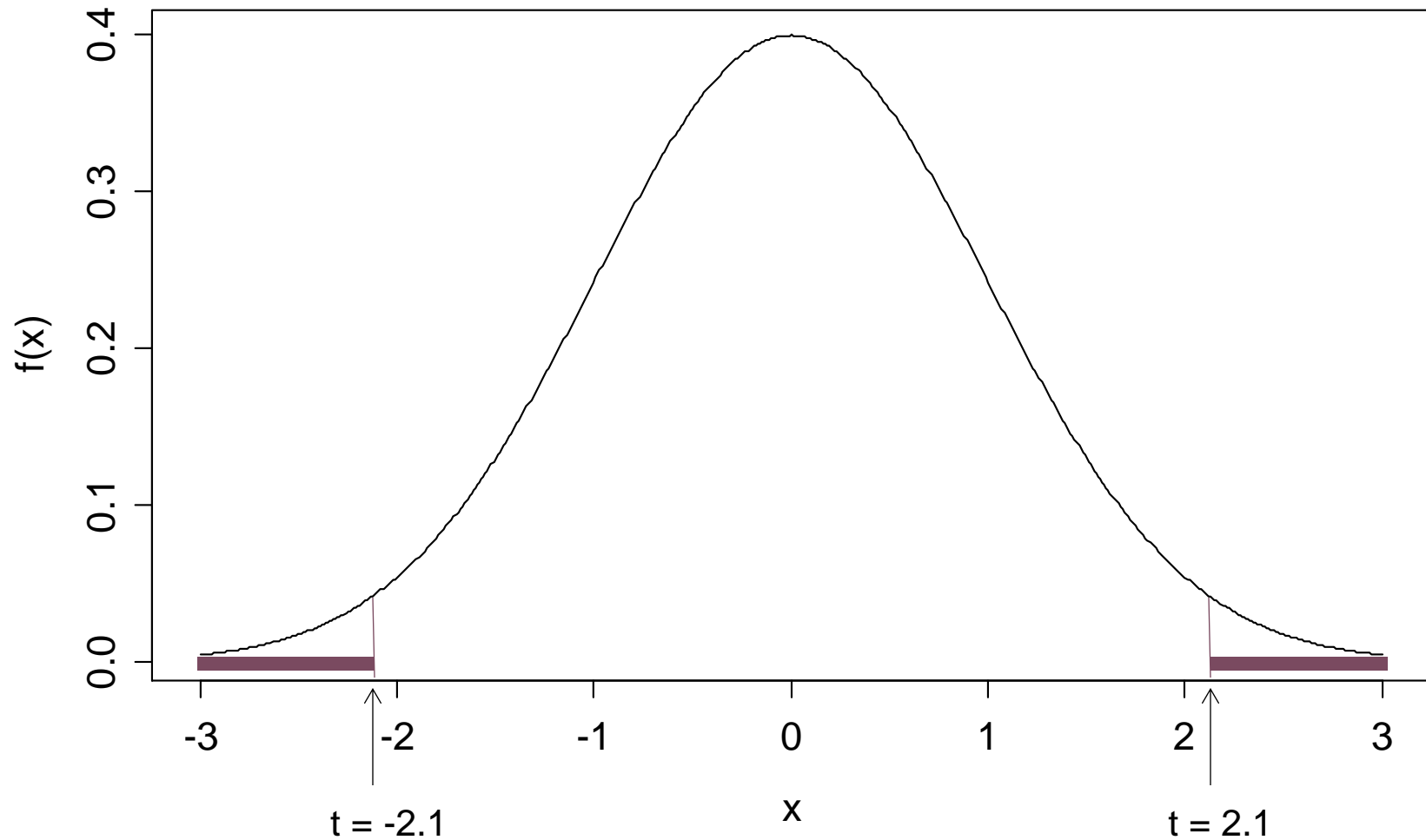
$$K = [t_{n-1; 1-\alpha}, \infty) \text{ bei } H_A : \mu > \mu_0.$$

6. **Testentscheid:** Überprüfe, ob der beobachtete Wert der Teststatistik im Verwerfungsbereich liegt.

Wdh: Fehler 1. Art & Macht



P-Wert (zweiseitiger Test)



$(1-\alpha)$ -Vertrauensintervall für μ

- Äquivalente Definitionen:
 - enthält wahren Wert mit Wa. $1 - \alpha$
 - enthält alle Werte μ_0 , bei denen $H_0: \mu = \mu_0$ vs. $H_A: \mu \neq \mu_0$ mit Sig.niveau α **nicht** verwirft
- Im t-Test, Schritt 5: Nicht verwerfen, falls
$$\left| \frac{\bar{X}_n - \mu}{\frac{\hat{\sigma}_X}{\sqrt{n}}} \right| < t_{n-1; 1-\frac{\alpha}{2}} ; \text{ nach } \mu \text{ auflösen}$$
- VI: $\left[\bar{x}_n - t_{n-1; 1-\frac{\alpha}{2}} * \frac{\hat{\sigma}_x}{\sqrt{n}} ; \bar{x}_n + t_{n-1; 1-\frac{\alpha}{2}} * \frac{\hat{\sigma}_x}{\sqrt{n}} \right]$
- Im Bsp: $\left[-10 - 2.06 * \frac{60}{\sqrt{26}} ; -10 + 2.06 * \frac{60}{\sqrt{26}} \right] =$
 $= [-34; 14] \text{ ms}$

Annahmen und Ausblick

- Stichprobenvarianz bekannt (z-Test)
→ gelöst im t-Test
- Daten normalverteilt (z-Test, t-Test)
→ abgeschwächt im Wilcoxon Test
→ gelöst im Vorzeichen-test
- Beobachtungen unabhängig und gleich verteilt
→ keine einfache Lösung

Vorzeichentest = Binomialtest

1. Modell:

$$X_1, \dots, X_n \quad iid,$$

wobei X_i eine beliebige Verteilung hat.

2. Nullhypothese: $H_0 : \mu = \mu_0$, (μ ist der Median)

Alternative: $H_A : \mu \neq \mu_0$ (oder einseitige Variante)

3. Teststatistik: V : Anzahl X_i s mit ($X_i > \mu_0$)

Verteilung der Teststatistik unter H_0 : $V \sim \text{Bin}(n, \pi_0)$ mit $\pi_0 = 0.5$

4. Signifikanzniveau: α

5. Verwerfungsbereich für die Teststatistik: $K = [0, c_u] \cup [c_o, n]$ falls $H_A :$

$\mu \neq \mu_0$,

Die Grenzen c_u und c_o müssen mit der Binomialverteilung oder der Normalapproximation berechnet werden.

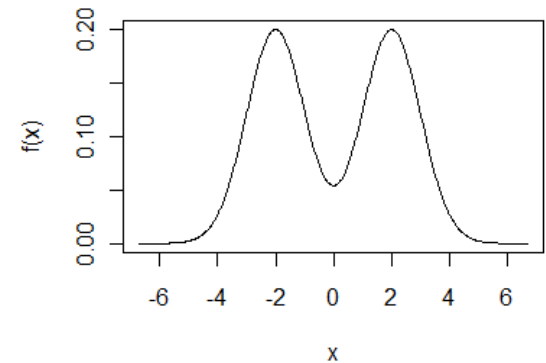
6. Testentscheid: Entscheide, ob der beobachtete Wert der Teststatistik im Verwerfungsbereich der Teststatistik liegt.

Bsp: Vorzeichentest

- Angenommen: $H_0: \mu = \mu_0 = 10$, $H_A: \mu \neq 10$
- Beobachtet: $x_1 = 13, x_2 = 9, x_3 = 17, x_4 = 8, x_5 = 14$
- Vorzeichen von $x_i - \mu_0$: +, -, +, -, +
- Mache Binomialtest mit
 $H_0: \pi = 0.5$, $H_A: \pi \neq 0.5$, $n=5$, $x=3$ (Anzahl '+')
- Der Vorzeichentest kann genau dann verworfen werden, wenn der entsprechende Binomialtest verworfen wird.
- **Vorteil:** Keine Annahme an Verteilung
- **Nachteil:** Kleinere Macht

Wilcoxon-Test: Intuition

- Kompromiss zw. Vorzeichen- und t-Test
- Annahme: $X_i \sim F$ iid, F ist symmetrisch
- Teste Median μ : $H_0: \mu = \mu_0$
(einseitig oder zweiseitig)



Bsp: Wilcoxon-Test

- Bsp: $H_0: \mu_0 = 0$
- Beobachte -1.9, 0.2, 2.9, -4.1, 3.9
- Absolutbeträge: 1.9, 0.2, 2.9, 4.1, 3.9
- Ränge der Absolutbeträge: 2,1,3,5,4
- Rangsumme der positiven Gruppe: $1+3+4=8$
Minimale Rangsumme: 0
Maximale Rangsumme: $1+2+3+4+5 = 15$

- Mit dem Computer:

```
> wilcox.test(c(-1.9, 0.2, 2.9, -4.1, 3.9), mu=0)
```

```
wilcoxon signed rank test
```

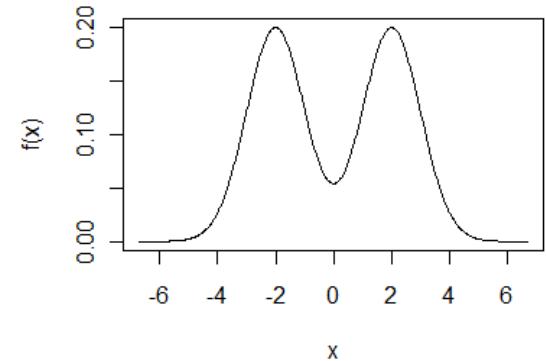
```
data: c(-1.9, 0.2, 2.9, -4.1, 3.9)
```

```
V = 8, p-value = 1
```

```
alternative hypothesis: true location is not equal to 0
```

Wilcoxon-Test: Intuition

- Kompromiss zw. Vorzeichen- und t-Test
- Annahme: $X_i \sim F$ iid, F ist symmetrisch
- Teste Median μ : $H_0: \mu = \mu_0$
(einseitig oder zweiseitig)
- Intuition der Teststatistik
 - Rangiere $|x_i - \mu_0| \rightarrow r_i$
 - Gib Rängen ursprüngliches Vorzeichen von $(x_i - \mu_0)$
("signed ranks")
 - Teststatistik T: Summe aller Ränge, bei denen $(x_i - \mu_0)$ positiv ist
- Falls H_0 stimmt, sollte diese Rangsumme nicht zu gross und nicht zu klein sein



Übersicht der Tests

Test	Annahmen				n_{min} bei $\alpha = 0.05$	Macht für ein Beispiel (1)
	σ_X bekannt	$X_i \sim N$	Symm. Verteilung	iid		
z	x	x	x	x	1	89 %
t		x	x	x	2	79 %
Wilcoxon			x	x	6	79 %
VZ				x	5	48 %

(1): $X_i \sim N(\mu, \sigma^2)$, $n = 10$; $H_0: \mu = 0$; $H_A: \mu \neq 0$; $\alpha = 0.05$

Macht berechnet für konkrete Alternative: $X_i \sim N(1,1)$