

Woche 4

Benjamin Schlegel

1 R^2 , angepasstes R^2 und RMSE

Material

- Video: <https://youtu.be/npE0TkcVj8w>
- *Regression Analysis: A Primer for the Social Sciences*, Kapitel 6.1

Übung

1. Was ist der Vor-/Nachteil des angepassten R^2 gegenüber dem normalen R^2 ?
 - Das angepasste R^2 kann nur Werte zwischen 0 und 1 annehmen, was die Interpretation vereinfacht.
 - Das angepasste R^2 bestraft Prädiktoren, welche kaum zur Erklärung der abhängigen Variable beitragen.
 - Das normale R^2 schützt besser davor, nicht zu viele Variablen ins Modell aufzunehmen (Overfitting).
 - Das angepasste R^2 kann auch verwendet werden, wenn sich der Stichprobenumfang ändert.
 - Das angepasste R^2 kann auch verwendet werden, wenn die abhängige Variable transformiert wird.
2. Schätze das Modell von letzter Woche ($acceptpo_i = \beta_0 + \beta_1 \cdot suppshare2_i + \beta_2 \cdot swjalager_i + \beta_3 \cdot nationalrat_i + \beta_4 \cdot srf2ja_i + \epsilon_{i.}$). Welchen Anteil der Varianz der abhängigen Variablen kann das Modell erklären? Wie gross ist das angepasste R^2 .
3. Schätze ein 2. Modell, diesmal ohne die Variable **suppshare2**. Wie stark verändert sich das R^2 und wie stark das angepasst R^2 ? Welches der beiden Modelle hat den besseren RMSE Wert?

Lernziele

- Du hast ein gutes Verständnis des R^2 , kannst es interpretieren und die Formel herleiten anhand grafischer Überlegungen. Ebenso weist du, was die Begriffe SSR, SSE und SST bedeuten.
- Du kennst die Vorteile des angepassten R^2 (\bar{R}^2) gegenüber dem normalen R^2 und weisst, dass es von den Anzahl Prädiktoren abhängt.
- Du kannst Modelle mit Hilfe von R^2 , \bar{R}^2 und RMSE vergleichen.

2 genestete/verschachtelte Modell

Material

- Video: <https://youtu.be/dyb2aoTjHMw>
- *Regression Analysis: A Primer for the Social Sciences*, Kapitel 6.2

Übungen

4. Wir haben folgende Modelle:

- A: $y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \epsilon_i$
- B: $y_i = \beta_0 + \beta_1 x_2 + \beta_2 x_3 + \epsilon_i$
- C: $y_i = \beta_0 + \beta_1 x_2 + \beta_2 x_3 + \beta_3 x_4 + \epsilon_i$
- D: $y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_3 + \epsilon_i$

Welche der folgenden Aussagen sind richtig und welche falsch?

Aussage	richtig	falsch
Modell A ist in Modell B genestet.	<input type="checkbox"/>	<input type="checkbox"/>
Modell B ist in Modell A und Modell C genestet.	<input type="checkbox"/>	<input type="checkbox"/>
Modell D ist in Modell A, aber nicht in Modell C genestet.	<input type="checkbox"/>	<input type="checkbox"/>
Modell A ist das Vollmodell von Modell C.	<input type="checkbox"/>	<input type="checkbox"/>

Lernziele

- Du kennst den Unterschied zwischen verschachtelten (nested) und nicht verschachtelten Modellen.

3 Informationskriterium

Material



- Video: <https://youtu.be/kqtYaqFaptQ>
- *Regression Analysis: A Primer for the Social Sciences*, Kapitel 6.3 und 6.4

Übung

5. Vergleiche die Modelle aus Aufgabe 2 und 3 mit Hilfe von AIC oder BIC. Welches Modell ist besser?

Lernziele

- Du kennst AIC, BIC und AICc und kannst damit Modelle vergleichen.
- Du kannst AIC und BIC ohne Nachschlagen in R mit den Funktionen `AIC()` und `BIC()` berechnen.
- Du weisst, wo man `AICc()` nachschlagen kann, um es in R zu berechnen.