

Fortgeschrittenenmodul Regressionsanalyse

Einige Anwendungsgebiete der Regressionsanalyse:

- **Strukturtestende Anwendung** (Modellbildung)
- **Ursachenanalyse** (Wie stark ist der Einfluss von unabhängigen Variablen?)
- **Wirkungsprognose** (Um wie viele Einheiten ändert sich die abhängige Variable bei Änderung einer unabhängigen Variable um eine Einheit?)
- **Vorhersage** (aber: Vorsicht bei Extrapolation!)
- **Kontrolle** (z.B. Überwachung der Qualität in der Fertigung)

Inhalt des Kurses I:

- Streudiagramm der Beispieldaten und grafische Darstellung der Abstandszersetzung
- Gütemaß für die Anpassungsgüte (Fit) des Modells
- Einfache Regression – Beispielgleichung
- Einfache Regression – Basisergebnisausgabe
- Zur Modellbildung das Beispiel von *Anscombes Quartet*
- Einfache Regression: Überprüfung der Modellannahmen
- Einfache Regression – Diagnose 1. Schritt:
Prüfen auf Ausreißer, Hebelwerte und einflussreiche Punkte
(Methode: Influence-Plot)
- Multiple Regression – Beispielgleichung
- Multiple Regression – Basisergebnisausgabe
- Multiple Regression – Diagnose 1. Schritt:
Prüfen auf Ausreißer, Hebelwerte und einflussreiche Punkte
(Methoden: Influence-Plot, Partielle Regressionsdiagramme (Partial Regression Plots oder Added Variable Plots))
- Multiple Regression – Diagnose 1. Schritt:
Ausreißer, Hebelwerte und einflussreiche Punkte gefunden – Maßnahmen
- Multiple Regression – Diagnose 2. Schritt:
Prüfen auf Linearität → Transformation der abhängigen Variable
(Methoden: Bulging Rule, klassisch (Streudiagramm mit der abhängigen Variable auf der X-Achse und den studentisierten ausgeschlossenen Residuen auf der Y-Achse), Box-Cox-Transformation, Inverse Response Plot)

Inhalt des Kurses II:

- Multiple Regression – Diagnose 2. Schritt:
Nichtlinearität gefunden (abhängige Variable) – Maßnahmen
- Multiple Regression – Diagnose 2. Schritt:
Prüfen auf Linearität → Transformation der unabhängigen Variable(n)
(Methoden: Bulging Rule, Component+Residual Plots (auch Partial Residual Plots genannt), CERES Plots, Box-Tidwell-Methode)
- Multiple Regression – Diagnose 2. Schritt:
Nichtlinearität gefunden (unabhängige Variable(n)) – Maßnahmen
- Multiple Regression – Diagnose 3. Schritt:
Prüfen auf konstante Fehlervarianz – Homoskedastizität
(Methoden: Residuenplots, Glejser-Test, Goldfeld-Quandt-Test)
- Multiple Regression – Diagnose 3. Schritt:
Heteroskedastizität gefunden – Maßnahmen
- Multiple Regression – Diagnose 4. Schritt:
Prüfen auf Normalverteilung der Residuen
(Methoden: Normalverteilungsdiagramme (P-P-Diagramme), Test auf Normalverteilung)
- Multiple Regression – Diagnose 4. Schritt:
Nichtnormalverteilung gefunden – Maßnahmen
- Regression mit Dummy-Variablen
- Kollinearität und Modellbildung

SPSS-Dialoge zur Diagnose:

- Die Dialoge (daneben auch Literaturangaben und Tutorials) sind unter folgender URL zu finden: userpage.gruener.fu-berlin.de
- Einschlägige Dialoge sind:
RegBoxCoxTransforms, RegCompResPlots , RRegDiagGraph und RRegDiagTest