# Start:Bayes!

Abschluss

Sebastian Sauer





# Lernziele dieses Kapitels

Sie können …

- eine typische, sozialwissenschaftliche Forschungsfrage (quantitativ) untersuchen
- typische “Lieblingsfehler” benennen und vermeiden
- zwischen Frequentistischer Statistik und Bayes-Statistik übersetzen
- die Grundideen der Bayes-Statistik einordnen

# Begleitmaterial

- Fragestunde QM2
- Playlist QM2
- Lieblingsfehler

# Probeklausur



# Probeklausuren

- Aufgabensammlungen für 2022, 2023, 2024 auf dem Datenwerk
- Kategorien: `qm2-pruefung`, `exam-22`
- Liste ist im Aufbau — zusätzlich selbständig alle relevanten Kapitel-Aufgaben konsultieren

# Lieblinglingsfehler



# Lieblingsfehler im Überblick 🤷

1. Quantile und Verteilungsfunktion verwechseln
2. Prädiktoren nicht zentrieren, wenn ein Interaktionsterm vorliegt
3. Interaktion falsch interpretieren
4. Regressionskoeffizienten kausal interpretieren, ohne kausale Fundierung
5. Post-Prädiktiv-Verteilung (PPV) und Post-Verteilung verwechseln

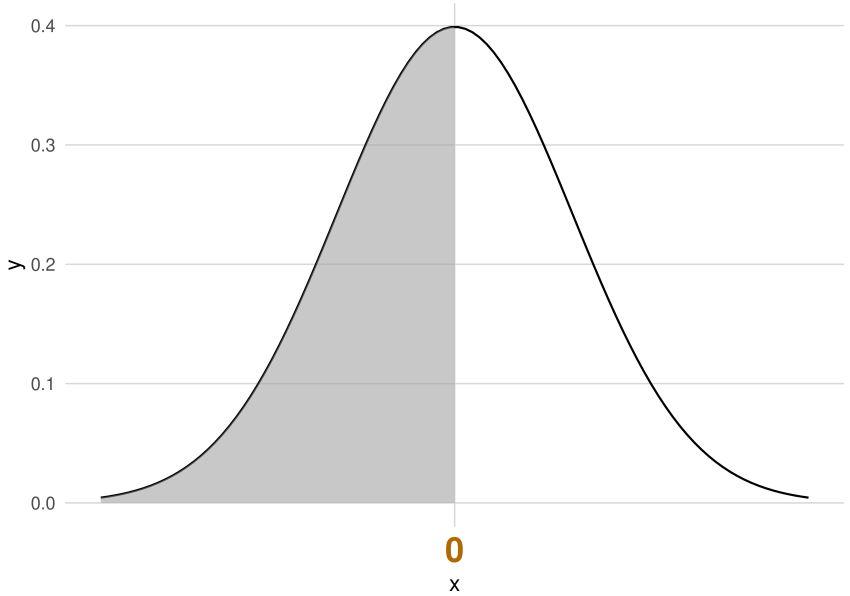
# Fehler: PPV vs. Post-Verteilung 🤷

- **Post-Verteilung**: Stichproben zu den *Parameterwerten* des Modells
- **Posterior-Prädiktiv-Verteilung (PPV)**: Vorhersagen für *Beobachtungen*, keine Parameterwerte
- Modell `mpg ~ hp`: `parameters(m1)` liefert i. d. R. bequemer Post-Intervalle/Punktschätzer als die rohe Stichprobentabelle

# Fehler: Quantil vs. Verteilungsfunktion 🤷

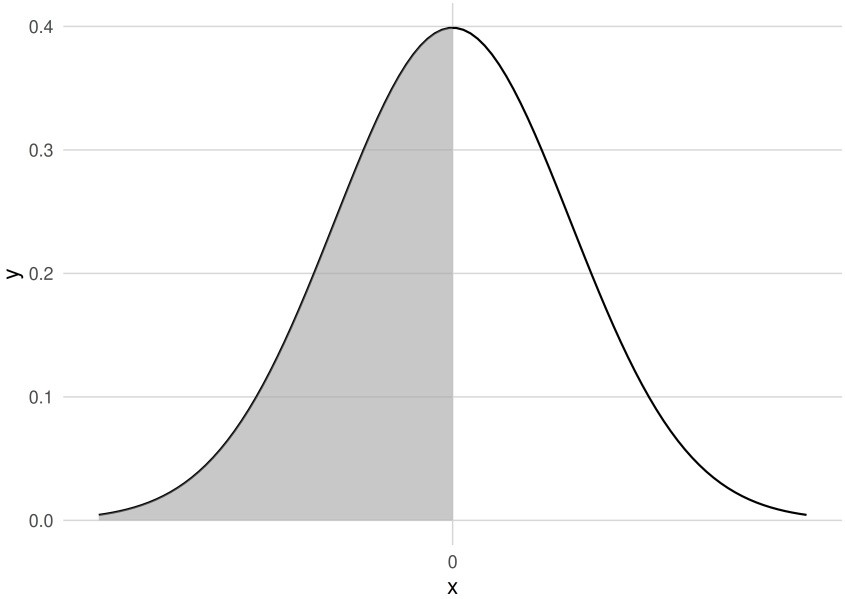
### p-Quantil

Teilt eine Verteilung so, dass mind. Anteil $p$ kleiner-gleich dem Quantil ist.



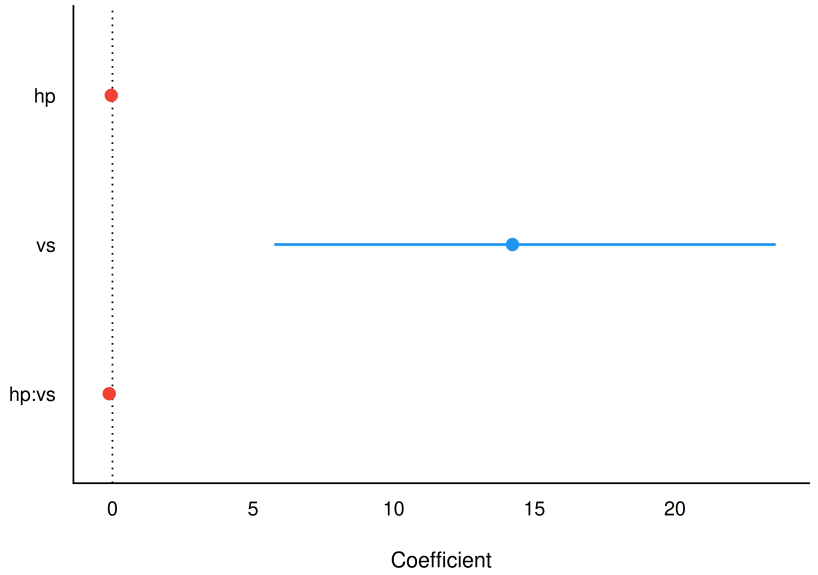

### Verteilungsfunktion F(x)

$F(x)$: Wahrscheinlichkeit, dass $X \leq x$.



# Fehler: Interaktion falsch interpretieren 🤷

Modell: `mpg ~ hp*vs`




**Falsch** 😈: Unterschied zwischen `vs=0`/`vs=1` beträgt ca. −0.11.

**Richtig** 👼: Unterschied beträgt ca. −0.11 — *nur wenn* `hp=0`. Da `hp=0` unrealistisch ist, sind zentrierte Prädiktoren sinnvoll (Gelman et al., 2021, Kapitel 10; McElreath, 2020, Kapitel 8).

# Kochrezepte 

# Kochrezept: Forschungsfrage untersuchen

**Theoretische Phase**

1. Staunen über ein Phänomen $y$, Kausalfrage finden
2. Literatur wälzen zu möglichen Ursachen $x$
3. Forschungsfrage/Hypothese präzisieren
4. Modell präzisieren (DAG(s), Prioris)

**Empirische Phase**

5. Versuch planen
6. Daten erheben

**Analytische Phase**

7. Daten aufbereiten
8. Modell berechnen
9. Modell prüfen/kritisieren
10. Forschungsfrage beantworten

# Parameter schätzen vs. Hypothesen prüfen

- **Hypothesenprüfung**: “Frauen parken im Schnitt schneller ein als Männer.”
- **Parameterschätzung**: “Wie groß ist der mittlere Unterschied in der Einparkzeit?”

Je ausgereifter ein Forschungsfeld, desto **kühnere** Hypothesen sind möglich — kühne Hypothesen sind wünschenswert 

# Formalisierung von Forschungsfragen

Mittelwert in Gruppe A höher als in Gruppe B:

$$\mu_1 > \mu_2 \Leftrightarrow \mu_1 - \mu_2 > 0 \Leftrightarrow \mu_d > 0$$

# Kerngedanken Bayes

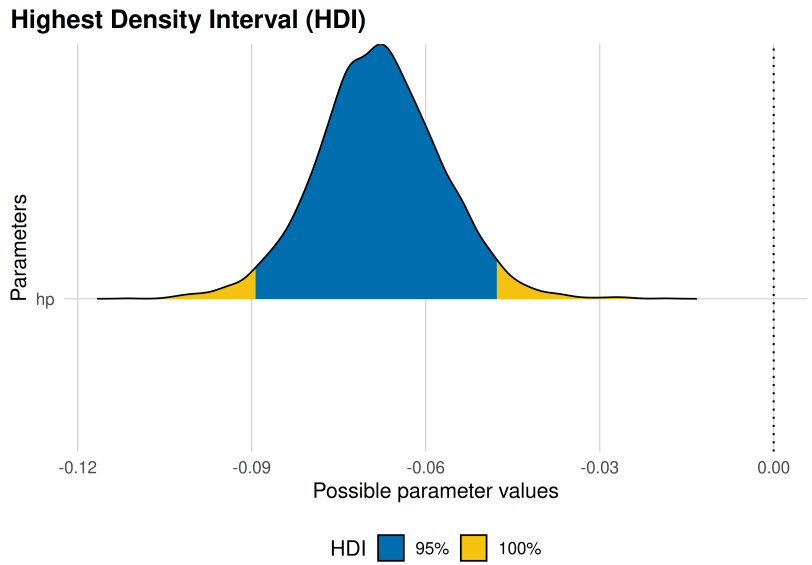


# Videos & zentraler Kennwert

📺 Bayes in fünf Minuten

📺 Bayes in zehn Minuten

Zentraler Kennwert: die **Post-Verteilung**




Zusammenfassung (z. B. 95 %-PI) ist oft sinnvoll — aber nur die *gesamte* Post-Verteilung trägt alle Information.

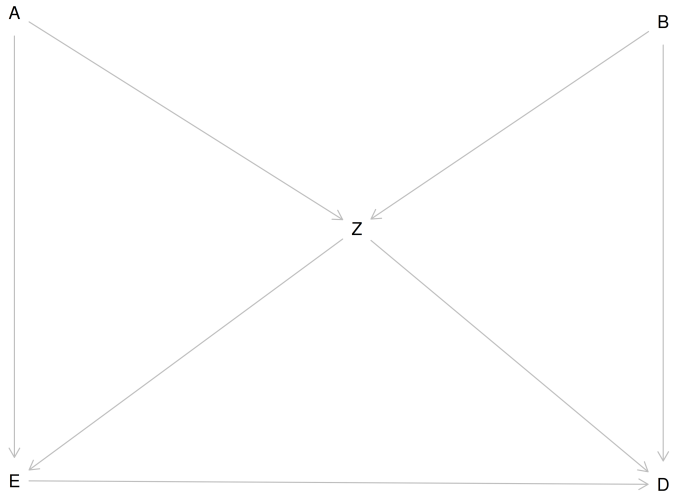
# Bayes-Formel

$$\text{Posteriori} = \frac{\text{Likelihood} \times \text{Priori}}{\text{Evidenz}}$$

# Beispiele für Prüfungsaufgaben

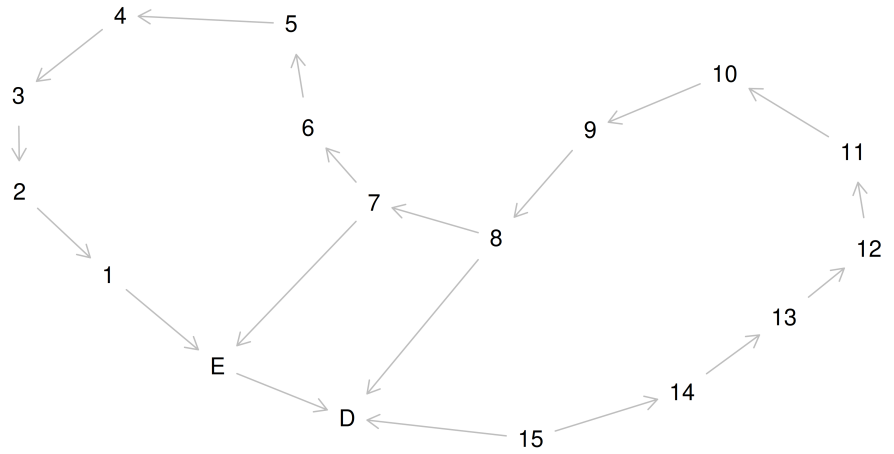
# Prüfungsaufgaben: DAGs analysieren

Minimale Adjustierungsmenge: kleinstmögliche Knotenmenge eines DAGs, die zu adjustieren ist, um einen kausalen Effekt zu identifizieren.




Adjustierungsmenge für den totalen Effekt von E auf D: {A,Z} **oder** {B,Z}.

# Prüfungsaufgaben: auch komplexe DAGs

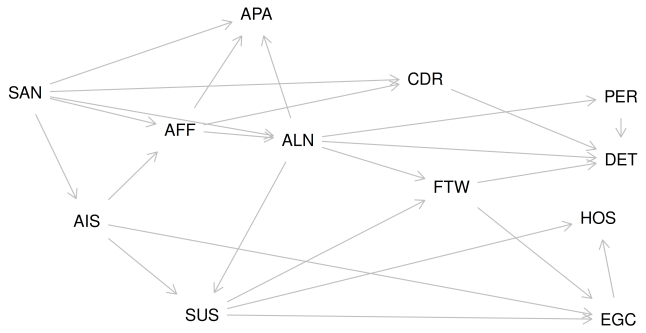



Trotz vieler Variablen: Adjustierungsmenge für E → D bleibt klein — hier {7} oder {8}.

# Beispiel: Kausalmodell der Schizophrenie

van Kampen (2014): SSQ-Modell des schizophrenen Prodromalverlaufs (van Kampen, 2014)

UV: *SUS* (Suspiciousness), AV: *EGC* (Egocentrism)




Minimales Adjustment-Set für den totalen Kausaleffekt: {AIS, ALN}

# Prüfungsfragen zu Modellen — Beispiele

- Punktschätzer (Median) für Prädiktor X in Modell Y?
- 89 %-HDI für Parameter X?
- $R^2$ angeben
- Interaktionsmodell formulieren
- Welches Modell modelliert den kausalen Effekt korrekt?
- Modell mit gegebenen Prioris formulieren
- Liegt Effekt X im ROPE?
- CI- vs. HDI-Breite für Prädiktor X vergleichen
- Was ändert sich bei Zentrierung/Z-Standardisierung der Prädiktoren?

# Aufgabensammlungen



# Relevante Tags im Datenwerk

Relevante Tags auf dem Datenwerk:

- bayes, bayes-grid
- dag, qm2
- probability, post, rope

Besondere Prüfungsnähe: `qm2-pruefung`, `exam-22`, `quiz1-qm2-ws23`, `Verteilungen-Quiz`





# Viel Erfolg bei der Prüfung!

# Wir drücken die Daumen!



Viel Erfolg bei der Prüfung!

# Vielen Dank!



Fragen?

Gelman, A., Hill, J., & Vehtari, A. (2021). *Regression and Other Stories*. Cambridge University Press.

McElreath, R. (2020). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan* (2. Aufl.). Taylor and Francis, CRC Press.

van Kampen, D. (2014). The SSQ Model of Schizophrenic Prodromal Unfolding Revised: An Analysis of Its Causal Chains Based on the Language of Directed Graphs. *European Psychiatry*, *29*(7), 437–448. https://doi.org/10.1016/j.eurpsy.2013.11.001