

# Geodatenanalyse I: Statistisches Testen

Kathrin Menberg



# Stundenplan

	08:30 – 12:30 Uhr	13:30 – 17:30 Uhr
Montag	Tag 1 / Block 1	Tag 1 / Block 2
Dienstag	Tag 2 / Block 1	Tag 2 / Block 2
Mittwoch	Tag 3 / Block 1	Tag 3 / Block 2
Donnerstag	Tag 4 / Block 1	Tag 4 / Block 2
Freitag	Tag 5 / Block 1	Tag 5 / Block 2

- ▶ 2.1 Einführung und Deskriptive Statistik
- ▶ **2.2 Statistischen Testen**
- ▶ 2.3 Schließende Statistik und Wahrscheinlichkeiten

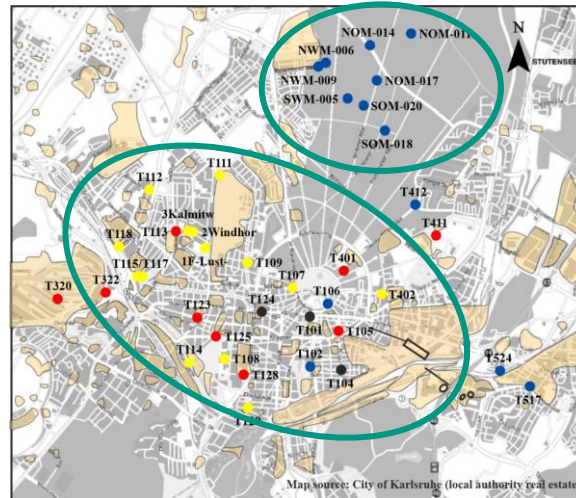
## Lernziele Block 2.2

Am Ende der Stunde werden die Teilnehmer :

- ▶ ... die theoretischen Grundlagen des klassischen statistischen Testens kennen.
- ▶ ... verschiedene statistische Tests für unterschiedliche Zwecke in Python kennen und anwenden können.
- ▶ ... die Testergebnisse in Bezug auf Signifikanz und p-Wert bewerten und kritisch diskutieren können.

# Problemstellung

- Grundwassertemperaturen in Karlsruhe und im Hardtwald:



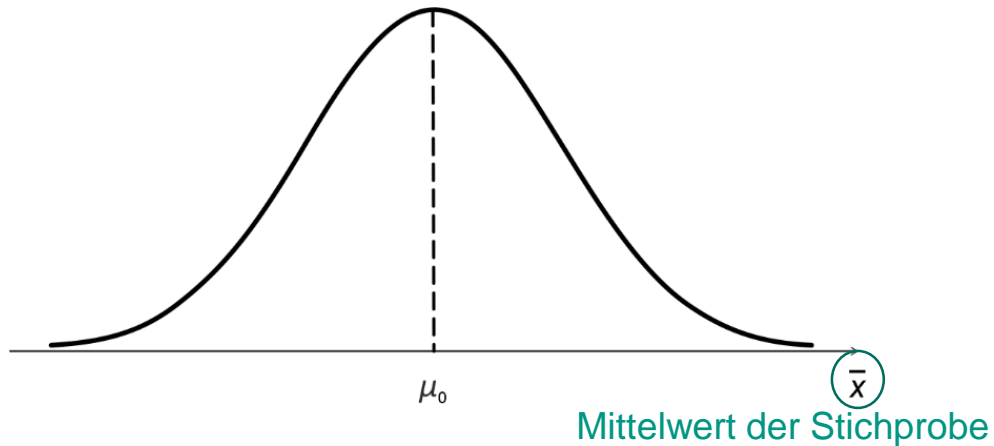
- ... passt unsere ursprüngliche Vermutung, dass die Temperatur im Wald 11°C beträgt?
- ... sind die Temperaturen in der Stadt höher, oder doch eher gleich?

# Klassisches statistisches Testen

- ▶ Aufstellen einer Hypothese
  - ▶ z.B. „Die mittlere Grundwassertemperatur im Hardtwald beträgt  $11^{\circ}\text{C}$ “.
- ▶ Prüfen der Hypothese
  - ▶ Vergleich von dem was man sieht, mit dem was man beobachten würde, wenn die Hypothese stimmt.
  - ▶ Je besser die Beobachtung zur Hypothese passt, desto eher wird man ihr vertrauen
- ▶ Eine Hypothese kann nicht endgültig bestätigt oder widerlegt werden!
- ▶ ... wir können uns aber dafür entscheiden sie anzunehmen oder abzulehnen

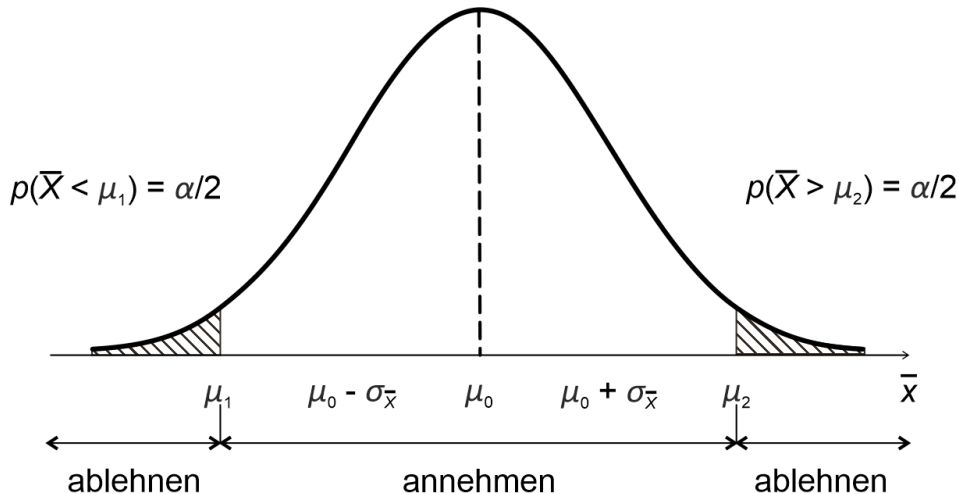
# Ablauf eines statistisches Tests

- ▶ **Nullhypothese  $H_0$ :** „Die mittlere Temperatur beträgt  $11^\circ\text{C}$ .“
- ▶ Hypothetischer Wert  $\mu_0 = 11$
- ▶ Alternative Hypothese: „Die mittlere Temperatur beträgt **nicht**  $11^\circ\text{C}$ .“



# Ablauf eines statistisches Tests

- ▶ Definition eines Annahme, bzw. Ablehnungsbereichs
- ▶ Bedingte Wahrscheinlichkeit, dass wir  $H_0$  ablehnen, obwohl  $H_0$  stimmt
- ▶ Signifikanzniveau  $\alpha$  (oftmals 0.05 oder 0.01)



# Vier mögliche Ergebnisse

	$H_0$ ist richtig	$H_0$ ist falsch
$H_0$ wird angenommen	richtig entschieden	$\beta$ -Fehler
$H_0$ wird abgelehnt	$\alpha$ -Fehler	richtig entschieden

Tschirk (2014)

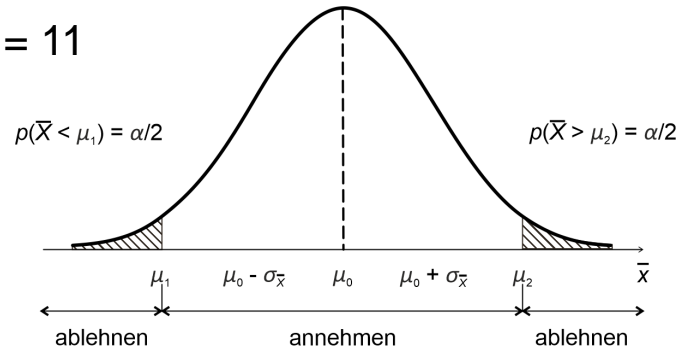
## ► 2 mögliche Fehler:

- Ablehnen einer richtigen Nullhypothese:  $\alpha$ -Fehler oder Fehler 1. Art
- Annehmen einer falschen Nullhypothese:  $\beta$ -Fehler oder Fehler 2. Art

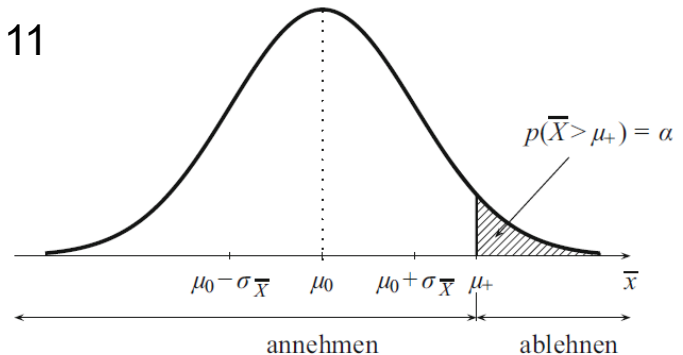


# Ein- und zweiseitige Tests

- Zweiseitig: Nullhypothese  $\mu_0 = 11$



- Einseitig: Nullhypothese  $\mu_0 \leq 11$

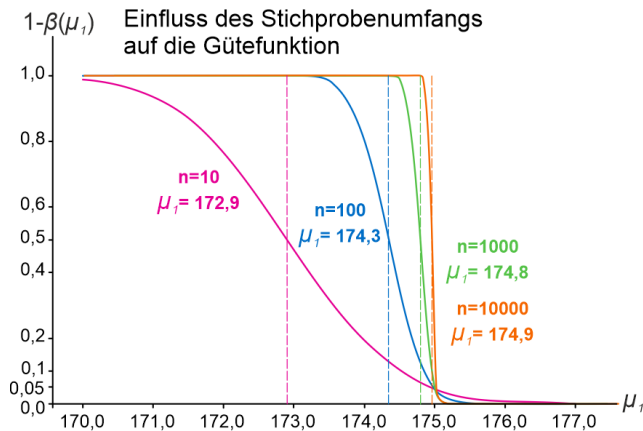
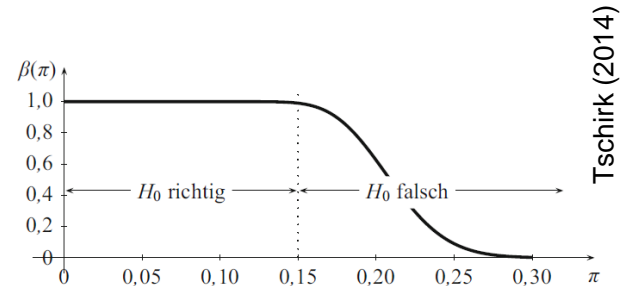


# Interpretation des Testergebnisses

- ▶ Signifikanzwert,  $p$ -Wert
- ▶ Wahrscheinlichkeit den beobachteten Wert zu erhalten, unter der Bedingung dass  $H_0$  stimmt
- ▶ Ablehnen der Nullhypothese, wenn  $p\text{-Wert} \leq \alpha$ 
  - ▶ Annahme der alternativen Hypothese
  - ▶ „statistisch signifikant“ = „überzufällig“
- ▶ Ermöglicht Vergleich verschiedener Testergebnisse
- ▶ Gibt keine Aussage über die Größe des wahren Effekts**
- ▶ Sagt nicht aus wie wahrscheinlich die Nullhypothese ist**

# Trennschärfe eines Tests

- ▶ Funktion  $1 - \beta$ : auch Güte, Stärke, engl. power
- ▶  $\beta$ : Wahrscheinlichkeit, dass  $H_0$  korrekterweise abgelehnt wird



- ▶ Abhängig von Anzahl der Proben

# Übersicht einiger typischer Tests

Student's t-test (one sample)	Mittelwert einer Verteilung entspricht einem bestimmten Wert
Student's t-test (two sample)	Mittelwert zweier Verteilungen sind identisch
F-Test	Vergleicht die Varianz zweier Proben
Mann-Whitney U-Test	Differenz des Median zweier Verteilungen
Shapiro-Wilk Test	Test auf Normalverteilung

► ... viele mehr!

# Parametrische und nicht-parametrische Tests

- ▶ Parametrische Tests setzen eine Normalverteilung der Stichproben voraus → Überprüfen!
- ▶ ggfs. müssen Datensätze normalisiert, bzw. standardisiert werden
- ▶ Parametrische Test:
  - ▶ Student's t-test
  - ▶ F-test
  - ▶ Analysis of Variance (ANOVA)
  - ▶ ...
- ▶ Nicht-Parametrische Test:
  - ▶ Mann-Whitney U-test
  - ▶ ...

# Limitierungen statistischer Tests

- ▶ Eine Hypothese kann nicht endgültig bestätigt oder widerlegt werden
- ▶ Prüfung der Übereinstimmung von Stichprobe und Hypothese
- ▶ Der Test bevorzugt die Nullhypothese (kleiner  $p$ -Werte)
- ▶ Die Nullhypothese muss von der Stichprobe unabhängig sein

## ... was ein Test nicht kann

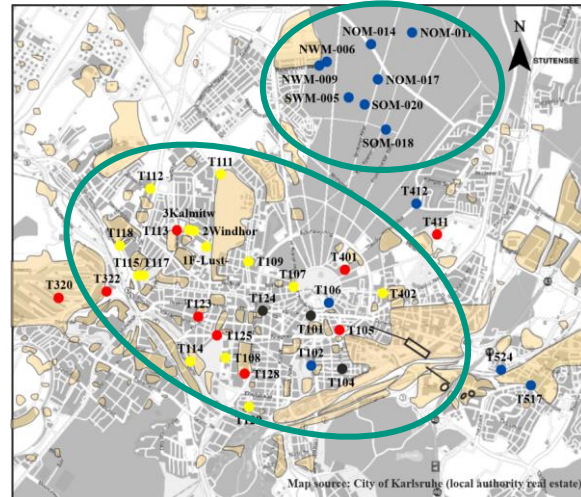
12 Missverständnisse zu  $p$ -Werten (Goodman, 2008):

- ▶ 1. mit  $p = 0.05$  hat die Nullhypothese eine Chance von 5% wahr zu sein.
- ▶ 2. ein nicht-signifikanten Unterschied ( $p > 0.05$ ) bedeutet, dass kein Unterschied zwischen den Gruppen besteht.
- ▶ 3. ein statistisch signifikantes Ergebnis ist wissenschaftlich bedeutsam.
- ▶ 7.  $p = 0.05$  und  $p \leq 0.05$  bedeuten das Gleiche.
- ▶ ...

# Übung 2.2: Statistisches Testen

## ► Grundwasserdatensatz Karlsruhe

- Hypothesen testen
- Verschiedene Tests
- p-Werte bestimmen



- Aufgaben in Jupyter Notebook:  
geodatenanalyse\_1-2-2



# Aufgabenbesprechung

## ► Hypothese 1:

- Temperatur im Wald normalverteilt  $\rightarrow H_0$  annehmen
- Temperatur = 11°C?  $\rightarrow H_0$  nicht annehmen ( $p = 0.005$ )
- Mittelwert Temperatur = 10.7°C,  $n = 8 \rightarrow$  Trennschärfe!

## ► Hypothese 2:

- Sauerstoffsättigung normalverteilt  $\rightarrow H_0$  annehmen
- $F = 1.79$ ,  $p = 0.12 \rightarrow H_0$  annehmen
- $T = 3.46$ ,  $p = 0.0007$ , und  $T > T_{\text{kritisch}} \rightarrow H_0$  ablehnen

## ► Mann-Whitney U-test:

- z.B. Phosphat, nicht normalverteilt, gleiche Verteilung in Wald und Stadt

# Literatur

- ▶ Tschirk (2014) Statistik: Klassisch oder Bayes, Springer
- ▶ Steve Goodman (2008) A Dirty Dozen: Twelve P-Value Misconceptions

## Nützliche Links:

- ▶ <https://machinelearningmastery.com/statistical-hypothesis-tests-in-python-cheat-sheet/>

