

Multivariate Regression

M10: Empirische Wirtschaftsforschung

Präsenzveranstaltung 3

Agenda

- Fragen zu E-Learning Materialien
- Besprechung Hausarbeit 2
- Annahmen der Multivariaten Regression
- Partielle oder Multivariate Regression
- Interpretation der Koeffizienten
- Hypothesen-Tests
- Nachbesprechung von Verständnistests
- Vorbesprechung Hausaufgabe

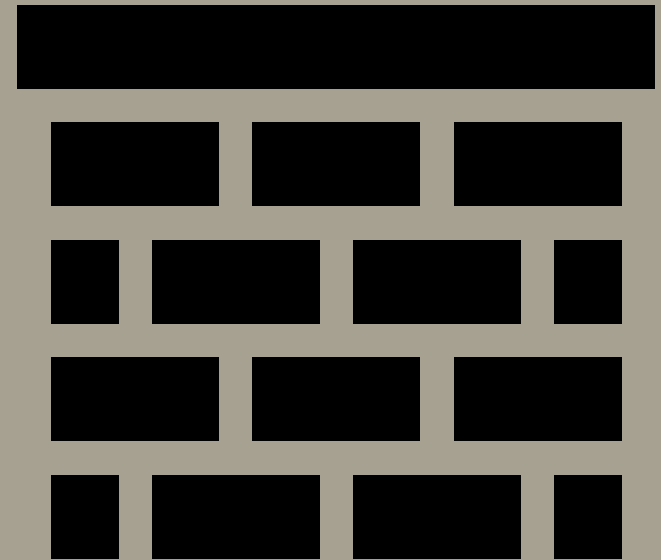
Fragen zu E-Learning Materialien



Besprechung Hausarbeit 2



Annahmen der Multivariaten Regression



Annahmen der Multivariaten Regression

- MLR.1 Linearität in Parametern
- MLR.2 Zufallsstichprobe
- MLR.3 Varianz und keine perfekte Multikollinearität
- MLR.4 Exogene Regressoren

MLR.4 Exogene Regressoren

- Es gibt keine unbeobachteten Faktoren, welche sowohl die abhängige als auch die unabhängige Variable beeinflussen.
- Die Annahme ist nicht testbar und muss durch ökonomische Argumente plausibilisiert werden.
- Gängige Verletzungen dieser Annahme durch 'Omitted Variable' oder 'Reverse Causality'
- Schätzungen können nicht mehr kausal interpretiert werden, nur noch für Vorhersagen verwendet werden.

Boston Housing Data

Aggregierte Gebäudedaten für 506 Stadtteile von Boston mit folgenden 14 Variablen:

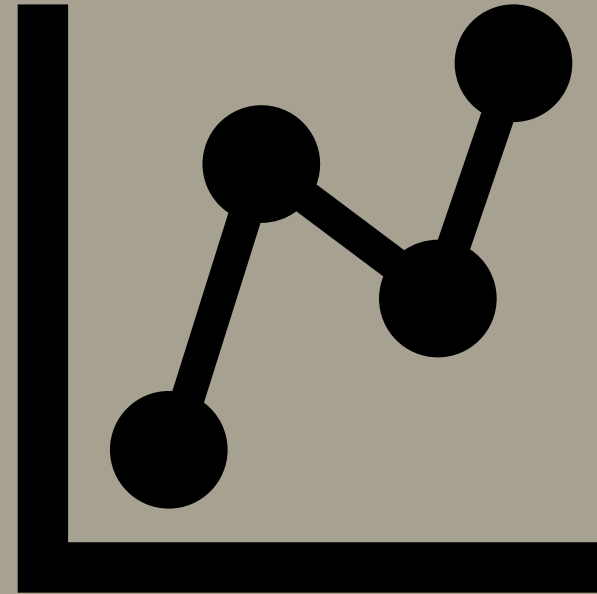
- medv: Durchschnittlicher Wert eines Wohnhauses
- crim: Kriminalitätsrate
- zn: Flächenanteil von Wohnzone
- indus: Flächenanteil von Industriezone
- chas: Gebiet liegt am Charles River
- nox: Stickstoffoxid Konzentration der Luft
- rm: Durchschnittliche Anzahl der Zimmer in Gebäude
- age: Anteil der Gebäude älter als 1940
- dis: Distanz vom Stadtzentrum
- rad: Index für Verkehrsanbindung
- tax: Steuersatz für Immobilien
- ptratio: Lehrer/Schüler Verhältnis
- b: Index für Anteil der schwarzen Bevölkerung
- lstat: Anteil an SOB-schwacher Bevölkerung

Aufgabe: MLR-Annahmen

Sie verwenden 'Boston Housing Data' für ihre Regression. Entscheiden Sie sich für ein geeignetes Modell für eine kausale Schätzung und diskutieren Sie die Plausibilität der Exogenitäts-Annahme.

- Breakout rooms
 - Raum 1: Sie sind interessiert am Effekt der **Kriminalitätsrate** (crim) auf die **Wohnpreise** (medv)
 - Raum 2: Sie sind interessiert am Effekt der **Stickstoffoxid Konzentration** (nox) auf die **Wohnpreise** (medv)

Partielle oder Multivariate Regression



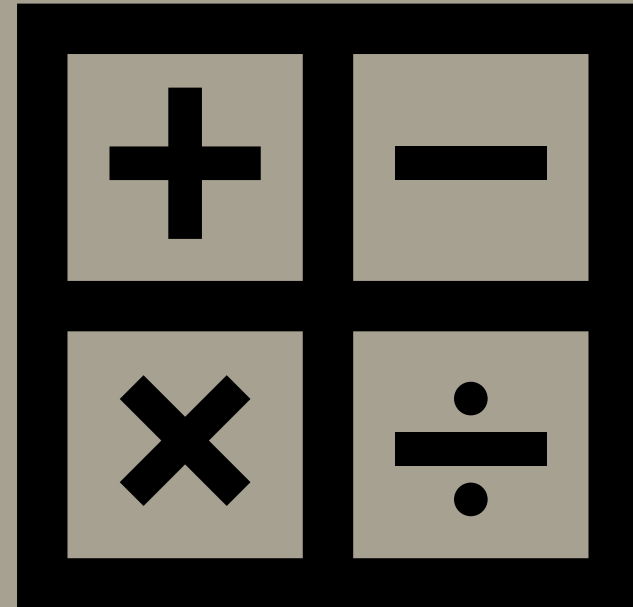
Partielle oder Multivariate Regression

- Was ist der Unterschied zwischen einer partiellen Regression und der multivariaten Regression?
- Wann liefern beide die äquivalenten Resultate?

R Aufgabe: Partielle Regression

- Öffnen Sie das Jupyter Notebook zu der Präsenzveranstaltung 3 auf dem Server:
http://54.235.20.131/user/<Ihre_Email_Adresse>
- Wir lösen das Modul ‘Partielle Regression’ gemeinsam.

Interpretation der Koeffizienten



R-Aufgabe: Interpretation

- Öffnen Sie das Jupyter Notebook zu der Präsenzveranstaltung 3 auf dem Server:
http://54.235.20.131/user/<Ihre_Email_Adresse>
- Wir lösen das Modul ‘Interpretation der Koeffizienten’ gemeinsam.

R-Aufgabe: Interpretation

- Schätzen Sie die Koeffizienten mittels partieller Regression, sowie multivariater Regression. Verwenden Sie dazu das Modell welches sie zuvor bestimmt haben. Sind die Resultate identisch?
- Vergleichen Sie die Resultate mit den Resultaten des maximalen Modells.
- Breakout rooms:
 - Raum 1: Sie sind interessiert am Effekt der **Kriminalitätsrate** (crim) auf die **Wohnpreise** (medv)
 - Raum 2: Sie sind interessiert am Effekt der **Stickstoffoxid Konzentration** (nox) auf die **Wohnpreise** (medv)

Hypothesen- Tests



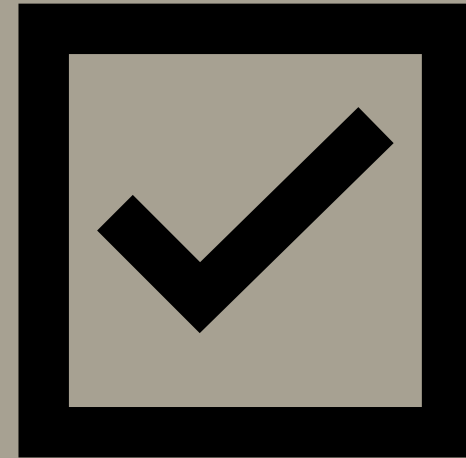
Hypothesen-Tests

- Was ist das Ziel von Hypothesen-Tests?
- Was ist der Unterschied zwischen dem t-Test und dem F-Test?
- Was sind die Vor-/Nachteile des Signifikanzniveaus?

R-Aufgabe: Hypothesen-Test

- Präsenzveranstaltung 3 auf dem Server:
http://54.235.20.131/user/<Ihre_Email_Adresse>
- Wir lösen das Modul ‘Hypothesen-Test’ gemeinsam.

Nachbesprechung von Verständnistests



Verständnistest

Nehmen Sie an, dass das wahre Modell

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u$$

ist. Es wird jedoch

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + u$$

geschätzt, wobei x_3 mit x_1 und x_2 unkorreliert ist.

Was ist der Effekt der Einbeziehung einer irrelevanten Variable? Entscheiden sie für jede Aussage, ob diese wahr oder falsch ist.

Multiple Choice Fragen:

- 1) Die geschätzten OLS Koeffizienten $\hat{\beta}_0$ und $\hat{\beta}_1$ sind weder unverzerrt noch effizient.
- 2) Die geschätzten OLS Koeffizienten $\hat{\beta}_0$ und $\hat{\beta}_1$ sind unverzerrt und effizient.
- 3) Die geschätzten OLS Koeffizienten $\hat{\beta}_0$ und $\hat{\beta}_1$ sind unverzerrt aber wahrscheinlich nicht effizient.
- 4) Keine der anderen Antworten ist wahr

Vorbesprechung Hausaufgabe

