

Block 3: Multiple-Choice-Fragen

Komplexität, Kausalität & Generalisierung

January 11, 2026

Ein VIF-Wert von 12 bedeutet:

- A) Keine Probleme mit Multikollinearität
- B) Leichte Multikollinearität
- C) Schwere Multikollinearität
- D) Perfekte Kollinearität

Ein VIF-Wert von 12 bedeutet:

- A) Keine Probleme mit Multikollinearität
- B) Leichte Multikollinearität
- C) Schwere Multikollinearität
- D) Perfekte Kollinearität

Lösung: C)

VIF > 10 gilt als schwer problematisch. Die Variable teilt $>90\%$ ihrer Varianz mit anderen Prädiktoren.

Ein Modell mit hohem Bias und niedriger Varianz:

- A) Ist zu komplex (Overfitting)
- B) Ist zu einfach (Underfitting)
- C) Ist optimal
- D) Hat zu wenig Daten

Ein Modell mit hohem Bias und niedriger Varianz:

- A) Ist zu komplex (Overfitting)
- B) Ist zu einfach (Underfitting)
- C) Ist optimal
- D) Hat zu wenig Daten

Lösung: B)

Hoher Bias = systematischer Fehler = Modell zu einfach = Underfitting. Das Modell erfasst wichtige Zusammenhänge nicht.

Was unterscheidet Lasso von Ridge Regression?

- A) Lasso ist schneller
- B) Lasso kann Koeffizienten auf genau 0 setzen
- C) Ridge hat hoesheres R^2
- D) Es gibt keinen Unterschied

Was unterscheidet Lasso von Ridge Regression?

- A) Lasso ist schneller
- B) Lasso kann Koeffizienten auf genau 0 setzen
- C) Ridge hat hoheres R^2
- D) Es gibt keinen Unterschied

Lösung: B)

Lasso (L1-Strafe) kann Koeffizienten auf exakt 0 setzen und führt damit automatische Variablenselektion durch. Ridge schrumpft nur, entfernt nie.

Bei 10-Fold Kreuzvalidierung, wie viel Prozent der Daten sind in jedem Fold Testdaten?

- A) 5%
- B) 10%
- C) 50%
- D) 90%

Bei 10-Fold Kreuzvalidierung, wie viel Prozent der Daten sind in jedem Fold Testdaten?

- A) 5%
- B) 10%
- C) 50%
- D) 90%

Lösung: B)

Bei K-Fold CV ist $1/K$ der Daten Testset. Bei 10-Fold sind das 10%. Die restlichen 90% sind Trainingsset.

Eisverkauf und Ertrinkungsfaelle korrelieren positiv. Was ist die wahrscheinlichste Erklaerung?

- A) Eis essen verursacht Ertrinken
- B) Ertrinken verursacht Eiskauf
- C) Beide werden durch Temperatur (Confounder) beeinflusst
- D) Es ist reiner Zufall

Eisverkauf und Ertrinkungsfaelle korrelieren positiv. Was ist die wahrscheinlichste Erklaerung?

- A) Eis essen verursacht Ertrinken
- B) Ertrinken verursacht Eiskauf
- C) Beide werden durch Temperatur (Confounder) beeinflusst
- D) Es ist reiner Zufall

Lösung: C)

Hohe Temperatur führt zu mehr Eisverkauf UND mehr Schwimmen (und damit Ertrinkungsfaellen). Temperatur ist der Confounder.

Warum funktioniert Randomisierung in A/B-Tests?

- A) Sie macht die Stichprobe größer
- B) Sie balanciert alle Confounders zwischen den Gruppen
- C) Sie reduziert die Varianz
- D) Sie erhöht die Power

Warum funktioniert Randomisierung in A/B-Tests?

- A) Sie macht die Stichprobe größer
- B) Sie balanciert alle Confounders zwischen den Gruppen
- C) Sie reduziert die Varianz
- D) Sie erhöht die Power

Lösung: B)

Durch zufällige Zuteilung sind alle bekannten und unbekanntes Stoervariablen im Erwartungswert zwischen den Gruppen ausgeglichen.

Ein A/B-Test zeigt $d = 0.3$ mit $p < 0.001$ bei $n = 100'000$. Was bedeutet das?

- A) Grosser und signifikanter Effekt
- B) Kleiner aber hochsignifikanter Effekt
- C) Nicht signifikant, kleiner Effekt
- D) Grosser aber nicht signifikanter Effekt

Ein A/B-Test zeigt $d = 0.3$ mit $p < 0.001$ bei $n = 100'000$. Was bedeutet das?

- A) Grosser und signifikanter Effekt
- B) Kleiner aber hochsignifikanter Effekt
- C) Nicht signifikant, kleiner Effekt
- D) Grosser aber nicht signifikanter Effekt

Lösung: B)

$d = 0.3$ ist ein kleiner Effekt (unter 0.5). Bei sehr grossen Stichproben wird aber fast alles signifikant – daher $p < 0.001$ trotz kleinem Effekt.

Warum ist OLS für binaere Outcomes ungeeignet?

- A) OLS ist zu langsam
- B) OLS kann Werte ausserhalb $[0,1]$ vorhersagen
- C) OLS braucht mehr Daten
- D) OLS kann keine Koeffizienten berechnen

Warum ist OLS für binaere Outcomes ungeeignet?

- A) OLS ist zu langsam
- B) OLS kann Werte ausserhalb $[0,1]$ vorhersagen
- C) OLS braucht mehr Daten
- D) OLS kann keine Koeffizienten berechnen

Lösung: B)

Bei binären Outcomes modellieren wir $P(Y=1)$. OLS kann Werte < 0 oder > 1 vorhersagen, was als Wahrscheinlichkeit unsinnig ist.

Ein Logit-Koeffizient von 0.7 entspricht einem Odds Ratio von:

- A) 0.7
- B) 1.7
- C) 2.01
- D) 7.0

Ein Logit-Koeffizient von 0.7 entspricht einem Odds Ratio von:

- A) 0.7
- B) 1.7
- C) 2.01
- D) 7.0

Lösung: C)

Odds Ratio = $e^{\beta} = e^{0.7} \approx 2.01$. Die Odds verdoppeln sich bei einer Einheit Erhöhung des Prädiktors.

Welche Metrik sollte primär zur Modellwahl verwendet werden?

- A) Training R^2
- B) Anzahl signifikanter Koeffizienten
- C) Cross-Validation Fehler
- D) Modellkomplexität

Welche Metrik sollte primär zur Modellwahl verwendet werden?

- A) Training R^2
- B) Anzahl signifikanter Koeffizienten
- C) Cross-Validation Fehler
- D) Modellkomplexität

Lösung: C)

Der CV-Fehler schätzt die Out-of-Sample Performance – das ist entscheidend für Generalisierung. Training R^2 ist optimistisch.