

# Modulpfad Statistik

**Der folgende Pfad strukturiert das Modul und dient Ihnen als Leitfaden zur besseren Orientierung.**

Jochen Kruppa-Scheetz

Invalid Date

## Gültigkeit

Dieses Dokument ist für das **Wintersemester 2025/26** und das darauffolgende **Sommersemester 2026** gültig.

Alle notwendigen zusätzlichen Informationen zum Modul finden Sie auf ILIAS, einschließlich:

- Mitschriften der Vorlesung und der Übungen
- R-Skripte und die in der Vorlesung verwendeten Dateien
- Die gesammelten Klausurfragen für Bio Data Science

Beachten Sie bitte auch die zusätzlichen Informationen zu einzelnen Terminen auf ILIAS.

## Skript als Lernhilfe

Das Openbook *Bio Data Science* dient als Skript der Veranstaltung und ist online unter <https://jkruppa.github.io/> verfügbar. Dort finden Sie ergänzend weiterführende Informationen zu R, Excel und der Datenauswertung. *Für die Klausur sind jedoch primär die Inhalte der Veranstaltungsmitschriften relevant.* Beachten Sie die Veranstaltungsmitschriften daher besonders.

## Videos als Lernhilfe

Ergänzend zum Openbook *Bio Data Science* und den Veranstaltungsmitschriften stehen Ihnen unterstützende Lernvideos zur Veranstaltung auf YouTube unter <https://www.youtube.com/c/JochenKruppa> zur Verfügung. Dort finden Sie thematisch sortierte Playlists zu R- und Excel-Analysen sowie allgemeine Statistik-Videos.

## **Programmieren in R**

Zum Verständnis der Programmierung in R stehen Ihnen das Openbook *Bio Data Science* und YouTube-Videos unter <https://www.youtube.com/c/JochenKruppa> zur Verfügung.

### Aktive Teilnahme

- Programmieren in R ist integraler Bestandteil der Veranstaltung.
- Die optionale Hausarbeit erfordert die selbstständige Anwendung von R.
- Theoretische R-Kenntnisse sind klausurrelevant.

## **Inhalt der Vorlesung**

Nachfolgend finden Sie eine Übersicht der Modulinhalte. Die Weblinks führen Sie zu den entsprechenden YouTube-Videos. Alle Inhalte sind verbindlich für die Klausur, auch wenn einzelne Vorlesungen krankheitsbedingt ausfallen sollten. Eine umfassende Übersicht über die Modulinhalte bietet Ihnen die [YouTube-Playlist Openbook Bio Data Sciene](#). Diese Videos sind als Ergänzung und zur Wiederholung konzipiert. Insbesondere folgende Inhalte aus dem Openbook *Bio Data Science* sind prüfungsrelevant:

- 1) [Programmieren in R](#)
- 2) [Deskriptive Statistik](#)
- 3) [Visualisierung von Daten](#)
- 4) [Die Testentscheidung](#)
- 5) [Die Testtheorie](#)
- 6) [Der t-Test](#)
- 7) [Die ANOVA](#)
- 8) [Der Chi-Quadrat-Test](#)

## **Prüfungsleistung**

Die Prüfungsleistung umfasst eine schriftliche Klausur und optional eine Hausarbeit. Für beide werden Punkte vergeben. Zum Bestehen ist das Gesamtergebnis entscheidend, nicht das Ergebnis einzelner Teilleistungen. Die Teilnahme an der schriftlichen Klausur ist verbindlich. Die Abgabe der Hausarbeit ist freiwillig.

Die Prüfungsphase ist im Wintersemester im Januar und im Sommersemester im Juni. Den Prüfungsplan für das laufende Semester finden Sie im Intranet der Hochschule Osnabrück unter <https://intranet.hs-osnabrueck.de/infothek/fakultaet-aul/studium/pruefungen/>.

## Schriftliche Klausur

Auf ILIAS finden Sie eine PDF-Datei mit allen potenziellen Klausurfragen des Moduls. Zu fast allen Aufgaben sind erklärende YouTube-Videos verlinkt.

**Die schriftliche Klausur gliedert sich in:**

- *Multiple-Choice-Teil:* 10 Aufgaben (5 Antwortmöglichkeiten pro Frage, davon 1 richtig). Falsche Antworten führen zu keinem Punktabzug.
- *Text- und Rechenteil:* 6 Aufgaben.

**Erlaubte Hilfsmittel:**

- Taschenrechner (nicht kommunikationsfähig, kein Handy!).
- Eine beidseitig handschriftlich beschriebene DIN A4-Formelsammlung (keine Ausdrucke).

**Bearbeitungszeit:** 120 Minuten (2 Stunden).

## Inhalt der Klausur

Die Klausurinhalte werden jedes Wintersemester aktualisiert. Aus diesem Grund beziehen sich die hier bereitgestellten Aufgaben ausschließlich auf die Hauptklausur des Wintersemesters und die Nachklausur des darauffolgenden Sommersemesters. Für Wiederholungsprüfungen in folgenden Jahren ist es entscheidend, den jeweils aktuellen Stoff zu beachten.

*Es werden sechs der folgenden Statistikaufgaben in der Klausur drankommen.*

- 1) Grundlegende Kenntnisse der Programmierung in R
- 2) Zerforschen des Barplots
- 3) Visualisierung des Barplots
- 4) Zerforschen des Boxplots
- 5) Visualisierung des Boxplots
- 6) Visualisierung des Scatterplots
- 7) Visualisierung des Mosaicplots
- 8) Grundgesamtheit und experimentelle Stichprobe
- 9) Das Nullritual - Die statistische Testtheorie
- 10) Visualisierung der Teststatistik und dem p-Wert
- 11) Berechnung des Student t-Test oder Welch t-Test
- 12) Berechnung des Student t-Test
- 13) Interpretation des t-Tests in R: Teststatistik und p-Wert
- 14) Interpretation des t-Tests in R: die Visualisierung
- 15) Visualisierung der einfaktoriellen ANOVA
- 16) Ergebnistabelle der einfaktoriellen ANOVA

- 17) Die einfaktorielle ANOVA in R
- 18) Den Chi-Quadrat-Test berechnen
- 19) Den Chi-Quadrat-Test mit Effektmaß berechnen
- 20) Den Chi-Quadrat-Test in einem Fragebogen berechnen

## Hausarbeit

Alle relevanten Informationen zur Hausarbeit für das Modul “*Statistik*” sind nachfolgend aufgeführt. Die Abgabe ist im Laufe des Semesters jederzeit möglich.

Die Abgabe der Hausarbeit ist für das Bestehen des Moduls “*Statistik*” nicht erforderlich.

### Abgabe der Hausarbeit als PDF-Datei

Für die Abgabe gehen Sie bitte wie folgt vor:

- 1) Laden Sie Ihre PDF-Datei in den ILIAS-Abgabeordner hoch.
- 2) Benennen Sie die Datei mit Ihrem Namen und Ihrer Matrikelnummer.
- 3) Die Abgabe muss bis zum **21. Dezember 2025** erfolgen.

## Datenanalyse in R

Die Datenanalyse ist eine Gruppenarbeit. Die Gruppenarbeit beschränkt sich dabei auf die Erstellung der Daten und die Analyse der Hausarbeit. Die finale Abgabe des Textes muss von jedem Einzelnen persönlich erfolgen.

### Formalia

Für Ihre Hausarbeit gelten folgende Anforderungen:

- 1) Die Hausarbeit ist im IMRaD-Schema zu verfassen.
- 2) Der Umfang beträgt 8 bis maximal 10 Seiten reinen Fließtextes. Aufzählungen sind nicht zulässig.
- 3) Die Datenanalyse ist mit der Software RStudio und R durchzuführen.

Im Modul “*Statistik*” arbeiten Sie in Gruppen an einem Projekt, bei dem Sie einen Fragebogen entwickeln. Dieser wird von den anderen Teilnehmenden des Moduls ausgefüllt. Die daraus gewonnenen Daten verwenden Sie als Grundlage für Ihre Hausarbeit. Ihre wissenschaftliche Fragestellung erarbeiten Sie eigenständig in Ihrer Gruppe.

## Vorgaben für die Fragen

Achten Sie beim Erstellen Ihres Fragebogens darauf, dass die folgenden Vorgaben erfüllt sind.

- 1) **Erste Frage: Dichotome Frage** Mit dieser Frage unterteilen Sie die Teilnehmenden in zwei Gruppen. Hier ist nur eine Ja/Nein-Antwort möglich.

*Beispiel: "Haben Sie bereits Erfahrung mit statistischer Datenauswertung?"*

2. **Zweite Frage: Kontinuierliche Frage** Diese Frage sollte eine kontinuierliche Antwort ermöglichen, also eine Antwort, die jeden beliebigen Wert in einem bestimmten Bereich annehmen kann.

*Beispiel: "Wie viele Stunden pro Woche verbringen Sie durchschnittlich mit der Vorbereitung auf das Statistik-Modul?"*

3. **Dritte Frage: Kategoriale Frage mit drei Optionen** Diese Frage hat eine kategoriale Antwort mit drei festen Auswahlmöglichkeiten.

*Beispiel: "Welche Vorkenntnisse in Statistik bringen Sie mit?"*

- ( ) Keine
- ( ) Geringe
- ( ) Umfangreiche

4. **Vierte Frage: Kategoriale Frage mit fünf Optionen** Diese Frage hat eine kategoriale Antwort mit fünf festen Auswahlmöglichkeiten.

*Beispiel: "Wie bewerten Sie die bisherige Qualität der Vorlesung?"*

- ( ) Sehr schlecht
- ( ) Schlecht
- ( ) Mittelmäßig
- ( ) Gut
- ( ) Sehr gut

5. **Fünfte Frage: Kategoriale Frage mit zwei Optionen** Diese Frage hat eine kategoriale Antwort mit zwei festen Auswahlmöglichkeiten.

*Beispiel: "Nehmen Sie regelmäßig an den Tutorien teil?"*

- ( ) Ja
- ( ) Nein

Ihre Hausarbeit wird nach folgenden Kriterien bewertet:

- **Einleitung (1 Punkt):** Entwicklung einer wissenschaftlichen Fragestellung auf Basis der ausgewählten Daten.
- **Materialien und Methoden (2 Punkte):**
  - Beschreibung des Fragebogens anhand der gegebenen Informationen.
  - Beschreibung der angewandten statistischen Methoden zur Datenauswertung.
- **Ergebnisse (6 Punkte):**
  - Visualisierung der Daten im Kontext der Fragestellung (**2 Punkte**).
  - Bericht der Ergebnisse und Erkenntnisse aus der statistischen Analyse (**4 Punkte**).
- **Diskussion (1 Punkt):** Bewertung der Ergebnisse im Kontext der Fragestellung.

### **Benotung**

Ihre Leistung im Bericht und der Datenauswertung wird mit bis zu 10 Punkten bewertet. Diese Punkte fließen direkt in Ihre Gesamtnote ein, indem sie zu Ihrer Klausurpunktzahl addiert werden. Die erreichte Summe bestimmt dann Ihre endgültige Modulnote.