Signifikanz von Parametern: der t-Test

Dan Li

Gliederung

- Regressionsanalyse

  1.Defiinition

  2. Funktion

  3. Modelltypen

- Hypothesentest

  Durchführung der Hypothesentest

- t-Test

  1.Definition

  2.Signifikanz einzelner Regressoren mittels t-Tes

  3.Beziehung für den kritischen Bereich

  4.Durchführung

- Beispiel und Lösung
- Quellen

Definition

- Die **Regressionsanalyse** :

  -ist ein statistisches Analyseverfahren.

  -Ziel ist es, Beziehungen zwischen einer abhängigen und einer oder mehrerer unabhängigen Variablen festzustellen.

Regressionsfunktion

- Einfache lineare Regression:

$$\hat{Y} = b_0 + b_1 X$$

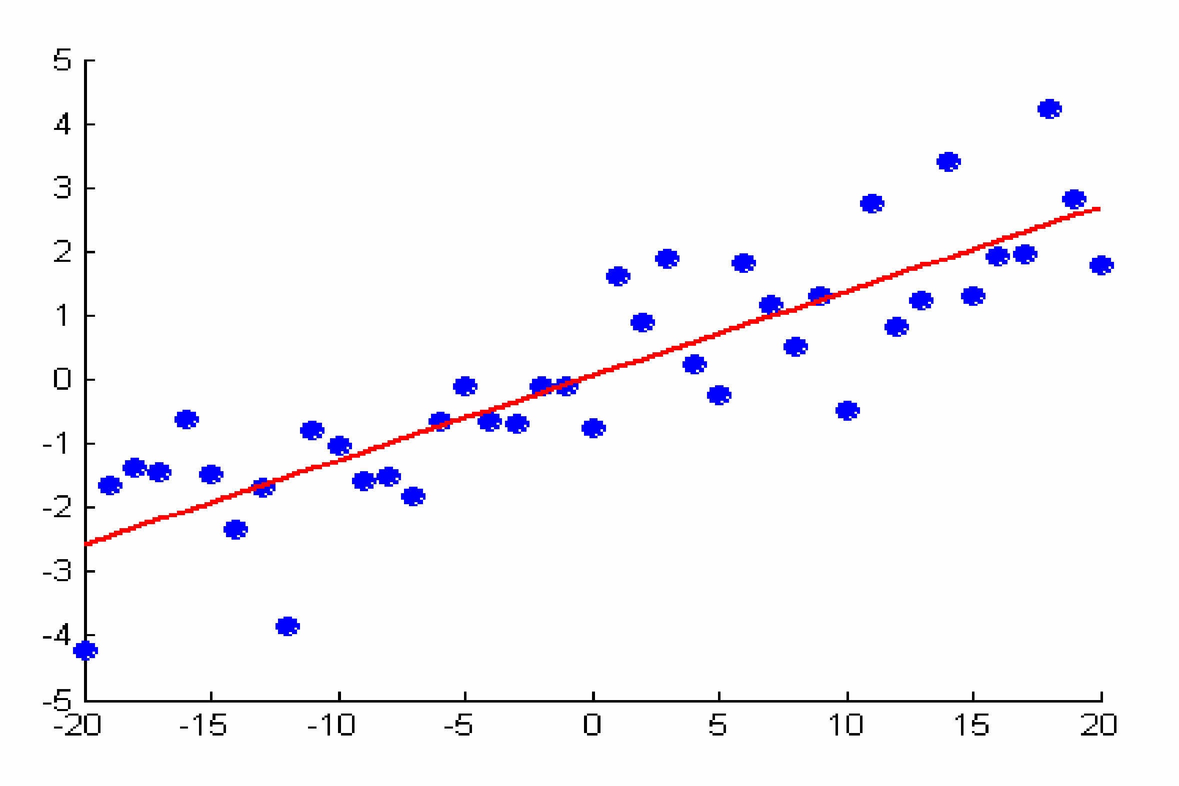
- Multiple lineare Regression:

$$\hat{Y} = b_0 + b_1 X + b_2 X_2 + \ldots + b_j X_j + b_J X_J$$

- $\hat{Y}$ = Schätzung der abhängigen Variablen Y
- b_0 = Konstantes Glied
- b_1, b_2 …b_j,b_J= Regressionskoeffizient
- X = unabhängige Variable

Modelltypen

- Einfache Regression mit linearem Zusammenhang:
  $Y_i = f(X_i) = a_0 + a_1 X_i + e_i$



Hypothesentest

- A_0: H_0 beibehalten, H_0 ist mit der Beobachtung vereinbar
- A_1: H_0 ablehnen, H_1 ist signifikant (im Sinne von „statistisch nachgewiesen“)

Wirklichkeit Testergebnis	**H_0 ist richtig**	**H_1 ist richtig**
H_0 beibehalten	Kein Fehler	Fehler 2.Art
H_0 ablehnen, H_1 signifikant	Fehler 1.Art (Wahrscheinlich-keit:α%)	Kein Fehler

Die Durchführung eines Hypothesentests : die 5 Schritte

- 1 Hypothesenaufstellung H_0 und H_1
- 2 Festlegung der Irrtumswahrscheinlichkeit
- 3 Bestimmung einer geeigneten Prüfgröße
- 4 Konstruktion des kritischen Bereichs
- 5 Vergleichung der berechneten Prüfgröße

  mit dem kritischen Bereich

Definition des t-Tests

- Der t-Test ist ein Hypothesentest
- Der t-Test ist auch ein typisches Beispiel dafür, dass es für bestimmte Hypothesentests durchaus sinnvoll sein kann, den Test als einseitige oder als zweiseitige Fragestellung zu formulieren.

Signifikanz einzelner Regressoren mittels t-Test

- T-Statistik

$$t_j = \frac{\hat{\beta}_j}{s\sqrt{x^{jj}}} = \frac{\hat{\beta}_j}{\hat{S}E}$$

$\hat{\beta}_j$ = Regressionskoeffizient des j-ten Regressors

$s\sqrt{x^{jj}}$ = Standardfehler von $\hat{\beta}_j$

$$\hat{\beta} = \frac{S_{xy}}{S_x^2} \qquad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

2

$$\hat{S}E(\hat{\beta}_j) = s\sqrt{x^{jj}}$$

Beziehung für den kritischen Bereich

- Man arbeitet bei der formulierten einseitigen Fragestellung also mit dem **α**-Quantil und nicht mit dem **α**/2-Quantil der t-Verteilung. Die Nullhypothese H_0 wird nur dann abgelehnt, wenn gilt: $t_j > t_{n-(k+1);1-\alpha}$ **.**
- Diese p Values unmittelbar mit den **α**-Wahrscheinlichkeiten verglichen werden, da die folgenden Äquivalenz-

  beziehungen gelten:

  p Value/2>**α**/2 $\leftrightarrow |t_j| < t_{n-(k+1);1-\alpha/2}$ **bzw.** $-|t_j| > t_{n-(k+1);\alpha/2}$

  → Beibehaltung H_0

  p Value/2<**α**/2 $\leftrightarrow |t_j| > t_{n-(k+1);1-\alpha/2}$ **bzw.** $-|t_j| < t_{n-(k+1);\alpha/2}$

  → Ablehnung H_0

Die Durchführung des t-Tests: Die 5 Arbeitsschritte

- 1 Formulierung der geeigneten Hypothesen
- 2 Festlegung der Irrtumswahrscheinlichkeit α
- 3 Festlegung einer geeigneten Prüfgröße
- 4 Konstruktion des kritischen Bereiches
- 5 Berechnung der entsprechenden Prüfgröße t und Entscheidung über H_0

Beispiel

- Im folgenden Beispiel sollen die jährlichen DAX-Veränderungen (ΔDAX) mit Hilfe der jährlichen Veränderungsraten des Bruttoinlandproduktes (ΔBIP) und der Exporte (ΔEXP) mittels eines linearen Regressions-modells erklärt werden.

t=1,...,4	ΔDAX	ΔBIP	ΔEXP
1 (1993-1994)	5	2	-1
2 (1994-1995)	-2	-1	3
3 (1995-1996)	3	0	1
4 (1996-1997)	7	1	2

- Die Regressionsgleichung lautet in diesem Falle:

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t \,, \qquad t = 1,\ldots,n$$

Lösung

$\hat{b} = \begin{bmatrix} 2 \\ 2{,}5 \end{bmatrix}$ $\hat{SE}(\hat{\beta}_0)$ **= 1.42** $\hat{SE}(\hat{\beta}_1)$ **= 1,162**

1. $H_0: \beta_0 = 0$ gegen $H_1: \beta_0 \neq 0$ bzw. $H_0: \beta_1 = 0$ gegen $H_1: \beta_1 \neq 0$

2. $\alpha = 0{,}05$, $t_j > t_{n-(k+1);1-\alpha/2}$ bzw. $t_j > t_{n-(k+1);\alpha/2}$

$t_{n-(k+1);1-\alpha/2} = t_{4-(1+1);0{,}975} = 4{,}303$

$t_j > t_{n-(k+1);\alpha/2} = t_{4-(1+1);0{,}025} = -4{,}303$

3. $t_0 = \frac{2}{1{,}423} = 1{,}405$ $t_1 = \frac{2{,}5}{1{,}162} = 2{,}151$

Die ermittelten Werte für beide Regressionsparameter sind weder größer als 4,303, noch kleiner als -4,303. Damit kann die Nullhypothese, dass die Regressionskonstante β_0 einen Wert von null besitzt, nicht verworfen werden. Die gleiche Aussage gilt für den Regressionskoeffizienten β_1.

Quellen

- Poddig Th.u.a: Statistik, Ökonometrie, Optimierung, 2.Aufl.,Bad Soden 2001
- Backhaus. Erichson, Plinke.Weiber: Multivariate Analysemethoden, 10.Aufl.,Springer
- http://de.wikipedia.org/wiki/Statistischer_Test
- David Schneider, Markus Kettern „ SeminarData Mining and Prediction“ WS 04/05, Institut für Informatik Freie Universität