

Stochastisches Denken im KI-Zeitalter: KI als Medium für Verstehen, Intuition und didaktische Transformation

Manfred Borovcnik, Universität Klagenfurt

- A. Stochastisches Denken und stochastische Literalität mit KI**
- B. Statistische Analyse mit KI als Werkzeug**
- C. KI in der Datenanalyse – Konsequenzen für Lehre und Praxis**

KI ist umso effizienter, je besser man das Fach beherrscht.
KI kann allerdings – mit erheblicher Variabilität zwar –
auch dem Laien vieles „erklären“.

A. Stochastisches Denken und stochastische Literalität mit KI

B. Statistische Analyse mit KI als Werkzeug

C. KI in der Datenanalyse – Konsequenzen für Lehre und Praxis

1. Stochastisches Denken und Literalität
2. Veränderungen in der Unterweisung durch KI
3. Didaktisch kontrollierter KI-Einsatz, der Denkformen & Literalitäten fördert
4. Statistisches Denken mit KI fördern
5. KI selbst ist nicht die interaktive Umgebung
6. Fragen nach der Gestaltung in GeoGebra
7. Conclusio

KI ist umso effizienter, je besser man das Fach beherrscht.

KI kann allerdings – mit erheblicher Variabilität zwar – auch dem Laien vieles „erklären“.

A. Stochastisches Denken und stochastische Literalität mit KI

B. Statistische Analyse mit KI als Werkzeug

C. KI in der Datenanalyse – Konsequenzen für Lehre und Praxis

1. Orientierung: Analyseleitfäden durch KI
2. Erster Datensatz: Einfach strukturierte ANOVA-Situation
3. Zweiter Datensatz: ANOVA mit zusätzlicher Verletzung der Voraussetzungen
4. Stabilitätsprüfungen
5. Weitere Datensätze (t-Test und Regression) – nicht im Artikel abgedeckt
6. Fazit aus der Analyse der Datensätze

KI ist umso effizienter, je besser man das Fach beherrscht.

KI kann allerdings – mit erheblicher Variabilität zwar – auch dem Laien vieles „erklären“.

A. Stochastisches Denken und stochastische Literalität mit KI

B. Statistische Analyse mit KI als Werkzeug

C. KI in der Datenanalyse – Konsequenzen für Lehre und Praxis

1. Generative KI als Ermöglicher statistischer Datenanalyse
2. Standardisierung statistischer Analyse als Voraussetzung für Automatisierung
3. Leistungsfähigkeit der KI: Zwischen Effizienz und Oberflächlichkeit
4. Inkonsistenz und mangelnde Zuverlässigkeit
5. Grenzen generativer KI: Kein konzeptionelles Verständnis
6. Demokratisierung vs. Qualitätsverlust
7. Konsequenzen für die Statistikausbildung
8. Veränderung der Rolle von Software und Programmierung
9. Fokus auf Datenverständnis und kritisches Denken
10. Schlussfolgerungen

KI ist umso effizienter, je besser man das Fach beherrscht.
KI kann allerdings – mit erheblicher Variabilität zwar –
auch dem Laien vieles „erklären“.

A. Stochastisches Denken und stochastische Literalität mit KI

B. Statistische Analyse mit KI als Werkzeug

C. KI in der Datenanalyse – Konsequenzen für Lehre und Praxis

1. Stochastisches Denken und Literalität
2. Veränderungen in der Unterweisung durch KI
3. Didaktisch kontrollierter KI-Einsatz, der Denkformen & Literalitäten fördert
4. Statistisches Denken mit KI fördern
5. KI selbst ist nicht die interaktive Umgebung
6. Fragen nach der Gestaltung in GeoGebra
7. Conclusio

KI ist umso effizienter, je besser man das Fach beherrscht.

KI kann allerdings – mit erheblicher Variabilität zwar – auch dem Laien vieles „erklären“.

1. Denken und Literalität

Statistisches Denken – Schlüsselmerkmale

Datenorientierung Statistisches Denken basiert auf realen Daten. Diese werden gesammelt, aufbereitet & analysiert, um Muster oder Zusammenhänge zu erkennen. Dabei steht die Beobachtung im Mittelpunkt. Ohne Daten keine statistische Analyse.

Variabilität als zentrales Konzept Kernaspekt ist das Verständnis, dass Daten streuen und nie exakt gleich sind. Statistisches Denken beschäftigt sich mit der Beschreibung & Erklärung dieser Variabilität. Ursachen können Zufall oder systematische Effekte sein. Streuung ist kein Fehler, sondern ein wesentliches Erkenntniselement.

Stichprobe vs. Grundgesamtheit Statistisches Denken unterscheidet klar zwischen Stichprobe und Population. Schlussfolgerungen über die Grundgesamtheit werden aus Stichprobendaten gezogen. Dabei spielt die Repräsentativität eine entscheidende Rolle. Fehlerquellen wie Verzerrungen müssen berücksichtigt werden.

Schätzen und Testen Statistisches Denken arbeitet mit Schätzungen & Hypothesentests. Ergebnisse sind immer mit Unsicherheit behaftet. Konfidenzintervalle & Signifikanzniveaus „quantifizieren“ diese Unsicherheit mittels „Wahrscheinlichkeiten“.

Kontextabhängigkeit Daten sind immer im Kontext zu interpretieren. Statistisches Denken berücksichtigt den Entstehungskontext der Daten (z. B. Messmethoden, Studiendesign). Ohne Fachwissen sind Fehlinterpretationen leicht möglich.

Modellierung realer Prozesse Statistische Modelle dienen dazu, reale Zusammenhänge zu beschreiben. Diese sind vereinfachte Darstellungen der Wirklichkeit. Ihre Gültigkeit hängt von Annahmen ab. Modellkritik ist daher wesentlich.

Probabilistisches Denken – Schlüsselmerkmale

Umgang mit Zufall und Unsicherheit Probabilistisches Denken fokussiert auf die Beschreibung von Zufallsphänomenen. Unsicherheit wird quantitativ durch Wahrscheinlichkeiten erfasst. Dabei wird nicht beobachtet, sondern modelliert. Zufall wird als strukturiertes Konzept verstanden.

Wahrscheinlichkeitsmodelle Im Zentrum stehen Modelle wie Zufallsvariablen & Verteilungen. Diese beschreiben **mögliche** Ergebnisse & deren Eintrittswahrscheinlichkeiten. Sie sind „unabhängig“ von konkreten Daten.

A-priori-Betrachtung Probabilistisches Denken betrachtet Situationen häufig vor der Datenerhebung. Wahrscheinlichkeiten werden theoretisch bestimmt (z. B. bei idealen Zufallsexperimenten). Es geht um Erwartungen über zukünftige Ereignisse. Beobachtungen sind nicht zwingend erforderlich.

Regelbasierte Struktur nach festen mathematischen Regeln. Diese ermöglichen exakte Berechnungen. Schlussfolgerungen folgen logisch aus den Annahmen. Die Struktur ist formal & axiomatisch.

Gedankenexperimente Viele probabilistische Überlegungen erfolgen in idealisierten Szenarien. Beispiele sind Münzwürfe oder Würfelexperimente. Diese dienen dem Verständnis von Zufallsgesetzen. Reale Komplexität wird oft ausgeblendet.

Langfristige Häufigkeiten Zentral ist die Interpretation von Wahrscheinlichkeit als relative Häufigkeit. Wiederholte Experimente führen zu stabilen Schätzungen. Diese Perspektive verbindet Theorie & Empirie. Der Fokus liegt auf theoretischen Erwartungen.

Unterschiede zwischen statistischem und probabilistischem Denken

Empirie vs. Theorie Während statistisches Denken von empirischen Beobachtungen ausgeht, ist probabilistisches Denken modellbasiert und beginnt oft mit Annahmen. Die Ausgangspunkte unterscheiden sich grundlegend.

Richtung der Schlussfolgerung Statistik arbeitet induktiv: von Stichproben zu allgemeinen Aussagen. Probabilistik ist deduktiv: von Modellen zu konkreten Wahrscheinlichkeiten. Das bedeutet, Statistik generalisiert aus Daten, während Wahrscheinlichkeit aus Regeln ableitet. Diese Denklogiken sind gegensätzlich.

Rolle von Daten Im statistischen Denken sind Daten zwingend notwendig. Ohne Daten gibt es keine Analyse oder Interpretation. Im probabilistischen Denken können viele Aussagen ohne Daten getroffen werden. Modelle funktionieren auch rein theoretisch.

Umgang mit Unsicherheit Statistik quantifiziert Unsicherheit aufgrund begrenzter Daten. Probabilistik beschreibt sie als inhärente Eigenschaft eines „Experiments“. Die Ursachen der Unsicherheit werden unterschiedlich interpretiert → Methoden verschieden.

Zielsetzung Statistisches Denken zielt darauf ab, reale Phänomene zu verstehen & Entscheidungen zu treffen. Probabilistisches Denken, Zufallsprozesse formal zu beschreiben. Statistik ist stärker anwendungsorientiert, stärker theoretisch ausgerichtet.

Modellbezug Statistische Modelle werden an Daten angepasst. Probabilistische Modelle werden oft vorab festgelegt und analysiert. In der Statistik können Modelle verworfen werden, wenn sie nicht passen. In der Wahrscheinlichkeit gelten Modelle als gegeben.

Ist eine Unterscheidung Denken vs. Literalität sinnvoll?

Kognitive Prozesse & beobachtbaren Kompetenzen „Denken“ bezeichnet kognitive Prozesse (Modellieren von Zufall, Ziehen von Schlussfolgerungen etc.). „Literalität“ hingegen bezieht sich auf manifeste Fähigkeiten (Verstehen von Darstellungen, Kommunizieren etc.). Die Unterscheidung zwischen innerem Verständnis und äußerer Performanz zu differenzieren ist insbesondere in der Diagnostik bedeutsam.

Didaktische Planbarkeit Lernziele werden präziser, wenn zwischen Denkprozessen und Literalität unterschieden wird. Denkformen betreffen eher konzeptuelle Einsichten, Literalität eher anwendungsnahe Kompetenzen. Dadurch können Curricula strukturierter aufgebaut werden. Ohne diese Differenzierung besteht die Gefahr unscharfer Lernziele.

Messbarkeit und Evaluation Literalität ist in der Regel leichter operationalisierbar und empirisch erfassbar als Denken. Tests, Aufgabenformate & Assessments zielen auf sichtbare Leistungen. Die Unterscheidung verhindert, dass gemessene Performanz mit tiefem Verständnis gleichgesetzt wird. Dies erhöht die Validität von Bildungsforschung.

Gefahr künstlicher Trennung Literalität ohne zugrunde liegendes Denken bleibt oberflächlich. Umgekehrt kann Denken ohne Ausdrucksfähigkeit schwer sichtbar werden. Eine zu starke Trennung kann daher theoretisch problematisch sein.

Kontextabhängigkeit der Literalität Literalität ist an konkrete Kontexte gebunden. Denken ist stärker konzeptuell & transferierbar. Die Unterscheidung trennt kontextspezifische Fähigkeiten von allgemeinerem Verständnis. → Transferfragen.

Charakterisierung von Statistischer Literalität – zentrale Merkmale

Verstehen statistischer Darstellungen Statistische Literalität umfasst die Fähigkeit, Tabellen, Diagramme & Kennzahlen korrekt zu interpretieren. Auch das Erkennen von Verzerrungen oder irreführenden Darstellungen. Die Interpretation erfolgt im Kontext der Daten. Visuelle & numerische Informationen werden integriert.

Kontextsensibilität Daten werden nicht isoliert betrachtet, sondern im Hinblick auf ihre Entstehung und Bedeutung eingeordnet. Dies schließt Kenntnisse über Datenerhebung und möglichen Bias ein. Statistische Aussagen werden kritisch hinterfragt. Kontext ist Voraussetzung für valide Interpretation.

Grundverständnis zentraler Konzepte Begriffe wie Mittelwert, Streuung oder Korrelation werden in ihrer Bedeutung verstanden. Dabei geht es mehr um inhaltliche Deutung als um formale Berechnung. Fehlinterpretationen, etwa von Korrelation als Kausalität, sollen vermieden werden. Konzeptuelles Verständnis steht im Vordergrund.

Bewertung von Informationen Beinhaltet die Fähigkeit, Datenargumente in Medien oder Wissenschaft zu beurteilen. Erkennen von Unsicherheit und methodischen Schwächen. Aussagen werden nicht unreflektiert übernommen. Skepsis ist ein integraler Bestandteil.

Kommunikationsfähigkeit Statistische Informationen können verständlich dargestellt und erklärt werden. Dies betrifft sowohl mündliche als auch schriftliche Kommunikation. Ziel ist eine adressatengerechte Vermittlung. Fachsprache wird angemessen eingesetzt.

Alltagsbezug Eng mit realen Anwendungen verbunden. Statistische Literalität ermöglicht informierte Entscheidungen. Der Fokus liegt auf praktischer Relevanz. Abstrakte Modelle treten in den Hintergrund.

Probabilistische Literalität – zentrale Merkmale

Interpretation von Wahrscheinlichkeiten Umfasst das Verständnis von Wahrscheinlichkeitsaussagen in verschiedenen Darstellungsformen (Prozent, Brüche). Schließt verbale Unsicherheitsaussagen mit ein. Die Bedeutung von Wahrscheinlichkeit wird situationsgerecht interpretiert. Missverständnisse werden vermieden.

Umgang mit Unsicherheit Unsicherheit wird als grundlegendes Merkmal vieler Situationen erkannt. Probabilistische Literalität ermöglicht es, Risiken und Chancen abzuwägen. Entscheidungen werden unter Unsicherheit reflektiert getroffen.

Verständnis einfacher Modelle Grundlegende probabilistische Modelle wie Zufallsexperimente oder Verteilungen werden verstanden. Dabei steht die Idee möglicher Ergebnisse und deren Wahrscheinlichkeiten im Zentrum. Formale Herleitungen sind zweitrangig. Modelle dienen als Denkwerkzeuge.

Alltagsrelevanz von Wahrscheinlichkeit Probabilistische Literalität zeigt sich in der Anwendung auf reale Situationen, etwa Wetterprognosen oder medizinische Risiken. Wahrscheinlichkeiten werden nicht nur berechnet, sondern interpretiert. Entscheidungen werden daran ausgerichtet. Der Bezug zur Lebenswelt ist zentral.

Erkennen typischer Fehlvorstellungen Fehlkonzepte wie der Spielerfehlschluss werden erkannt. Probabilistische Literalität beinhaltet ein reflektiertes Verständnis von Zufall. Intuitive Fehltritte werden hinterfragt. Dies erhöht die Qualität von Entscheidungen.

Kommunikation probabilistischer Aussagen Wahrscheinlichkeiten können verständlich erklärt & dargestellt werden. Umfasst auch die Wahl geeigneter Darstellungsformen. Unsicherheit wird transparent kommuniziert. Klare und korrekte Vermittlung.

Wesentliche Unterschiede Denken vs. Literalität

Denken vs. Literalität Denkformen beziehen sich auf interne kognitive Strukturen und Prozesse. Literalität beschreibt beobachtbare Fähigkeiten im Umgang mit Inhalten. Während Denken die Grundlage bildet, ist Literalität die manifeste Ausprägung. Beide Ebenen sind analytisch trennbar, aber praktisch eng verbunden.

Abstraktionsgrad Probabilistisches und statistisches Denken sind stärker theoretisch und konzeptuell orientiert. Literalitäten sind stärker an konkrete Anwendungssituationen gebunden. Sie operationalisieren Denkprozesse in realen Kontexten. Dadurch sind sie oft leichter zugänglich, aber auch kontextabhängiger.

Anwendung vs. Struktur Literalität betont die Nutzung von Wissen zur Interpretation & Kommunikation. Denken fokussiert stärker auf Strukturverständnis & Modellbildung. Beispielsweise kann jemand statistisch denken, ohne komplexe Darstellungen sicher zu kommunizieren. Umgekehrt ist oberflächliche Literalität ohne tiefes Denken möglich.

Statistisch vs. probabilistisch Statistische Literalität ist datengetrieben & rückblickend (Analyse vorhandener Daten). Probabilistische Literalität ist auf Unsicherheit und zukünftige Ereignisse ausgerichtet. Erstere interpretiert empirische Evidenz, letztere bewertet mögliche Entwicklungen. Ergänzen sich, sind aber konzeptionell verschieden.

KI im Unterricht von Wahrscheinlichkeitsrechnung & Statistik kann sowohl (statistische & probabilistische) *Denkformen* als auch (statistische & probabilistische) *Literalitäten* substanziell beeinflussen. Entscheidend ist, ob KI als kognitives Werkzeug zur Vertiefung oder als bloßes Antwortsystem genutzt wird. Im Folgenden wird dies systematisch strukturiert.

2. Veränderungen in der Unterweisung durch KI

Traditioneller Unterricht (ohne KI)

- Lineare Vermittlung von Inhalten (Lehrperson → Lernende)
- Fokus auf Verfahren (Rechnen, Formeln, Standardaufgaben)
- Begrenzte Datenbasis (Lehrbuchbeispiele, kleine Datensätze)
- Verzögerte Rückmeldung (Korrektur durch Lehrperson)
- Geringe Individualisierung

Auswirkungen auf Denkformen:

- Statistisches Denken bleibt oft oberflächlich (Deskription statt Interpretation)
- Probabilistisches Denken reduziert sich häufig auf Rechenregeln
- Fehlvorstellungen bleiben teilweise unerkannt

Auswirkungen auf Literalitäten:

- Interpretation realer Daten ist eingeschränkt
- Kommunikation statistischer Aussagen wird wenig geübt
- Transfer auf Alltagssituationen bleibt begrenzt

KI-unterstützter Unterricht

- Interaktive, dialogische Lernumgebung
- Dynamischer Zugriff auf reale oder simulierte Daten
- Sofortige, adaptive Rückmeldungen
- Hohe Individualisierung (Tempo, Niveau, Darstellungsform)
- Unterstützung bei Modellbildung und Interpretation

Von Rechnen & Reproduzieren hin zu Verstehen, Interpretieren & Reflektieren.

Einfluss von KI auf Statistisches Denken

- KI kann große, realitätsnahe Datensätze bereitstellen und analysieren
- Lernende können Hypothesen formulieren und explorativ prüfen
- Visualisierungen werden dynamisch erzeugt und variiert
- Kontextualisierung wird erleichtert (z. B. durch Fallbeispiele)

Der Fokus verschiebt sich von der Berechnung statistischer Kennzahlen hin zur Interpretation von Daten und Modellen. KI kann gezielt Fragen stellen („Was bedeutet diese Streuung im Kontext?“) und so Reflexion anregen.

Risiko Automatisierte Auswertungen können dazu führen, dass Lernende Ergebnisse übernehmen, ohne deren Entstehung zu verstehen.

Einfluss von KI auf Probabilistisches Denken

- Simulationen (z. B. Zufallsexperimente) beliebig oft & variabel durchführen
- KI kann alternative Szenarien generieren und vergleichen
- Komplexe Wahrscheinlichkeitsprobleme werden schrittweise erklärt
- Intuitive Fehlvorstellungen können gezielt adressiert werden

Vom statischen Aufgabenformat hin zu dynamischem Experimentieren mit Zufall
Lernende erleben Wahrscheinlichkeit als langfristiges Muster und nicht nur als Formel.

Risiko Wenn KI Lösungen sofort liefert, kann die eigenständige Modellbildung unterdrückt werden.

Einfluss von KI auf Statistische Literalität

- Analyse aktueller Medienberichte mit realen Daten
- Automatisierte Generierung unterschiedlicher Darstellungen (Diagrammtypen etc.)
- Unterstützung beim Formulieren und Kommunizieren von Ergebnissen
- Vergleich unterschiedlicher Interpretationen

Stärkere Ausrichtung auf kritische Bewertung von Daten in realen Kontexten

KI kann als Diskussionspartner fungieren und Argumente herausfordern.

Risiko KI-generierte Texte können scheinbar überzeugend, aber inhaltlich fehlerhaft sein. Ohne kritische Prüfung entsteht Scheinsicherheit.

Einfluss von KI auf Probabilistische Literalität

- Alltagssituationen (z. B. Risiken, Prognosen) können simuliert und analysiert werden
- KI kann unterschiedliche Darstellungsformen von Wahrscheinlichkeiten anbieten
- Unterstützung bei der Interpretation unsicherer Aussagen
- Reflexion über Entscheidungsprozesse unter Unsicherheit

Vom abstrakten Rechnen hin zur reflektierten Entscheidung unter Unsicherheit

KI kann alternative Interpretationen anbieten und zur Bewertung auffordern.

Risiko Lernende könnten probabilistische Aussagen der KI ungeprüft übernehmen, insbesondere wenn diese plausibel formuliert sind.

Chancen eines KI-Zugangs

Individualisierung des Lernprozesses

KI kann Aufgaben und Erklärungen an den Kenntnisstand anpassen. Dies ermöglicht differenziertes Lernen ohne zusätzlichen Aufwand für Lehrkräfte.

Förderung von Exploration und Hypothesenbildung

Durch Simulationen und flexible Datenzugänge können Lernende eigenständig Fragestellungen entwickeln und untersuchen.

Unmittelbares Feedback

Fehlvorstellungen können sofort erkannt und adressiert werden. Dies ist besonders für probabilistisches Denken relevant.

Integration realer Kontexte

Aktuelle Daten und Anwendungen können leicht eingebunden werden. Dies stärkt insbesondere die Literalitäten.

Unterstützung komplexer Denkprozesse

KI kann bei Modellbildung, Interpretation und Argumentation helfen, ohne diese vollständig zu ersetzen.

Förderung metakognitiver Reflexion

Durch dialogische Interaktion kann KI Lernende zur Begründung & Überprüfung ihrer Aussagen anregen.

Gefahren eines KI-Zugangs

Oberflächliches Verständnis durch Automatisierung

Wenn KI Rechen- und Analyseprozesse übernimmt, besteht die Gefahr, dass grundlegende Konzepte nicht verstanden werden.

Autoritätsbias gegenüber KI

Lernende neigen dazu, KI-Antworten als korrekt zu akzeptieren. Dies schwächt kritische Prüfung.

Verlust grundlegender Fertigkeiten

Manuelle Fähigkeiten (z. B. einfache Berechnungen oder graphische Interpretation) könnten vernachlässigt werden.

Intransparenz von Modellen

KI liefert Ergebnisse oft ohne vollständige Nachvollziehbarkeit. Dies steht im Spannungsfeld zu wissenschaftlicher Denkweise.

Fehlerhafte oder verzerrte Antworten

KI kann plausible, aber falsche Aussagen generieren. Ohne ausreichende Literalität bleiben diese unentdeckt.

Ungleichheit im Zugang und in der Nutzungskompetenz

Nicht alle Lernenden profitieren gleichermaßen von KI. Unterschiede in Nutzungskompetenz können bestehende Bildungsungleichheiten verstärken.

Bedingte Wirksamkeit eines KI-Zugangs

KI hat das Potenzial, sowohl Denkformen als auch Literalitäten erheblich zu fördern.

Dies gilt jedoch nur unter der Bedingung, dass sie didaktisch eingebettet wird:

- KI sollte Denkprozesse unterstützen, nicht ersetzen
- Aufgaben sollten auf Begründung, Interpretation und Kritik ausgerichtet sein
- Lehrpersonen behalten eine zentrale Rolle als epistemische Instanz

Ohne diese Einbettung besteht die reale Gefahr,

dass KI vor allem zu einer Beschleunigung oberflächlicher Lernprozesse führt, statt zu deren Vertiefung.

3. Didaktischer Einsatz von KI, der Denkformen & Literalitäten fördert

Für jeden Bereich wird ein Unterrichtsbeispiel, die notwendigen Umsetzungsschritte sowie Voraussetzungen & Optimierungsstrategien dargestellt. Der Fokus liegt auf einem KI-Einsatz, der Denkformen & Literalitäten gezielt fördert.

Statistisches Denken „*Beeinflusst Schlafdauer die Prüfungsleistung?*“

Lernende erhalten (oder generieren mit KI) einen Datensatz zu Schlafdauer und Testergebnissen. Die KI kann Visualisierungen erzeugen & einfache Analysen durchführen.

Unterscheidung Korrelation & Kausalität; Interpretation von Streuung & Zusammenhang.

Notwendige Schritte

1. Formulierung von Hypothesen durch die Lernenden
2. Exploration der Daten (Visualisierung, Kennzahlen)
3. Interpretation der Ergebnisse im Kontext
4. Kritische Reflexion möglicher Störvariablen (z. B. Stress, Vorbereitung)
5. Diskussion: Was lässt sich *nicht* aus den Daten schließen?

Voraussetzungen

- Grundverständnis von Variabilität und Zusammenhang
- Fähigkeit, Diagramme zu lesen
- Bewusstsein für Kontextabhängigkeit

Verbesserung des Impakts

- KI nicht nur zur Berechnung nutzen, sondern gezielt zur Begründung („Warum ist ...?“)
- Vergleich mehrerer Datensätze (echte vs. simulierte)
- Explizite Thematisierung von Fehlinterpretationen

Probabilistisches Denken „Ist eine Serie von 5-mal Kopf ,unwahrscheinlich‘?“

KI simuliert wiederholt Münzwürfe und visualisiert Häufigkeiten von Serien.

Lernende können Parameter variieren (Anzahl der Würfe, Wiederholungen).

Ziel: Verständnis von Zufall, Unabhängigkeit und langfristigen Häufigkeiten.

Notwendige Schritte

1. Intuitive Einschätzung der Wahrscheinlichkeit durch Lernende
2. Simulation mit KI (viele Wiederholungen)
3. Vergleich von Erwartung und Beobachtung
4. Diskussion von Fehlvorstellungen (z. B. „Ausgleichserwartung“)
5. Verallgemeinerung auf andere Zufallsexperimente

Voraussetzungen

- Grundidee von Zufallsexperimenten
- Verständnis relativer Häufigkeit
- Bereitschaft zur Reflexion eigener Intuition

Verbesserung des Impakts

- Vorhersagen **vor** der Simulation einfordern
- KI gezielt widersprüchliche Fälle erzeugen lassen
- Reflexion stärker gewichten als das Ergebnis

Statistische Literalität „*Medienbericht zu Impfwirksamkeit*“

Ein realer oder KI-generierter Artikel enthält statistische Angaben (z. B. Prozentwerte, Risiken).

Lernende analysieren diesen kritisch mit Unterstützung der KI.

Ziel: Kritische Interpretation statistischer Informationen im Kontext.

Notwendige Schritte

1. Identifikation zentraler Aussagen im Text
2. Übersetzung in statistische Begriffe (absolute vs. relative Risiken)
3. Prüfung der Datenbasis (Stichprobe, Repräsentativität)
4. Bewertung der Darstellung (irreführend oder korrekt?)
5. Formulierung einer eigenen, verbesserten Darstellung

Voraussetzungen

- Verständnis grundlegender statistischer Begriffe
- Lesekompetenz für Fachtexte
- Sensibilität für Verzerrungen

Verbesserung des Impakts

- KI gezielt alternative Darstellungen erzeugen lassen
- Fehlerhafte Interpretationen bewusst einbauen und analysieren
- Diskussion in Gruppen strukturieren (Argumente begründen lassen)

Probabilistische Literalität „*Was bedeutet 30 % Wahrscheinlichkeit für Regen?*“

KI generiert verschiedene Interpretationen (korrekt und inkorrekt) einer Wetterprognose. Lernende müssen diese bewerten.

Ziel

Verständnis probabilistischer Aussagen im Alltag; Vermeidung typischer Fehlinterpretationen.

Notwendige Schritte

1. Sammlung spontaner Interpretationen durch Lernende
2. Vergleich mit KI-generierten Aussagen
3. Klassifikation in korrekt vs. inkorrekt
4. Begründung der Entscheidungen
5. Übertragung auf andere Kontexte (z. B. medizinische Risiken)

Voraussetzungen

- Grundverständnis von Wahrscheinlichkeit als Maß für Unsicherheit
- Sprachliche Differenzierungsfähigkeit
- Reflexionsfähigkeit

Verbesserung des Impakts

- KI als „Diskussionspartner“ einsetzen, nicht als Autorität
- Mehrdeutigkeit bewusst thematisieren
- Verknüpfung mit realen Entscheidungen herstellen

Übergreifende Voraussetzungen für erfolgreichen KI-Einsatz

Didaktische Einbettung

KI darf nicht isoliert eingesetzt werden, sondern muss in klar strukturierte Lernprozesse integriert sein.

Aufgabenqualität

Aufgaben müssen auf *Interpretation, Begründung und Kritik* abzielen, nicht auf reine Ergebnisproduktion.

Transparenz der KI-Nutzung

Lernende müssen verstehen, dass KI ein Werkzeug mit Grenzen ist. Ergebnisse sind zu prüfen, nicht zu übernehmen.

Lehrkompetenz im Umgang mit KI

Lehrpersonen müssen in der Lage sein, KI gezielt zu steuern und didaktisch sinnvoll einzusetzen.

Technische Zugänglichkeit

Zuverlässiger Zugang zu KI-Systemen ist Voraussetzung für kontinuierliches Arbeiten.

Förderung metakognitiver Fähigkeiten

Lernende müssen lernen, ihre eigenen Denkprozesse zu reflektieren und zu hinterfragen.

Maximierung des positiven Impakts

- KI als kognitives Werkzeug, nicht als Ersatz
- Primat der Begründung vor dem Ergebnis
- Konfrontation mit Fehlvorstellungen
- Verknüpfung von Modell und Realität

Maximierung des positiven Impakts

- KI als kognitives Werkzeug, nicht als Ersatz
- Primat der Begründung vor dem Ergebnis
- Konfrontation mit Fehlvorstellungen
- Verknüpfung von Modell und Realität

Konkrete Maßnahmen

„Explain-first“-Ansatz:

Lernende müssen ihre Überlegungen formulieren, bevor sie KI nutzen.

Vergleich von Mensch- und KI-Antworten

Unterschiede werden analysiert und bewertet.

Fehleranalyse als Lernziel

KI-generierte Fehler werden systematisch genutzt.

Mehrperspektivität

KI liefert mehrere Interpretationen, die gegeneinander abgewogen werden.

Reflexionsphasen fest verankern

Jede KI-Nutzung wird durch eine Reflexion abgeschlossen („Was habe ich verstanden?“).

Der Mehrwert von KI entsteht nicht automatisch durch ihre Verfügbarkeit,

sondern durch ihre epistemisch kontrollierte Nutzung. Entscheidend ist, dass Lernende nicht nur Antworten erhalten, sondern lernen, Fragen zu stellen, Unsicherheit zu interpretieren und Ergebnisse kritisch zu bewerten.

Nur unter diesen Bedingungen kann KI

sowohl Denkformen als auch Literalitäten nachhaltig fördern.

4. Statistisches Denken mit KI fördern

Im Folgenden wird für die sechs zuvor genannten Merkmale des *statistischen Denkens* konkretisiert, wie KI diese *interaktiv und graphisch* unterstützen kann. Der Fokus liegt auf funktionalen didaktischen Möglichkeiten, nicht auf spezifischen Tools.

Datenorientierung

Verstehen, dass statistische Aussagen auf Daten basieren und durch diese begrenzt sind.

KI-Unterstützung (interaktiv/graphisch):

- Dynamische Datengenerierung: KI erzeugt Datensätze mit variierbaren Parametern (z. B. Stichprobengröße, Verteilung).
- Interaktive Tabellen und Visualisierungen (Histogramme, Boxplots), die sich bei Datenänderung sofort anpassen.
- „Was-wäre-wenn“-Szenarien: Lernende verändern Datenpunkte und beobachten unmittelbar die Auswirkungen auf Kennzahlen und Graphiken.

Die Abhängigkeit statistischer Aussagen von den zugrunde liegenden Daten wird direkt erfahrbar.

Variabilität als zentrales Konzept

Erkennen und Interpretieren von Streuung und Unsicherheit in Daten.

KI-Unterstützung (interaktiv/graphisch):

- Animation von Datensätzen mit unterschiedlicher Streuung bei gleichem Mittelwert.
- Überlagerung mehrerer Verteilungen zur direkten visuellen Vergleichbarkeit.
- Interaktive Slider zur Veränderung von Varianz oder Standardabweichung mit sofortiger graphischer Rückmeldung.

Variabilität wird nicht abstrakt erklärt, sondern visuell & dynamisch erfahrbar gemacht.

Stichprobe vs. Grundgesamtheit

Verständnis der Unsicherheit bei Stichproben und deren Beziehung zur Population.

KI-Unterstützung (interaktiv/graphisch):

- Simulation von wiederholten Stichproben aus einer bekannten Population.
- Visualisierung der Verteilung von Stichprobenkennwerten (z. B. Mittelwerte).
- Vergleich: einzelne Stichprobe vs. tatsächliche Populationsverteilung.

Lernende erkennen, dass unterschiedliche Stichproben zu unterschiedlichen Ergebnissen führen, obwohl die Population gleichbleibt.

Schätzen und Testen

Verständnis von Unsicherheit in Schätzungen und inferenzstatistischen Aussagen.

KI-Unterstützung (interaktiv/graphisch):

- Visualisierung von Konfidenzintervallen als dynamische Bereiche um Schätzwerte.
- Simulation: Wie verändern sich Intervalle bei wachsender Stichprobe?
- Interaktive Darstellung von Testentscheidungen (z. B. Annahme-/Ablehnungsbereiche).

Unsicherheit wird als quantifizierbares Konzept sichtbar, nicht nur als abstrakte Idee.

Kontextabhängigkeit

Einordnung statistischer Ergebnisse im Entstehungskontext.

KI-Unterstützung (interaktiv/graphisch):

- Verknüpfung von Datenvisualisierungen mit Kontextinformationen (z. B. Beschreibung der Datenerhebung).
- Vergleich mehrerer Datensätze mit identischer Struktur, aber unterschiedlichem Kontext.
- Szenarienwechsel: gleiche Zahlen, unterschiedliche Interpretation je nach Kontext.

Lernende erkennen, dass Ergebnisse ohne Kontext keine eindeutige Bedeutung haben.

Modellierung realer Prozesse

Verstehen, dass statistische Modelle Vereinfachungen der Realität sind.

KI-Unterstützung (interaktiv/graphisch):

- Anpassung von Modellen (z. B. Regressionsgeraden) an Daten mit visueller Rückmeldung.
- Vergleich verschiedener Modelle für denselben Datensatz.
- Interaktive Manipulation von Modellannahmen und Beobachtung der Auswirkungen.

Modelle werden als veränderbare, überprüfbare Konstruktionen erfahrbar, nicht als feste Wahrheiten.

Übergreifende didaktische Leitidee

Die zentrale Stärke von KI liegt in der Kopplung von Interaktion und Visualisierung:

- Veränderungen werden sofort sichtbar
- Lernende können hypothesenbasiert explorieren
- Abstrakte Konzepte werden konkret erfahrbar

Diese Potenziale entfalten sich nur, wenn Lernende aktiv eingebunden werden:

- reine Betrachtung von Visualisierungen genügt nicht
- entscheidend ist die Verknüpfung von Handlung, Beobachtung und Reflexion

Ohne diese Einbettung besteht Gefahr, dass Visualisierungen lediglich konsumiert werden, ohne das statistische Denken tatsächlich zu vertiefen.

5. KI selbst ist nicht die interaktive Umgebung

KI ist nur ein Vermittler von Interaktivität, die in geeigneten Werkzeugen umgesetzt wird.

Vorbereitung interaktiver Lernumgebungen

Grundsätzliche Klärung

Aktuelle KI-Systeme (wie ChatGPT) können Diagramme beschreiben und statisch erzeugen; Code generieren, der Interaktivität ermöglicht; Interaktive Szenarien konzipieren und anleiten.

Aber:

Die eigentliche Interaktivität (Slider, dynamische Updates, Simulationen) entsteht in externen Umgebungen: Tabellenkalkulation (z. B. Excel); Dynamische Mathematiksoftware (z. B. GeoGebra); Programmierumgebungen (z. B. Python mit Bibliotheken wie *matplotlib*, *plotly*).

Das bedeutet:

KI ist *didaktischer und technischer Enabler*, nicht die primäre Interaktionsplattform.

Drei typische Nutzungsszenarien und deren Einschränkungen

KI als Code-Generator für Interaktivität

B: KI erzeugt Python-Code für ein Histogramm mit veränderbarer Stichprobengröße.

Lernende können Parameter aktiv verändern; Variabilität wird experimentell erfahrbar.

Erfordert minimale technische Kompetenz; Interaktivität liegt außerhalb der KI.

KI als Anleitung für bestehende Tools

B: KI erklärt schrittweise, wie man in GeoGebra einen Slider erstellt, der die Standardabweichung verändert.

Niedrigere Einstiegshürde; Fokus bleibt auf dem Konzept, nicht auf Programmierung.

Interaktivität ist durch das jeweilige Tool begrenzt; Qualität hängt von der Anleitung ab.

KI als quasi-interaktiver Dialogpartner

B: „Was passiert mit dem Mittelwert, wenn ich diesen Ausreißer entferne?“

Die KI simuliert gedanklich und beschreibt die Veränderung.

Niedrige Einstiegsschwelle; Förderung von Hypothesenbildung.

Keine echte visuelle Interaktivität; Gefahr der „Simulation im Kopf“ statt tatsächlicher Erfahrung.

Didaktisch entscheidender Punkt

Technische Interaktivität (Slider, Animationen) vs. Kognitive Interaktivität (Fragen, Hypothesen, Reflexion)

KI ist besonders stark in der kognitiven Interaktivität, während technische Interaktivität meist über externe Systeme realisiert wird.

Ein didaktisch wirksamer Unterricht kombiniert beide:

KI strukturiert das Denken, externe Tools machen die Effekte sichtbar.

Konsequenzen für den Unterricht

Realistische Einschätzung

Man kann nicht behaupten, dass KI allein interaktive Visualisierungen bereitstellt.

Die Interaktivität entsteht durch ein Ökosystem aus KI + Werkzeugen.

Ein effektives Setting besteht aus drei Komponenten:

- KI stellt Fragen, generiert Szenarien, unterstützt Interpretation
- Interaktives Tool (z. B. Excel, GeoGebra, Python) ermöglicht Manipulation von Daten/Parametern; visualisiert Effekte
- Reflexion durch Lernende verbindet Handlung und Verständnis

Positiver Effekt entsteht nicht durch Interaktivität, sondern durch deren gezielte Nutzung:

- Lernende müssen vor der Veränderung Vorhersagen treffen
- Veränderungen müssen begründet interpretiert werden
- KI sollte *nicht nur Ergebnisse liefern*, sondern Fragen stellen

B: Nicht „Hier ist das Diagramm“, sondern „Was erwartest du, wenn wir die Streuung verdoppeln? Teste es.“

- Echte Interaktivität entsteht typischerweise erst durch Tools wie Excel, GeoGebra oder Python. KI ersetzt diese Werkzeuge derzeit nicht vollständig.

Interaktivität didaktisch zugänglich, adaptiv und reflektiert nutzbar machen

Der eigentliche Mehrwert von KI liegt darin, Interaktivität didaktisch nutzbar zu machen. Ohne diese Einbettung bleibt auch die beste technische Interaktivität pädagogisch wirkungsschwach.

KI erzeugt GeoGebra-kompatiblen Code, allerdings mit einer wichtigen Präzisierung:

- GeoGebra verwendet keine klassische Programmiersprache, sondern eine Befehlsprache und algebraische Eingaben.
- KI kann diese Befehle korrekt formulieren (für Funktionen, Zufallsdaten, Slider, Regressionen etc.).
- Die Ausführung erfolgt in GeoGebra (oder einer anderen Software) selbst, nicht innerhalb von ChatGPT.

Was ChatGPT konkret leisten kann

Erzeugung von GeoGebra-Befehlen Regression und Daten:

$L1 = \{1, 2, 3, 4, 5\}$

$L2 = \{2, 3, 5, 4, 6\}$

`FitLine(L1, L2)`

ChatGPT kann Listen generieren, passende Befehle auswählen, die Struktur erklären.

Interaktive Elemente (Slider etc.)

```
a = Slider(0,5,0.1)
```

```
f(x) = a x
```

Damit kann man Parameter dynamisch verändern; die Auswirkung direkt visualisieren.

ChatGPT schlägt sinnvolle Parameterbereiche vor, entwirft didaktisch geeignete Szenarien.

Simulationen und Zufallsdaten

```
L = Sequence(RandomNormal(0,1), i, 1, 100)
```

```
Histogram(L)
```

Damit können Stichproben simuliert, Variabilität sichtbar gemacht werden.

Dynamische statistische Konzepte Beispiel (Mittelwert dynamisch):

```
Mean(L)
```

In Kombination mit variablen Daten verändert sich der Mittelwert automatisch;

Effekte von Ausreißern werden sichtbar.

Einschränkungen und typische Fehlerquellen

Syntaxabhängigkeit von GeoGebra-Versionen Classic vs. Web-Version.

Keine direkte Ausführung ChatGPT kann Code nicht testen, daher sind kleine Anpassungen manchmal nötig.

Komplexere Konstruktionen Bei sehr komplexen dynamischen Szenarien (z. B. mehrstufige Simulationen) kann Nacharbeit erforderlich sein.

Didaktische Qualität nicht automatisch gegeben Der Code funktioniert technisch, ist aber nicht automatisch lernwirksam.

Didaktisch besonders wertvolle Einsatzform

Der größte Nutzen entsteht, wenn ChatGPT nicht nur Code liefert, sondern eine Lernumgebung entwirft; Fragestellungen vorgibt; Interpretation anleitet.

Beispielstruktur:

1. Hypothesen formulieren
2. GeoGebra-Konstruktion erstellen (mit KI-Hilfe)
3. Parameter verändern
4. Beobachtungen interpretieren

ChatGPT kann GeoGebra-Befehle zuverlässig erzeugen und ist besonders hilfreich bei:

- der schnellen Erstellung interaktiver Modelle
- der didaktischen Strukturierung von Experimenten
- der Verbindung von Konzept und Visualisierung

Die eigentliche Stärke liegt jedoch nicht im Code selbst, sondern darin, interaktive mathematisch-statistische Lernumgebungen konzeptionell zugänglich zu machen.

6. Fragen nach der Gestaltung in GeoGebra

Kann GeoGebra Intervalle unter einer Dichtefunktion durch Code färben oder muss man das händisch, wie in Excel machen?

Ja, GeoGebra kann Intervalle unter einer Dichtefunktion per Befehl färben. Das erfolgt nicht händisch wie in Excel, sondern über integrierte Befehle und Objekte.

1. Zentrale Möglichkeit: `Integral(...)`

GeoGebra erlaubt es, Flächen unter Funktionen direkt zu definieren und automatisch zu schattieren. **B** (Normalverteilung):

$$f(x) = \text{Normal}(0, 1, x)$$

$$A = \text{Integral}(f, -1, 1)$$

- A ist dann die Fläche unter der Dichtefunktion im Intervall $[-1,1]$
- Diese Fläche wird automatisch graphisch dargestellt und gefärbt
- Farbe & Transparenz können auch angepasst werden (Eigenschaften)

Die Verbindung zwischen Wahrscheinlichkeit und Fläche wird unmittelbar sichtbar.

2. Alternative: `IntegralBetween`

Für zwei Funktionen:

$$\text{IntegralBetween}(f, g, a, b)$$

Bei Dichtefunktionen meist nicht notwendig, aber nützlich für allgemeine Fälle.

3. Dynamische Intervalle mit Slidern

Interaktivität entsteht durch Parametersteuerung:

```
a = Slider(-3, 3, 0.1)
```

```
b = Slider(-3, 3, 0.1)
```

```
A = Integral(f, a, b)
```

- Das gefärbte Intervall passt sich *dynamisch* an
- Lernende können Wahrscheinlichkeiten explorieren

4. Wahrscheinlichkeit direkt berechnen. Zusätzlich kann man den Wert anzeigen:

```
P = Integral(f, a, b)
```

GeoGebra liefert numerischen Wert (Wahrscheinlichkeit);
visuelle Fläche (Interpretation)

5. Gestaltung der Färbung wird nicht im Code, sondern über Eigenschaften gesteuert:

- Farbe wählen; Transparenz einstellen; Linien ausblenden

Kein „händisches Zeichnen“, sondern visuelle Anpassung eines definierten Objekts.

6. Vergleich zu Excel

- In Excel muss man Flächen über Tricks (z. B. Fehlerbalken) approximieren
- In GeoGebra ist die Fläche ein mathematisches Objekt, kein grafischer Workaround

7. Fazit

GeoGebra bietet eine direkte, mathematisch konsistente Möglichkeit, Intervalle unter Dichtefunktionen zu färben:

- formal korrekt (Integral als Objekt)
- dynamisch steuerbar (Slider)
- didaktisch besonders geeignet für Wahrscheinlichkeitsinterpretationen

Im Unterschied zu Excel

- ist dies keine manuelle Visualisierung,
- sondern eine modellbasierte Darstellung, die eng mit dem zugrunde liegenden Konzept verbunden ist.

Kann man in GeoGebra die Nullhypothese H_0 und eine ausgewählte Alternative H_1 durch eine normalverteilte Dichte darstellen, Ablehnzahl und Alternative variabel halten, den Ablehnbereich unter H_0 und den Nicht-Ablehnbereich unter H_1 färben und dann dynamisch die Alternative verschieben oder die Ablehnzahl ändern?

Ja, das ist in GeoGebra möglich und didaktisch sehr wirkungsvoll. Man kann sowohl *Nullhypothese* als auch eine *Alternative* durch Normaldichten darstellen, kritische Werte (Ablehnzahl) variabel machen & die relevanten Flächen dynamisch einfärben. Im Kern wird damit ein interaktives Modell eines Hypothesentests realisiert.

1. Grundidee der Konstruktion

Man definiert zwei Dichtefunktionen:

- Nullhypothese H_0 : $f_0(x)$, z. B. $N(\mu_0, \sigma)$
- Alternative H_1 : $f_1(x)$, z. B. $N(\mu_1, \sigma)$

Dazu kommt ein kritischer Wert c (Ablehnzahl), der den Ablehnbereich bestimmt.

2. Typische GeoGebra-Struktur

Definition der Dichten

$$\mu_0 = 0$$

$$\mu_1 = \text{Slider}(-2, 2, 0.1)$$

$$\sigma = 1$$

$$f_0(x) = \text{Normal}(\mu_0, \sigma, x)$$

$$f_1(x) = \text{Normal}(\mu_1, \sigma, x)$$

- μ_1 ist variabel \rightarrow Alternative verschiebbar
- Beide Dichten werden gleichzeitig angezeigt

Kritischer Wert (Ablehnzahl)

$c = \text{Slider}(-3, 3, 0.1)$

- c definiert den Ablehnbereich (z. B. rechtsseitiger Test)

Ablehnbereich unter H_0 (Fehler 1. Art)

$A_0 = \text{Integral}(f_0, c, 5)$

- Fläche unter H_0 im Ablehnbereich
- entspricht Signifikanzniveau α (bei festem c)

Nicht-Ablehnbereich unter H_1 (Fehler 2. Art)

$B = \text{Integral}(f_1, -5, c)$

- Fläche unter H_1 im Annahmebereich
- entspricht β -Fehler

Teststärke (Power)

$\text{Power} = \text{Integral}(f_1, c, 5)$

- Wahrscheinlichkeit, H_0 korrekt zu verwerfen

3. Dynamische Exploration – Realisierung von zwei zentralen Variationen

Verschieben der Alternative (μ_1)

- Abstand zwischen H_0 und H_1 vergrößert sich
- β nimmt ab, Power steigt
- Lernende sehen direkt den Einfluss der Effektgröße

Veränderung der Ablehnzahl (c)

- Ablehnbereich wird größer oder kleiner
- Trade-off zwischen α und β wird sichtbar

4. Die Konstruktion erlaubt es, zentrale Einsichten gleichzeitig zu fördern:

1. Zusammenhang zwischen α , β und Power wird visuell erfahrbar
2. Einfluss der Effektgröße wird unmittelbar sichtbar
3. Entscheidungsregeln werden nicht nur formal, sondern visuell verstanden
4. Trade-offs (z. B. strenger Test vs. Teststärke) werden explorierbar

5. Wichtige Präzisierungen

- Die Integrationsgrenzen (z. B. $-5, 5$) sind numerische Approximationen für „ $\pm\infty$ “
- Für zweiseitige Tests müssten zwei Ablehnbereiche definiert werden
- Die grafische Färbung erfolgt über die Integralobjekte (Eigenschaften einstellbar)

6. Fazit: GeoGebra ermöglicht genau das beschriebene Szenario

- zwei Normaldichten (Nullhypothese und Alternative)
- variabler kritischer Wert
- dynamisch gefärbte Fehlerbereiche
- interaktive Veränderung von Effektgröße und Testniveau

Didaktisch entsteht damit ein sehr starkes Werkzeug

Abstrakte Begriffe wie

Fehler 1. und 2. Art, Teststärke und Entscheidungsregel

werden in eine *unmittelbar visuell erfahrbare Struktur* überführt.

Gibt es dafür eine Möglichkeit, Code in GeoGebra zu erzeugen, der das alles leistet und die Slider für die Veränderung parat stellt?

Ja. Das ist in GeoGebra per Code möglich. GeoGebra stellt dafür die nötigen Befehle bereit: Slider für veränderbare Parameter, Integral für Flächen unter einer Funktion sowie SetColor und SetFilling für Farbe und Transparenz.

Für die Normaldichte wird `Normal(μ , σ , x , false)` verwendet; der Parameter `false` sorgt dafür, dass die Dichte anstelle der Verteilungsfunktion verwendet wird.

Wichtig ist die Präzisierung:

Statt des „Fehlers“ wird die *zugehörige Fläche unter der jeweiligen Dichte* gefärbt.

Genau das lässt sich mit zwei Dichten und zwei Integralobjekten umsetzen: Ablehnbereich unter H_0 einerseits und Nicht-Ablehnbereich unter H_1 andererseits. Diese Flächen aktualisieren sich dynamisch, sobald man den kritischen Wert oder die Lage der Alternativverteilung per Slider verändert.

Hier ist ein GeoGebra-Code für einen *rechtsseitigen Test mit verschiebbarer Alternative und variablem kritischem Wert*. Ich habe ihn so formuliert, dass er in GeoGebra Classic als Folge von Eingaben funktionieren sollte; Details können je nach Version kleine Anpassungen erfordern.

ChatGPT erzeugt diesen GeoGebra-Code (zeilenweise eingeben)

```
mu0 = 0
sigma = 1
mu1 = Slider(-1, 4, 0.1, 1, 140, false, true, false, false)
c = Slider(-1, 4, 0.1, 1, 140, false, true, false, false)
f0(x) = Normal(mu0, sigma, x, false)
f1(x) = Normal(mu1, sigma, x, false)
A0 = Integral(f0, c, 6)
B1 = Integral(f1, -6, c)
SetColor(f0, 0, 0, 1)
SetColor(f1, 1, 0, 0)
SetColor(A0, 0, 0, 1)
SetFilling(A0, 0.35)
SetColor(B1, 1, 0, 0)
SetFilling(B1, 0.35)
crit = x = c
SetColor(crit, 0, 0, 0)
alpha = Integral(f0, c, 6)
beta = Integral(f1, -6, c)
Text("H0", (-0.6, 0.43))
Text("H1", (mu1 - 0.2, 0.38))
Text("kritischer Wert c = " + c, (3.2, 0.33))
Text("Alpha = " + Round(alpha, 4), (3.2, 0.28))
Text("Beta = " + Round(beta, 4), (3.2, 0.23))
```

Was dieser Code leistet

μ_1 verschiebt die Alternativverteilung, c verändert den kritischen Wert.

$A_0 = \text{Integral}(f_0, c, 6)$ färbt den rechten Ablehnbereich unter H_0 ,

$B_1 = \text{Integral}(f_1, -6, c)$ den linken Nicht-Ablehnbereich unter H_1 .

`SetColor` & `SetFilling` sorgen dafür, dass diese Flächen sichtbar und transparent eingefärbt sind.

Die Grenzen -6 und 6 sind eine Approximation für „hinreichend weit links/rechts“ bei der Standardabweichung 1 .

Nun kann man zwei Arten von Dynamik sehen

Wenn man μ_1 nach rechts verschiebt, wird die Trennung der beiden Dichten größer und der β -Bereich typischerweise kleiner.

Wenn man c nach rechts verschiebt, wird der Ablehnbereich unter H_0 kleiner, aber der Nicht-Ablehnbereich unter H_1 größer.

Der Zielkonflikt zwischen Fehler 1. und 2. Art wird graphisch, nicht nur rechnerisch erfahrbar.

Die kurze Antwort auf die Frage lautet also:

Ja, man kann in GeoGebra Code erzeugen, der beides gleichzeitig leistet: die Flächenfärbung unter beiden Dichten und Slider für kritischen Wert sowie Alternative.

GeoGebra muss dafür nicht „händisch wie Excel“ bearbeitet werden; die Flächen sind echte Integralobjekte, deren Farbe und Transparenz per Befehl gesetzt werden können.

Erweiterungen

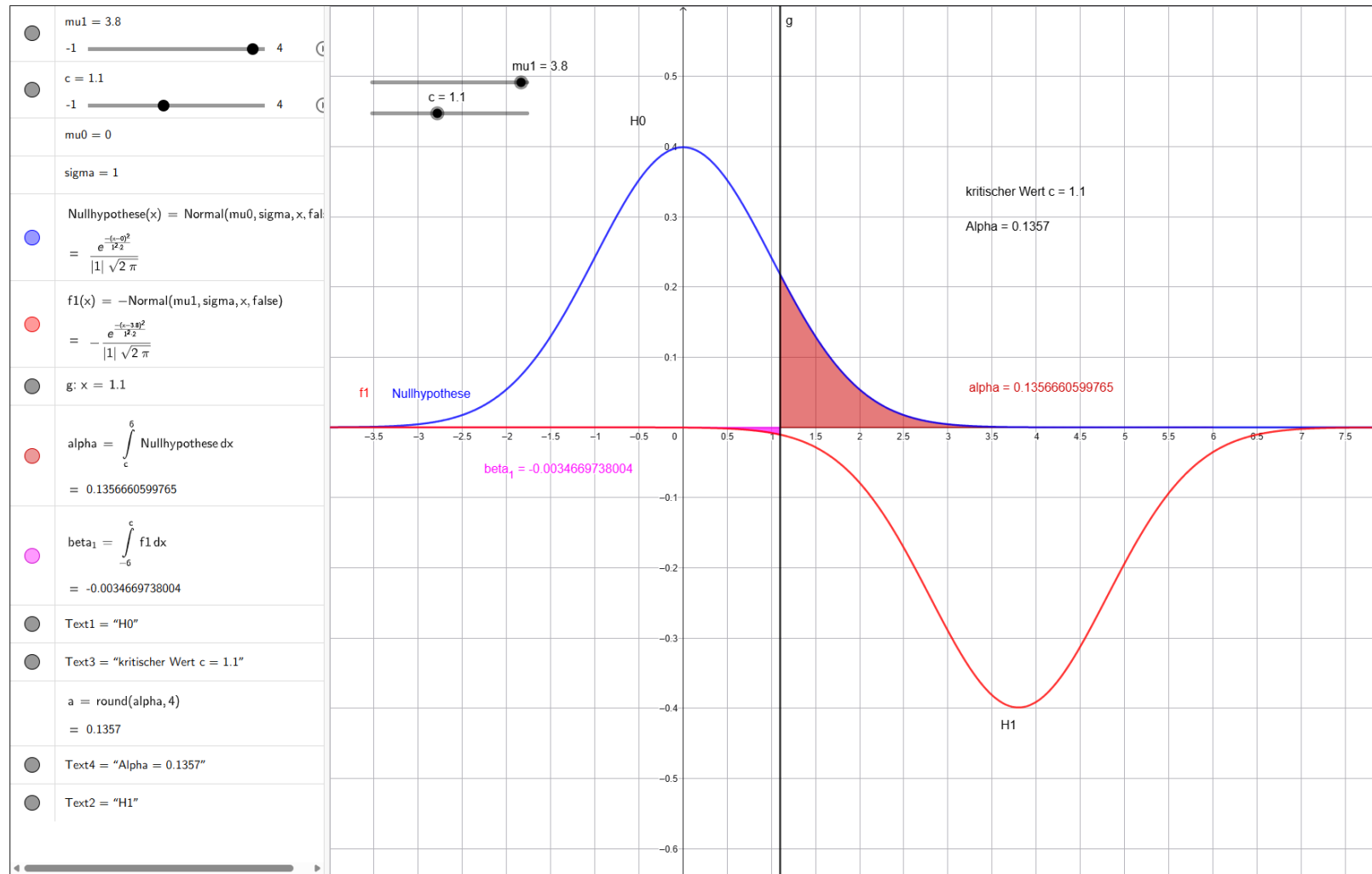
Für einen *linksseitigen Test*

müsste man die Integrale spiegeln, also etwa `Integral(f0, -6, c)` und `Integral(f1, c, 6)`.

Für einen *zweiseitigen Test*

braucht man kritische Werte c_1 und c_2 , und unter H_0 zwei gefärbte Randbereiche; unter H_1 wäre dann der mittlere Nicht-Ablehnbereich zu färben.

Die gleiche Logik bleibt erhalten, nur mit drei statt zwei Flächenobjekten.



7. Conclusio 1 – Grundverständnis von KI im Unterricht

Zentrale Einsicht des Beitrags:

Generative KI verändert die Bedingungen des Lernens in Statistik und Wahrscheinlichkeit grundlegend; sie eröffnet neue Zugänge zu Daten, Modellen und Interpretationen

Aber:

KI ersetzt weder statistisches noch probabilistisches Denken; sie kann Verständnis nicht automatisch erzeugen.

Kernthese:

KI ist kein Ersatz für Denken, sondern ein Medium zur Unterstützung von Denkprozessen.

Didaktische Konsequenz:

Nicht die Technologie entscheidet über den Lernerfolg, sondern die Art und Weise ihrer Einbettung in den Unterricht.

7. Conclusio 2 – Rolle von Fachwissen und Kompetenzverschiebung

Zentraler Befund:

KI ist umso effizienter, je besser das Fach beherrscht wird.

Implikationen:

KI kann auch Laien scheinbar tragfähige Erklärungen liefern; diese bleiben jedoch häufig oberflächlich oder kontextlos.

Neue Kompetenzanforderungen:

- Fähigkeit, geeignete Fragen zu stellen
- Ergebnisse kritisch zu interpretieren
- Annahmen und Voraussetzungen zu prüfen

Verschiebung im Lernen:

Weg von Durchführung, Rechnen, Softwarebedienung hin zu Verstehen, Bewerten, Reflektieren.

7. Conclusio 3 – Didaktische Potenziale von KI

Förderung statistischen Denkens:

- Zugang zu großen und realitätsnahen Datensätzen
- exploratives Arbeiten mit Hypothesen und Modellen
- dynamische Visualisierung von Variabilität und Zusammenhängen

Förderung probabilistischen Denkens:

- Simulation von Zufallsexperimenten in großer Zahl
- Vergleich alternativer Szenarien
- Sichtbarmachung langfristiger Häufigkeiten

Förderung von Literalitäten:

- Analyse realer Kontexte (z. B. Medienberichte)
- Vergleich unterschiedlicher Darstellungen und Interpretationen
- Unterstützung bei Kommunikation und Argumentation

Zentrale Stärke:

Verbindung von kognitiver Interaktivität (Dialog, Fragen)
und technischer Interaktivität (Visualisierung, Simulation).

7. Conclusio 4 – Risiken und strukturelle Grenzen

Zentrale Gefahren:

- Automatisierung von Analyseprozessen → oberflächliches Verständnis
- Autoritätsbias gegenüber KI-generierten Aussagen
- Übernahme von Ergebnissen ohne Prüfung

Methodische Grenzen der KI:

- fehlende zuverlässige Prüfung von Voraussetzungen
- mögliche Fehler und inkonsistente Ergebnisse
- begrenzte Transparenz der zugrunde liegenden Modelle

Didaktisches Kernproblem:

- Lernende sehen Ergebnisse, aber verstehen deren Entstehung nicht

Konsequenz:

Statistisches und probabilistisches Denken werden nicht ersetzt, sondern gewinnen an Bedeutung

7. Conclusio 5 – Didaktische Schlussfolgerung und Leitprinzipien

Bedingungen für erfolgreichen KI-Einsatz:

- klare didaktische Einbettung in strukturierte Lernprozesse
- Aufgaben mit Fokus auf Begründung, Interpretation und Kritik
- aktive Rolle der Lehrperson als epistemische Instanz

Zentrale Leitprinzipien:

- Primat der Begründung vor dem Ergebnis
- KI als kognitives Werkzeug, nicht als Autorität
- systematische Reflexion der KI-Nutzung

Konkrete Umsetzung:

- Vergleich von KI- und Schülerlösungen
- gezielte Analyse von Fehlern
- Einforderung von Begründungen vor und nach KI-Nutzung

Abschließende Kernaussage:

KI entfaltet ihr Potenzial nur dann, wenn sie Denkprozesse unterstützt und nicht ersetzt.

Transformation des Unterrichts: von Rechnen und Reproduzieren hin zu Verstehen, Interpretieren und Reflektieren.

7. Conclusio zum KI-Einsatz im Stochastikunterricht

Zentral:

Nicht die Technologie entscheidet über den Lernerfolg sondern die Art und Weise ihrer Einbettung in den Unterricht.

Rolle von Fachwissen und Kompetenzverschiebung

Weg von: Durchführung, Rechnen, Softwarebedienung
hin zu: Verstehen, Bewerten, Reflektieren.

Didaktische Potenziale von KI:

Verbindung von kognitiver Interaktivität (Dialog, Fragen)
und technischer Interaktivität (Visualisierung, Simulation).

Risiken und strukturelle Grenzen:

Statistisches und probabilistisches Denken werden nicht ersetzt,
sondern gewinnen an Bedeutung.

Didaktische Schlussfolgerung und Leitprinzipien:

KI entfaltet ihr Potenzial nur dann, wenn sie Denkprozesse unterstützt
und nicht ersetzt.

Transformation des Unterrichts: von Rechnen und Reproduzieren
hin zu Verstehen, Interpretieren und Reflektieren.

A. Stochastisches Denken und stochastische Literalität mit KI

B. Statistische Analyse mit KI als Werkzeug

C. KI in der Datenanalyse – Konsequenzen für Lehre und Praxis

1. Orientierung: Analyseleitfäden durch KI
2. Erster Datensatz: Einfach strukturierte ANOVA-Situation
3. Zweiter Datensatz: ANOVA mit zusätzlicher Verletzung der Voraussetzungen
4. Stabilitätsprüfungen
5. Weitere Datensätze (t-Test und Regression) – nicht im Artikel abgedeckt
6. Fazit aus der Analyse der Datensätze

Der Einsatz generativer KI in der statistischen Datenanalyse und ihre Auswirkungen auf die Lehre der Statistik

J. Schwarz, *Stochastik in der Schule*, open access,

<https://www.stochastik-in-der-schule.de/sisonline/46-01.htm>

1. Orientierung: Analyseleitfäden durch KI

Datensatz / Situation

Kein konkreter Datensatz im Vordergrund; es geht um allgemeine Datenanalyse.

Fragestellung

- „Welche Kennzahlen und Grafiken sind für eine deskriptive Datenanalyse erforderlich?“
- „Wie gehe ich bei einer konkreten Analysefrage vor?“

Prompts

- Allgemeine Fragen zur Vorgehensweise
- Problemorientierte Fragen in Alltagssprache

Ergebnis

Die KI liefert sinnvolle Übersichten zu Kennzahlen und Diagrammen und strukturierte Analyseabläufe (Deskription → Visualisierung → Test)

Einschränkung:

Leitfäden sind unvollständig; z. B. fehlende Prüfung von Varianzgleichheit oder Ausreißern.

2. Erster Datensatz: Einfach strukturierte ANOVA-Situation

Datensatz

Zwei Variablen: metrische Variable („MetricVar“); kategoriale Variable mit 3 Faktorstufen; 30 Beobachtungen (je 10 pro Gruppe);
unterschiedlich starke Streuung zwischen Gruppen.

Fragestellung

„Unterscheidet sich MetricVar zwischen den Factor-Stufen?“

Prompts

- „Bitte untersuche statistisch, ob sich MetricVar zwischen den Factor-Stufen unterscheidet.“
- später: „Kannst du mir das Ergebnis erklären?“

Analyse durch KI

Deskriptive Statistik; Boxplots (siehe Abb. 2); ANOVA-Test (siehe Abb. 3)

Ergebnisse

Kein signifikanter Unterschied zwischen den Gruppen ($p \approx 0.94$).
Interpretation wird verständlich erklärt (siehe Abb. 5)

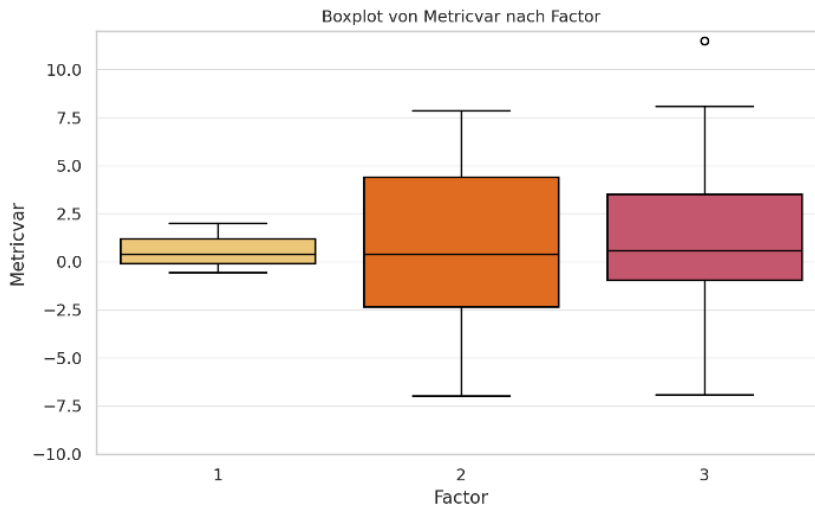
Bewertung

Analyse ist insgesamt ok (Kruskal-Wallis würde zu denselben Ergebnissen führen), **aber**:

- Varianzunterschiede werden ignoriert
- keine Prüfung der Voraussetzungen
- geringe Stichprobengröße wird nicht reflektiert

Beginnen wir mit den deskriptiven Statistiken und der Visualisierung.

Boxplot von MetricVar nach Factor



Factor 1

- Anzahl: 10
- Mittelwert: 0.5473
- Standardabweichung: 0.8354
- Minimum: -0.5647
- 25% Quantil: -0.0867
- Median: 0.3837
- 75% Quantil: 1.1864
- Maximum: 2.0184

Die deskriptiven Statistiken für MetricVar in den verschiedenen Factor - Stufen lauten:

Factor 2

- Anzahl: 10
- Mittelwert: 1.1096
- Standardabweichung: 5.7803
- Minimum: -6.9065
- 25% Quantil: -0.9381
- Median: 0.5902
- 75% Quantil: 3.5105
- Maximum: 11.4760

Factor 3

- Anzahl: 10
- Mittelwert: 0.5096
- Standardabweichung: 4.8918
- Minimum: -6.9694
- 25% Quantil: -2.3381
- Median: 0.3818
- 75% Quantil: 4.4129
- Maximum: 7.8599

Als Nächstes führen wir einen ANOVA-Test durch, um zu bestimmen, ob sich die Mittelwerte von MetricVar signifikant zwischen den Factor–Stufen unterscheiden.

[\[<_\]](#)

Die Ergebnisse des ANOVA-Tests lauten wie folgt:

- **F-Statistik:** 0.0584
- **p-Wert:** 0.9434

Da der p-Wert deutlich größer als 0.05 ist, können wir die Nullhypothese nicht verwerfen. Dies legt nahe, dass **kein statistisch signifikanter Unterschied** zwischen den Mittelwerten von MetricVar über die verschiedenen Factor–Stufen hinweg besteht.

Falls du weitere Fragen hast oder zusätzliche Analysen brauchst, sag mir einfach Bescheid!

[\[<_\]](#)

Natürlich! Hier ist eine vereinfachte Erklärung des Ergebnisses:

1. Was wir gemacht haben:

- Wir wollten prüfen, ob sich die durchschnittlichen Werte (Mittelwerte) von **MetricVar** zwischen den verschiedenen **Factor**-Stufen (1, 2 und 3) unterscheiden.
- Dafür haben wir einen statistischen Test namens **ANOVA** (Analyse der Varianz) verwendet. Dieser Test vergleicht die Mittelwerte von **MetricVar** für jede **Factor**-Stufe, um zu sehen, ob sie sich signifikant voneinander unterscheiden.

2. Ergebnis:

- Der ANOVA-Test liefert zwei wichtige Kennzahlen: die **F-Statistik** und den **p-Wert**.
- Die **F-Statistik** (0.0584) misst das Verhältnis der Varianz zwischen den Gruppen zur Varianz innerhalb der Gruppen. Für unsere Interpretation ist aber der **p-Wert** entscheidend.

3. p-Wert:

- Der **p-Wert** (0.9434) sagt aus, wie wahrscheinlich es ist, dass die beobachteten Unterschiede rein zufällig zustande gekommen sind.
- Allgemein gilt: Ist der p-Wert kleiner als 0,05, dann ist der Unterschied statistisch signifikant – es spricht also etwas dafür, dass sich die Gruppen wirklich unterscheiden.
- In unserem Fall ist der p-Wert **0.9434**, also deutlich größer als 0.05. Das heißt, es ist sehr wahrscheinlich, dass die beobachteten Unterschiede rein zufällig sind – also **keine echten Unterschiede zwischen den Gruppen** vorliegen.

3. p-Wert:

- Der **p-Wert** (0.9434) sagt aus, wie wahrscheinlich es ist, dass die beobachteten Unterschiede rein zufällig zustande gekommen sind.
- Allgemein gilt: Ist der p-Wert kleiner als 0,05, dann ist der Unterschied statistisch signifikant – es spricht also etwas dafür, dass sich die Gruppen wirklich unterscheiden.
- In unserem Fall ist der p-Wert **0.9434**, also deutlich größer als 0.05. Das heißt, es ist sehr wahrscheinlich, dass die beobachteten Unterschiede rein zufällig sind – also **keine echten Unterschiede zwischen den Gruppen** vorliegen.

4. Fazit:

- Weil der p-Wert so hoch ist, schließen wir, dass **kein signifikanter Unterschied** zwischen den Mittelwerten von **MetricVar** über die verschiedenen **Factor–**Stufen besteht.
- Einfacher gesagt: **MetricVar** unterscheidet sich **nicht nennenswert** zwischen den **Factor–**Stufen – zumindest basierend auf unseren Daten.

3. Zweiter Datensatz: ANOVA mit Verletzung der Voraussetzungen

Datensatz

Zwei Variablen (metrisch + Faktor mit 3 Stufen); 180 Beobachtungen (je 60 pro Gruppe). Enthält nun Ausreißer sowie fehlende Werte. **Robuste Verfahren liefern andere Ergebnisse.**

Fragestellung

„Unterscheiden sich die Werte zwischen den Faktorstufen?“

Prompts

Analog zum ersten Datensatz (alltagssprachlich formuliert)

Analyse durch KI

Erkennt fehlende Werte; erstellt deskriptive Statistik (nur Mittelwert, Standardabweichung). Führt ANOVA durch, aber keine Prüfung der Voraussetzungen.

Ergebnisse

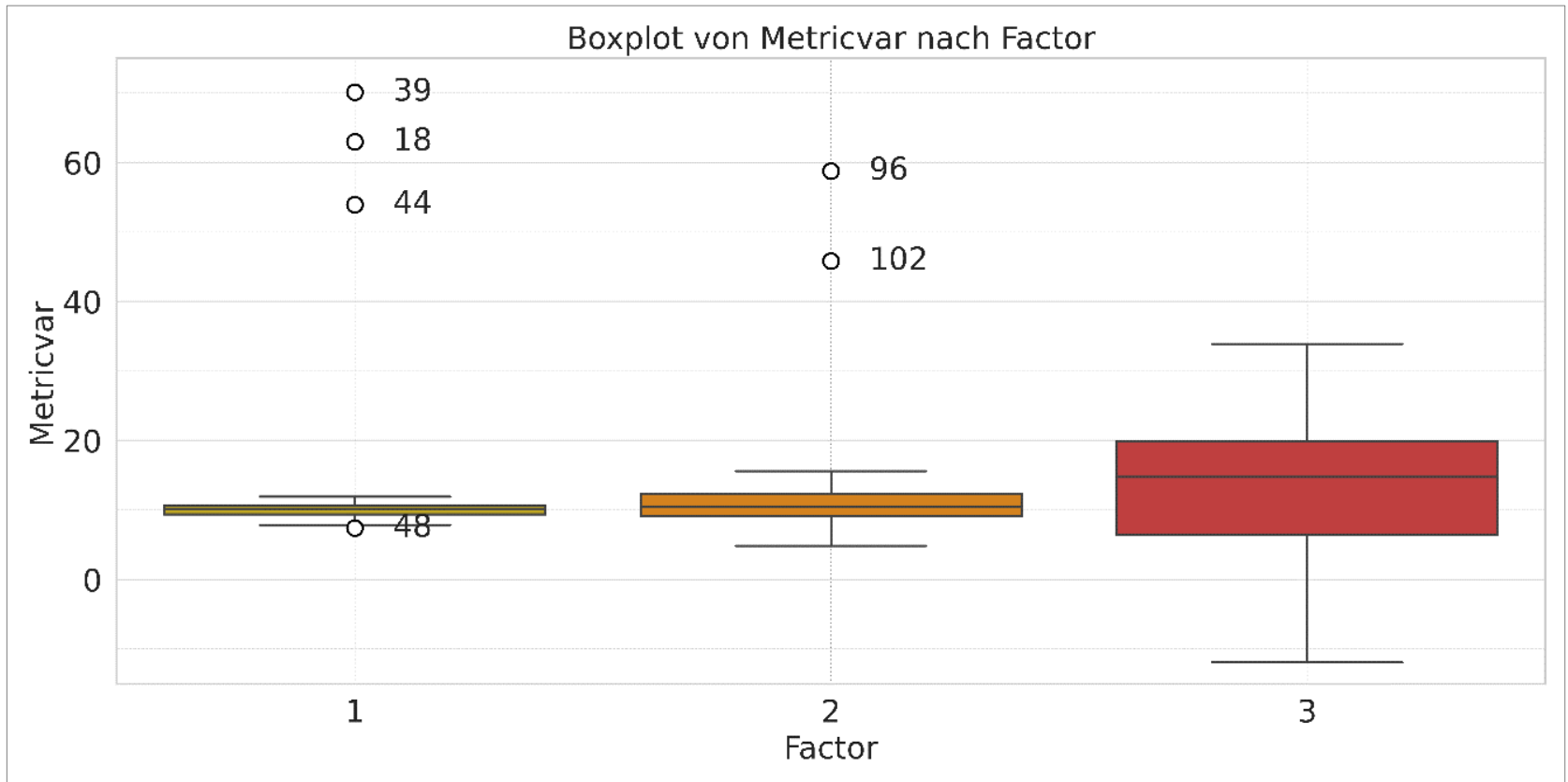
Kein signifikanter Unterschied laut ANOVA

Ergebnis mit Kruskal-Wallis (robust): signifikanter Unterschied vorhanden

Boxplots zeigen deutlich Ausreißer, diese werden von der KI wohl erkannt, aber nicht weiter diskutiert

Bewertung

Analyse ist formal falsch: Ausreißer werden ignoriert; Voraussetzungen nicht geprüft, falsches Verfahren gewählt.



Bitte untersuche statistisch, ob sich die Werte der Variable `MetricVar` zwischen den verschiedenen Stufen von `Factor` unterscheiden.

Die deskriptiven Statistiken für `MetricVar`, gruppiert nach `Factor`-Stufen, lauten wie folgt:

- **Factor 1:** Mittelwert = 12.87, Standardabweichung = 12.37
- **Factor 2:** Mittelwert = 12.08, Standardabweichung = 8.64
- **Factor 3:** Mittelwert = 11.13, Standardabweichung = 10.49

Ein ANOVA-Test wurde durchgeführt, um zu prüfen, ob sich `MetricVar` signifikant zwischen den verschiedenen `Factor`-Stufen unterscheidet. Das Ergebnis lautet:

- **F-Statistik** = 0.135
- **p-Wert** = 0.873

Da der p-Wert deutlich größer als 0,05 ist, gibt es **keinen statistisch signifikanten Unterschied** in `MetricVar` zwischen den verschiedenen `Factor`-Stufen.

Möchtest du weitere Analysen oder Visualisierungen? [[<_](#)]

4. Stabilitätsprüfungen

Datensatz

Gleiche Datensätze wie oben, mehrfach analysiert

Fragestellung

Liefert KI konsistente Ergebnisse?

Prompts

Identische Analyseanfragen wie zuvor

Ergebnisse

Teilweise unterschiedliche Vorgehensweisen:

- einmal ANOVA, einmal Welch-ANOVA oder Kruskal-Wallis
- Ergebnisse variieren entsprechend

Bewertung

KI ist nicht stabil und nicht reproduzierbar, zeigt inkonsistentes methodisches Verhalten

5. Weitere Datensätze (t-Test & Regression) – nicht im Artikel

Datensätze

Einfacher t-Test-Datensatz bzw. Datensatz für multiple lineare Regression

Fragestellungen

- Gruppenvergleich (t-Test)
- Zusammenhangsanalyse (Regression)

Prompts

Einfache, alltagssprachliche Analyseaufforderungen

Ergebnisse

t-Test: KI führt Test korrekt durch, aber keine Prüfung der Voraussetzungen, teilweise unvollständige Ausgabe.

Regression: KI führt multiple Regression durch; gelegentlich zusätzliche Analysen (Korrelationen); Annahmen werden selten geprüft.

Bewertung

Verfahren werden formal korrekt angewendet aber methodisch oft unvollständig reflektiert.

5. Fazit aus der Analyse der Datensätze

Die Beispiele im Beitrag zeigen ein konsistentes Muster:

KI kann:

- vollständige Analyseprozesse durchführen
- Visualisierungen erzeugen
- Ergebnisse verständlich erklären

KI kann nicht zuverlässig:

- Voraussetzungen prüfen
- Datenprobleme erkennen (Ausreißer, Varianzunterschiede)
- konsistent arbeiten

Diagramme (insbesondere Boxplots) enthalten oft wichtige Hinweise, die jedoch von der KI nicht interpretiert werden

Kernaussage

- Generative KI ist ein leistungsfähiges Werkzeug für standardisierte Analysen.
- aber kein verlässlicher Ersatz für statistisches Denken.

Die zentrale Kompetenz verschiebt sich daher von der Durchführung zur kritischen Interpretation und Kontrolle der Ergebnisse.

A. Stochastisches Denken und stochastische Literalität mit KI

B. Statistische Analyse mit KI als Werkzeug

C. KI in der Datenanalyse – Konsequenzen für Lehre und Praxis

1. Generative KI als Ermöglicher statistischer Datenanalyse
2. Standardisierung statistischer Analyse als Voraussetzung für Automatisierung
3. Leistungsfähigkeit der KI: Zwischen Effizienz und Oberflächlichkeit
4. Inkonsistenz und mangelnde Zuverlässigkeit
5. Grenzen generativer KI: Kein konzeptionelles Verständnis
6. Demokratisierung vs. Qualitätsverlust
7. Konsequenzen für die Statistikausbildung
8. Veränderung der Rolle von Software und Programmierung
9. Fokus auf Datenverständnis und kritischem Denken
10. Schlussfolgerungen

KI ist umso effizienter, je besser man das Fach beherrscht.
KI kann allerdings – mit erheblicher Variabilität zwar –
auch dem Laien vieles „erklären“.

Generative KI in der statistischen Datenanalyse – Potenziale, Grenzen und Konsequenzen für Lehre und Praxis

Der Einsatz generativer KI in der statistischen Datenanalyse
und ihre Auswirkungen auf die Lehre der Statistik

J. Schwarz, Stochastik in der Schule, open access,
<https://www.stochastik-in-der-schule.de/sisonline/46-01.htm>

Vorbemerkungen zum Artikel

Die rasante Entwicklung generativer KI-Systeme wie ChatGPT markiert einen tiefgreifenden Einschnitt in der statistischen Datenanalyse.

Während statistische Methoden traditionell spezialisiertes Fachwissen, methodische Kompetenz und häufig auch Programmierkenntnisse erforderten, verspricht generative KI eine weitgehende Demokratisierung dieser Prozesse.

Der Beitrag von Joachim Schwarz analysiert diese Entwicklung differenziert und zeigt sowohl die Chancen als auch die Grenzen dieses technologischen Wandels auf.

Im Zentrum steht dabei die Frage, wie sich die Arbeitsteilung zwischen Mensch und Maschine verändert & welche Konsequenzen sich daraus für die Statistikausbildung ergeben.

1. Generative KI als Ermöglicher statistischer Datenanalyse

Eine der zentralen Aussagen des Beitrags:

Generative KI erleichtert die Durchführung statistischer Analysen erheblich und macht sie sogar für Personen mit minimalen Statistikkenntnissen zugänglich.

Dies geschieht vor allem durch zwei Mechanismen:

Erstens kann KI auf Basis natürlicher Sprache Analyseprozesse strukturieren und durchführen. Nutzer müssen keine formalen statistischen Fachbegriffe oder Programmiersprachen beherrschen, sondern können ihre Fragestellungen alltagssprachlich formulieren.

Zweitens ist KI in der Lage, automatisch Code zu generieren und auszuführen, wodurch klassische Hürden wie die Beherrschung von R oder Python entfallen.

Diese Entwicklung hat weitreichende Implikationen.

Statistische Analyse wird nicht mehr primär durch technische Fähigkeiten begrenzt, sondern durch **die Fähigkeit, geeignete Fragen zu stellen und Ergebnisse zu interpretieren.**

Damit verschiebt sich die Kompetenzanforderung fundamental.

2. Standardisierung statistischer Analyse als Voraussetzung für Automatisierung

Ein wesentlicher Grund für die Leistungsfähigkeit generativer KI liegt in der strukturellen Natur statistischer Verfahren.

Viele statistische Analysen folgen standardisierten Abläufen: Datenaufbereitung, deskriptive Statistik, Visualisierung, Auswahl eines geeigneten Tests und Interpretation der Ergebnisse.

Diese Standardisierung macht statistische Verfahren besonders geeignet für Automatisierung.

Sobald das Skalenniveau von Variablen bekannt ist, kann die KI regelbasiert entscheiden, welche Verfahren anwendbar sind.

Dies gilt insbesondere für grundlegende Methoden wie t-Tests, ANOVA oder lineare Regression.

Die Konsequenzen sind klar:

- Für einfach strukturierte Probleme kann KI große Teile der Analyse vollständig übernehmen.
- Dies stellt einen Paradigmenwechsel dar, da die Durchführung statistischer Analysen nicht mehr exklusiv an Experten gebunden ist.

3. Leistungsfähigkeit der KI: Zwischen Effizienz und Oberflächlichkeit

Die empirischen Analysen im Beitrag zeigen jedoch ein ambivalentes Bild.

Einerseits ist KI durchaus in der Lage, vollständige Analysen durchzuführen, einschließlich Datenbeschreibung, Visualisierung und Hypothesentest. Selbst Nutzer ohne statistische Vorkenntnisse können so inferenzstatistische Verfahren anwenden.

Andererseits sind diese Analysen häufig unvollständig oder oberflächlich.

Ein zentrales Problem besteht darin, dass wesentliche Schritte – insbesondere die Überprüfung von Modellannahmen – nicht konsequent durchgeführt werden:

- Unterschiede in Varianzen werden ignoriert
- Ausreißer werden nicht erkannt
- fehlende Werte werden unzureichend behandelt

Diese Defizite sind nicht trivial, sondern können zu falschen Schlussfolgerungen führen. In einem der untersuchten Datensätze führte die Nichtbeachtung von Ausreißern dazu, dass ein signifikanter Unterschied nicht erkannt wurde.

Damit zeigt sich eine zentrale Grenze:

KI kann zwar Verfahren anwenden, aber nicht zuverlässig deren Voraussetzungen prüfen und bewerten.

4. Inkonsistenz und mangelnde Zuverlässigkeit

Ein besonders kritischer Befund ist die Instabilität der Ergebnisse.

Wiederholte Analysen desselben Datensatzes führten teilweise zu unterschiedlichen Vorgehensweisen und Ergebnissen.

- In einigen Fällen prüfte die KI die Voraussetzungen korrekt und wählte alternative Verfahren (z. B. Welch-ANOVA oder Kruskal-Wallis-Test), in anderen Fällen unterließ sie dies.
- Diese Inkonsistenz ist aus wissenschaftlicher Perspektive problematisch, da sie die Reproduzierbarkeit und Verlässlichkeit der Ergebnisse infrage stellt.

Für unerfahrene Nutzer entsteht hier ein erhebliches Risiko:

Sie sind nicht in der Lage zu erkennen, wann die KI korrekt arbeitet und wann nicht.

5. Grenzen generativer KI: Kein konzeptionelles Verständnis

Die im Beitrag diskutierten theoretischen Überlegungen zur Arbeitsteilung zwischen Mensch und Maschine liefern eine wichtige Erklärung für diese Defizite.

Generative KI besitzt kein echtes konzeptionelles Verständnis, sondern basiert auf der Kombination vorhandenen Wissens.

Das hat mehrere Konsequenzen:

- KI erkennt Probleme nur, wenn sie explizit danach „gefragt“ wird
- implizite Annahmen oder subtile Datenprobleme bleiben oft unentdeckt
- komplexe oder schlecht strukturierte Probleme überfordern die Systeme

Zudem besteht die Gefahr sogenannter „Halluzinationen“,

also plausibel wirkender, aber falscher Aussagen.

Diese Grenzen sind nicht nur technische Details, sondern betreffen den Kern wissenschaftlicher Arbeit:

- die Fähigkeit zur kritischen Prüfung und zum methodischen Urteil.

6. Demokratisierung vs. Qualitätsverlust

Die Möglichkeit, dass auch Personen ohne Statistikkennntnisse Analysen durchführen können, stellt eine der größten Chancen generativer KI dar.

- Sie eröffnet neue Zugänge zu Daten und kann Entscheidungsprozesse in Unternehmen und Organisationen unterstützen.

Gleichzeitig entsteht jedoch ein Spannungsfeld:

- Demokratisierung: Mehr Menschen können Daten analysieren
- Qualitätsrisiko: Analysen werden ohne ausreichendes Verständnis durchgeführt

Der Beitrag zeigt deutlich, dass diese beiden Aspekte untrennbar miteinander verbunden sind.

Technische Zugänglichkeit ersetzt nicht methodische Kompetenz.

7. Konsequenzen für die Statistikausbildung

Eine der wichtigsten Schlussfolgerungen betrifft die Neuausrichtung der Ausbildung.

Der Fokus auf die Durchführung von Analysen wird durch generative KI obsolet.

Stattdessen sollte die Ausbildung stärker auf folgende Aspekte ausgerichtet werden:

- Verständnis statistischer Konzepte
- Interpretation von Ergebnissen
- kritische Bewertung von Analysen
- Formulierung geeigneter Fragestellungen

Verschiebung vom „Wie rechne ich?“ hin zum „Was bedeutet das Ergebnis?“.

Studierende müssen lernen, KI-Ergebnisse zu überprüfen und zu hinterfragen.

8. Veränderung der Rolle von Software und Programmierung

Ein weiterer zentraler Punkt ist die veränderte Rolle von Programmiersprachen.

Da KI Code generieren kann, verliert das Erlernen spezifischer Syntax an Bedeutung.

Dies bedeutet jedoch nicht, dass technische Kompetenzen irrelevant werden. Vielmehr verschiebt sich der Fokus:

- von der manuellen Programmierung
- hin zum Verständnis, was der Code tut und ob er korrekt ist

Programmierung wird damit stärker zu einer interpretativen und kontrollierenden Tätigkeit.

9. Fokus auf Datenverständnis und kritischem Denken

Der Einsatz von KI eröffnet auch neue didaktische Möglichkeiten.

Datenverständnis

Wenn weniger Zeit für technische Aspekte benötigt wird, kann der Unterricht stärker auf das Verständnis von Daten ausgerichtet werden.

Dies umfasst insbesondere:

- Exploration von Daten
- Interpretation von Visualisierungen
- Diskussion von Unsicherheit und Modellannahmen

Dies steht im Einklang mit zentralen Bildungszielen zur Förderung statistischer Kompetenz.

Notwendigkeit kritischen Denkens

Ein wiederkehrendes Motiv im Beitrag ist die Bedeutung kritischen Denkens.

KI kann Analysen durchführen, aber nicht deren Qualität garantieren.

Daher wird kritisches Denken zur Schlüsselkompetenz:

- Erkennen von Fehlern
- Hinterfragen von Ergebnissen
- Bewertung von Annahmen

Ohne diese Fähigkeiten besteht die Gefahr, KI-Ergebnisse unreflektiert zu übernehmen.

10. Schlussfolgerungen

Generative KI wird die statistische Datenanalyse nicht ersetzen, sondern transformieren

- KI erleichtert die Durchführung standardisierter Analysen erheblich
- Sie ermöglicht neuen Nutzergruppen den Zugang zur Datenanalyse
- Gleichzeitig bleibt sie fehleranfällig und inkonsistent
- Statistisches Wissen bleibt unverzichtbar
- Die Rolle des Menschen verschiebt sich von der Durchführung zur Kontrolle

Diese Transformation entspricht einer allgemeinen Entwicklung in vielen Wissensbereichen: Automatisierung von Routinetätigkeiten bei gleichzeitiger Aufwertung anspruchsvoller kognitiver Kompetenzen.

Generative KI stellt einen tiefgreifenden Wandel in der statistischen Datenanalyse dar

Sie senkt Zugangshürden, erhöht Effizienz und eröffnet neue didaktische Möglichkeiten. Gleichzeitig bringt sie erhebliche Risiken mit sich, insbesondere im Hinblick auf Qualität & Verlässlichkeit von Analysen.

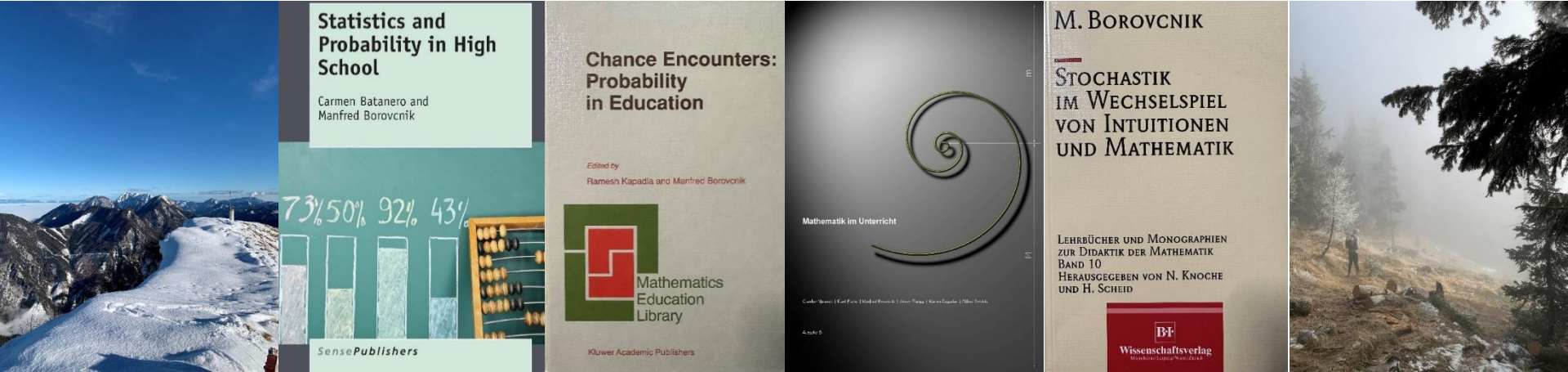
KI macht statistische Kompetenz nicht überflüssig, sondern im Gegenteil:

Sie macht diese Kompetenz noch wichtiger. Denn nur wer die zugrunde liegenden Konzepte versteht, kann die Ergebnisse der KI sinnvoll nutzen und bewerten.

Für die Zukunft der Statistik bedeutet dies eine klare Verschiebung:

Weg von der reinen Durchführung von Analysen hin zu deren kritischer Reflexion.

In dieser neuen Arbeitsteilung bleibt der Mensch unverzichtbar – nicht als Rechner, sondern als denkende, urteilsfähige Instanz.



Danke für Ihre Aufmerksamkeit!

manfred.Borovcnik@aau.at

www.researchgate.net/profile/Manfred-Borovcnik

