Statistik – Einführung

Lineare Regression
Kapitel 9

Statistik – WU Wien

Gerhard Derflinger · Michael Hauser · Jörg Lenneis · Josef Leydold ·
Günter Tirler · Rosmarie Wakolbinger

Lernziele

1. Beschreiben lineares Regressionsmodell.
2. Geben die einzelnen Modellierungsschritte an.
3. Erklären die (gewöhnliche) Methode der kleinsten Quadrate, (gewöhnliche Kleinste-Quadrate-Methode).
4. Berechnen Regressionskoeffizienten.
5. Prognostizieren die „Antwort“-Variable.
   Sie wird auch als „Repsonse“-, abhängige, zu erklärende Variable, oder Regressand bezeichnet.
6. Interpretieren einen Computerausdruck.

Modelle

1. Die Repräsentation **einer Auswahl** von Phänomenen.
2. Ein quantitatives Modell ist eine formale (mathematische) Darstellung einer Auswahl von Phänomenen.
3. Es beschreibt oft Zusammenhänge zwischen Variablen.
4. Typen von Modellen:
   - Deterministische Modelle
   - Stochastische Modelle

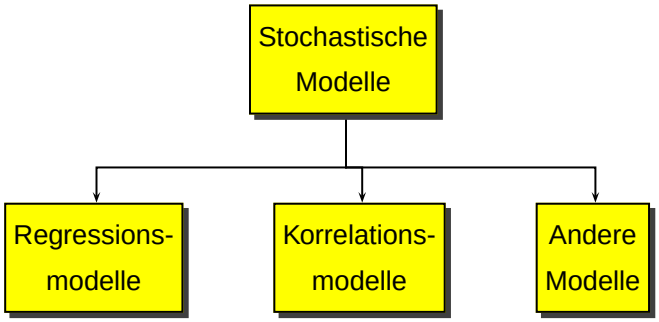
Deterministische Modelle

1. Formulieren Hypothesen über **exakte** Zusammenhänge.
2. Sind geeignet, wenn der Prognosefehler vernachlässigbar ist.

- Beispiele:
  - Kraft = Masse $\times$ Beschleunigung
    $F = m \cdot a$
  - Energie = Masse $\times$ Lichtgeschwindigkeit^2
    $E = m \cdot c^2$

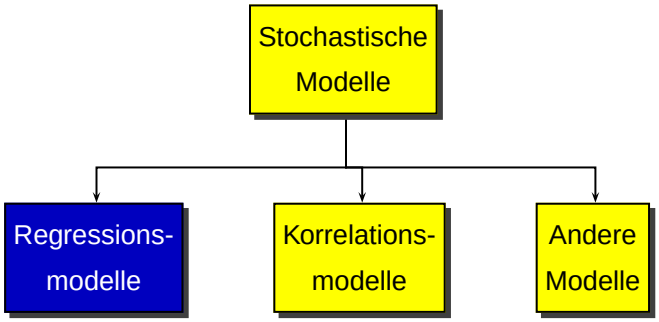
Stochastische Modelle

1. Formulieren Hypothesen über 2 Komponenten
   - Deterministische Komponente
   - Zufälligen Fehler
2. Beispiel:
   Das Verkaufsvolumen, Y, ist 10 mal so groß wie die Werbeausgaben, X, **plus** einem zufälligen Fehler, ϵ.
   - $Y = 10\,X + \epsilon$
   - Der zufällige Fehler, ϵ, wird durch Faktoren bestimmt, die nicht durch die Werbeausgaben beeinflusst werden: Wetter, Konkurrenzangebote, Konjunkturzyklus, Änderungen im Distributionsnetz, Kunde Mayer hat im Lotto gewonnen, ...



Regressionsmodelle

Typen von stochastischen Modellen



1. Beantworten die Frage: „Wie sieht der Zusammenhang zwischen Variablen aus?“
2. Die Gleichung besteht aus:
   - Eine numerische, **abhängige** (Response-) Variable, Y.
     - Was soll erklärt (prognostiziert) werden?
   - Eine oder mehrere numerische oder kategoriale, **unabhängige** (erklärende) Variable, $X_1, X_2, \ldots$.
     - Was hilft Y zu erklären (prognostizieren)?
3. Hauptsächlich verwendet für:
   - Schätzen und Testen:
     - Gibt es einen Zusammenhg. zwischen den Variablen?
     - Wie stark ist der Zusammenhang?
   - Prognose

Regression // Modellierungsschritte

1. Hypothese über die deterministische Komponente.
2. Schätzen die unbekannten Parameter im Modell.
3. Spezifizieren die Wahrscheinlichkeitsverteilung des Fehlerterms.
   - Schätzen die Standardabweichung des Fehlers.
4. Überprüfen das Modell.
5. Verwenden das Modell für die Prognose.

Modellspezifikation

- **Hypothese über die deterministische Komponente.**

2. Schätzen die unbekannten Parameter im Modell.

3. Spezifizieren die Wahrscheinlichkeitsverteilung des Fehlerterms.
   - Schätzen die Standardabweichung des Fehlers.

4. Überprüfen das Modell.

5. Verwenden das Modell für die Prognose.

Spezifikation eines Modells

1. Definieren die Variablen, die im Modell aufgenommen werden sollen.
   - Konzeptuell z.B., Werbeausgaben, Preis
   - Empirisch z.B., Listenpreis, tatsächlicher Verkaufspreis
   - Messung €, Einheiten,. . . ; tägl., monatl.,. . .

2. Hypothesen über die Art des Zusammenhangs
   - Erwartete Effekte: Größe der Koeffizienten, Vorzeichen
   - Funktionale Form: linear oder nicht-linear
   - Interaktionen

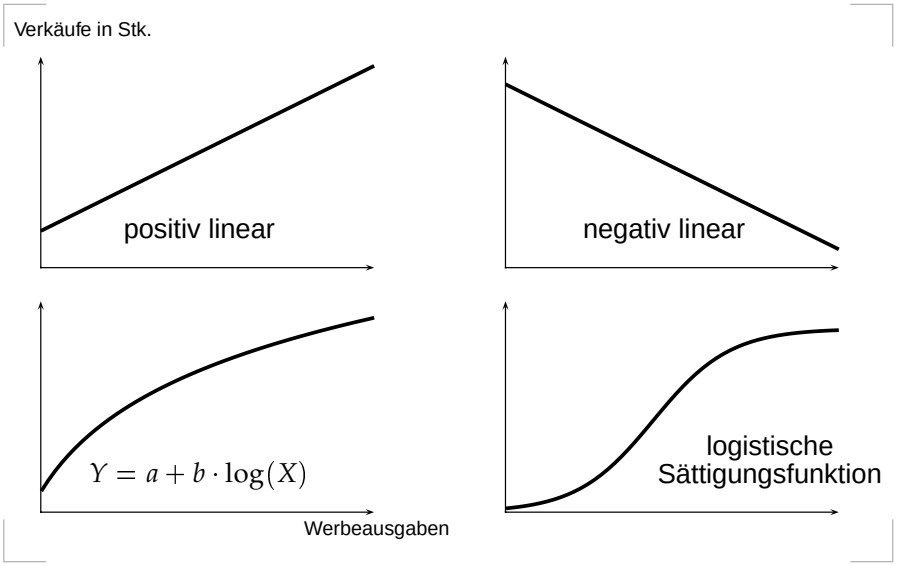
Modellspezifikation basiert auf Theorie

- Theorie des Gebiets (Substanzwissenschaften): Verhaltenswissenschaften, Konsumtheorie, Theorie der Firma, ...

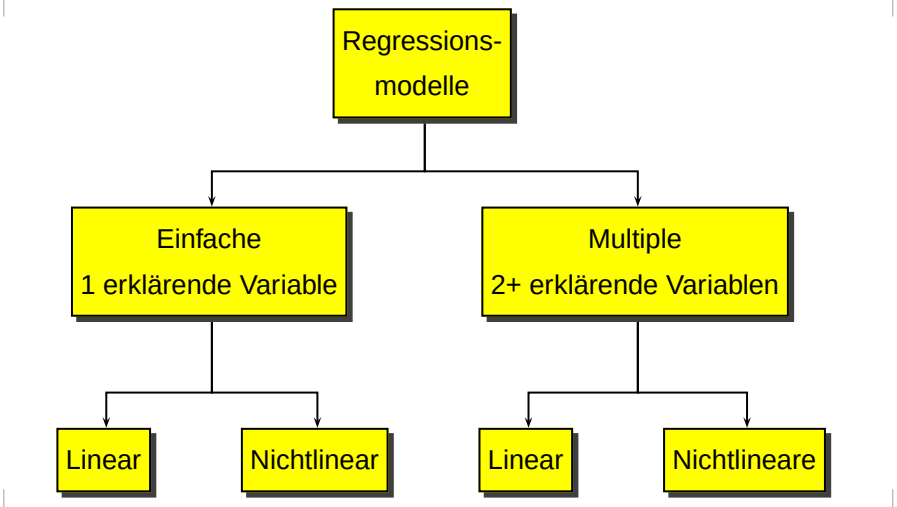
- Mathematische Theorie: Zur Quantifizierung, zur Modellierung von Maximierungsverhalten, ...

- Vorhergehende Forschung: Z.B., publizierte Arbeiten dazu.

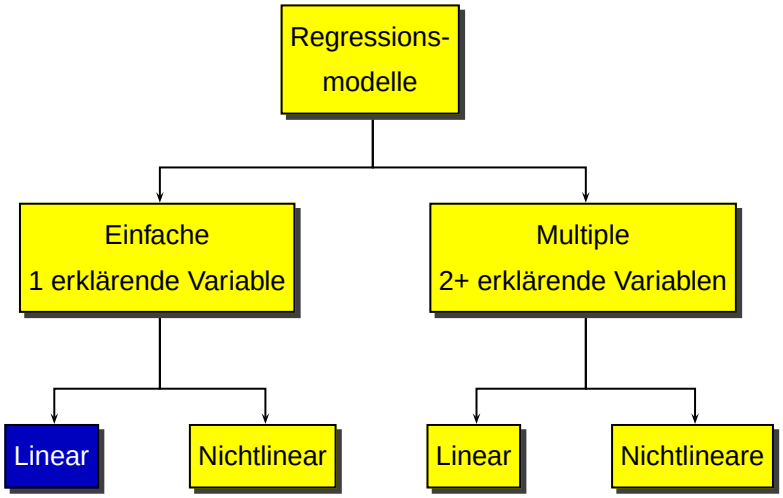
- Plausibilität, „Hausverstand“, begründbare Erfahrung.



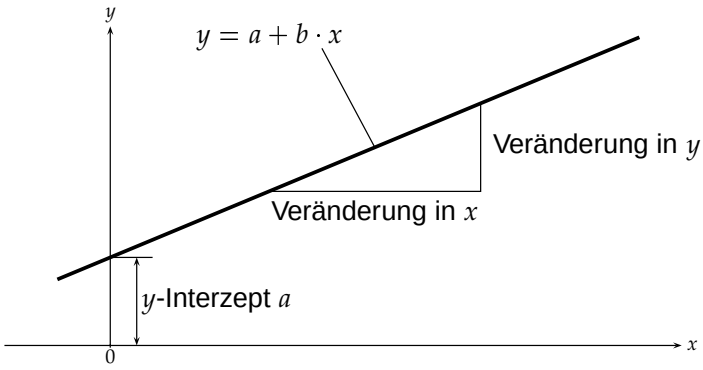
Typen von Regressionsmodellen



Lineares Regressionsmodell



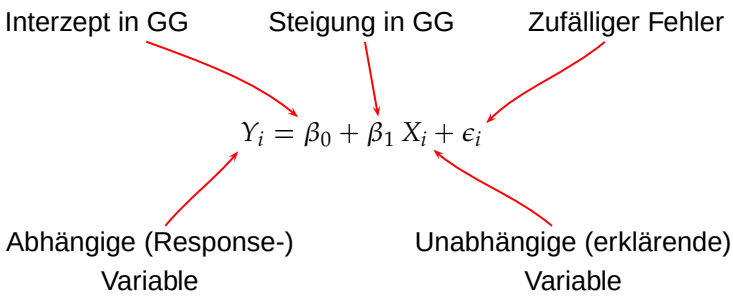
Linearer Zusammenhang



$$\text{Steigung } b = \frac{\text{Veränderung in } y}{\text{Veränderung in } x}$$

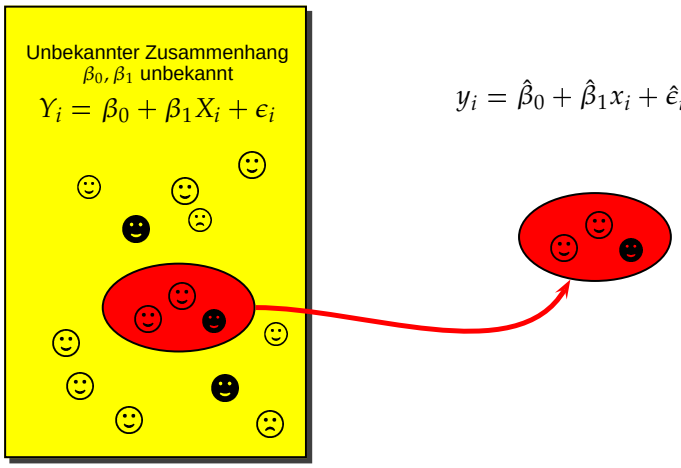
Lineares Regressionsmodell

1. Der Zusammenhang zwischen den Variablen ist eine lineare Funktion.



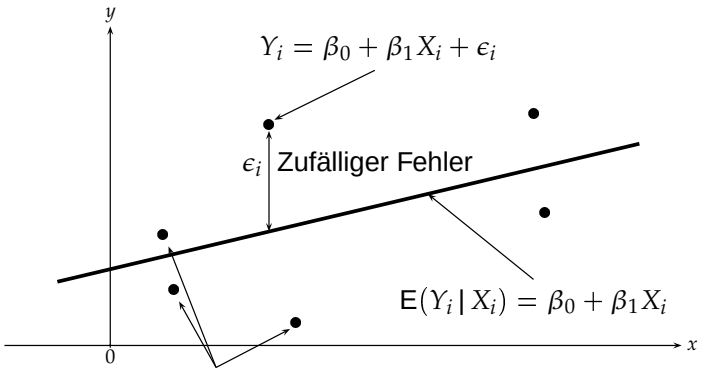
Grundgesamtheit

Stichprobe (Zufallsauswahl)



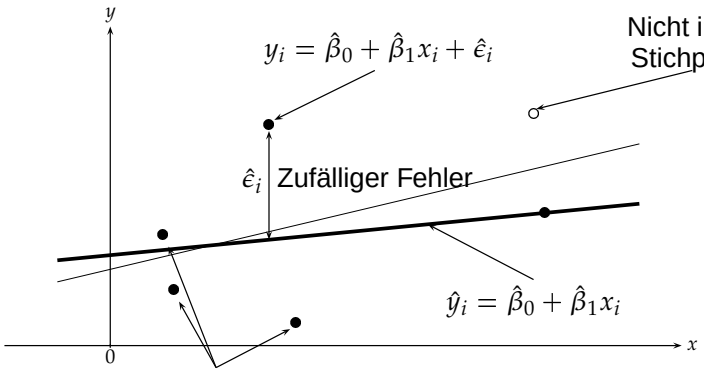
$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\epsilon}_i$$

Lineares Modell in der Grundgesamtheit



Datenpunkte der Grundgesamtheit

Lineares Modell in der Stichprobe



Beobachtete Werte

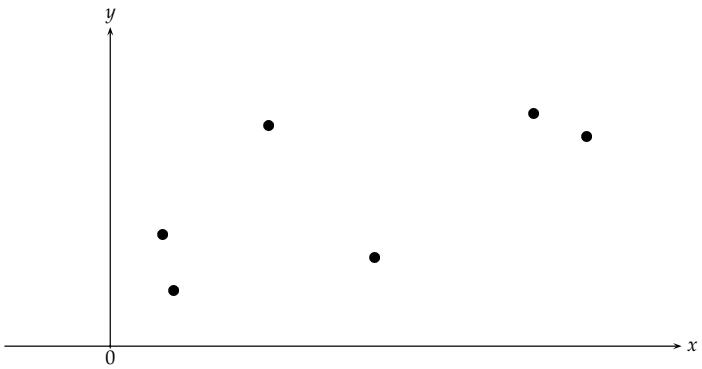
Schätzung der Parameter: Methode der kleinsten Quadrate

Regression // Modellierungsschritte

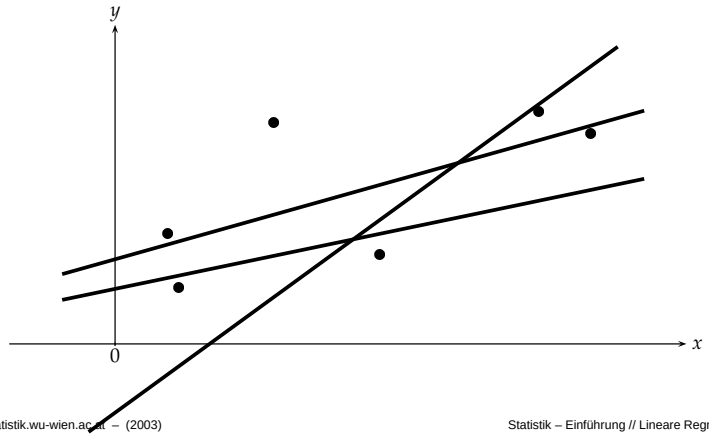
1. Hypothese über die deterministische Komponente.
2. **Schätzen die unbekannten Parameter im Modell.**
3. Spezifizieren die Wahrscheinlichkeitsverteilung des Fehlerterms.
   - Schätzen die Standardabweichung des Fehlers.
4. Überprüfen das Modell.
5. Verwenden das Modell für die Prognose.

Streudiagramm

1. Graphische Darstellung aller Paare (x_i, y_i).
2. Gibt einen Eindruck über die Anpassungsgüte des Modells.



Wie würden Sie eine Gerade durch die Punkte(wolke) legen?
Wie bestimmt man die Gerade, die „am besten“ passt?



Methode der kleinsten Quadrate

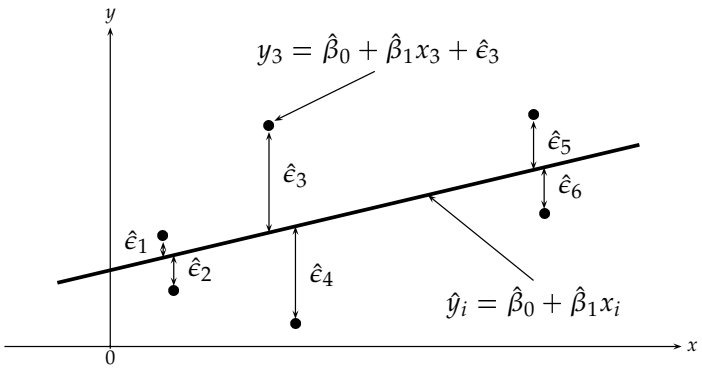
1. „Bester Fit“ (Beste Anpassung) bedeutet, dass die Differenz zwischen tatsächlichen y-Werten und vorhergesagten (modellierten) Werten, $\hat{y}$, minimal ist.
   - Aber: positive Differenzen löschen negative aus. Daher wird die quadratische Abweichung verwendet.

$$\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} \epsilon_i^2 \longrightarrow \min$$

2. Die Methode der kleinsten Quadrate minimiert die Summe der quadrierten Differenzen.
   (Engl.: (ordinary) least squares, (O)LS)

Kleinste-Quadrate (LS) graphisch

LS minimiert $\sum_{i=1}^{n} \hat{\epsilon}_i^2 = \hat{\epsilon}_1^2 + \ldots + \hat{\epsilon}_6^2$



- Vorhersagegleichung

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

- Schätzer für die Steigung

$$\hat{\beta}_1 = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2}$$

- Schätzer für den y-Interzept

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

Berechnungstabelle

x	y	x^2	y^2	$x\,y$
x_1	y_1	x_1^2	y_1^2	$x_1\,y_1$
x_2	y_2	x_2^2	y_2^2	$x_2\,y_2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_n	y_n	x_n^2	y_n^2	$x_n\,y_n$
$\sum x_i$	$\sum y_i$	$\sum x_i^2$	$\sum y_i^2$	$\sum x_i\,y_i$

Interpretation der Koeffizienten

1. Steigung, $\hat{\beta}_1$
   - Das geschätzte y ändert sich um $\hat{\beta}_1$, wenn x um eine Einheit erhöht wird.

     Wenn $\hat{\beta}_1 = 2$, dann steigen die Verkäufe, y, um 2 Einheiten für jede Einheit mehr in den Werbeausgaben, x.

2. y-Interzept, $\hat{\beta}_0$
   - $\hat{\beta}_0$ ist der durchschnittliche Wert von y, wenn $x = 0$.

     Ist $\hat{\beta}_0 = 4$, so werden 4 Verkaufseinheiten erwartet, wenn keine Werbeausgaben gemacht werden.

Sie sind ein Marketinganalyst für Haribo.
Sie sammeln folgende Daten:

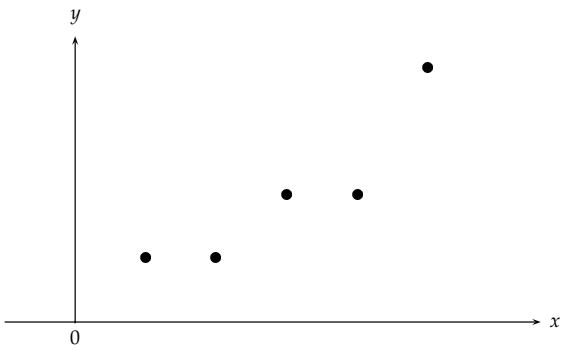
Monat	Werbeausgaben (1000 €)	Verkäufe (Einheiten)
1	1	1
2	2	1
3	3	2
4	4	2
5	5	4

Wie sieht die Beziehung zwischen Verkäufen und Werbeausgaben aus?

Streudiagramm: Werbung vs. Verkäufe

Vorhersagegleichung $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$

y . . . Verkäufe x . . . Werbeausgaben in 1000 €



Berechnungstabelle: Beispiel

x	y	x^2	y^2	$x\,y$
1	1	1	1	1
2	1	4	1	2
3	2	9	4	6
4	2	16	4	8
5	4	25	16	20
15	10	55	26	37

Schätzer für die Steigung

$$\hat{\beta}_1 = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2} = \frac{5 \cdot 37 - 15 \cdot 10}{5 \cdot 55 - 15^2} = 0.70$$

Schätzer für den y-Interzept

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 2 - 0.70 \cdot 3 = -0.10$$

Interpretation der Schätzung

Das geschätzte Modell lautet:

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\epsilon}_i \quad \text{bzw.} \quad \hat{y}_i = -0.10 + 0.7\, x_i$$

1. Steigung, $\hat{\beta}_1$
   - Das Verkaufsvolumen, y, steigt im Durchschnitt um 0.7 Einheiten pro eingesetzte Einheit Werbeausgaben, x.
2. y-Interzept, $\hat{\beta}_0$
   - Das durchschnittliche Verkaufsvolumen ist -0.10, wenn die Werbeausgaben Null sind.

     (In der Praxis erwartet man auch eine geringe Verkaufsmenge, wenn keine Werbung gemacht wird.)

Computer Output // R

```
> x=c(1,2,3,4,5)
> y=c(1,1,2,2,4)
> lm(y ~ x)

Call:
lm(formula = y ~ x)

Coefficients:
(Intercept)           x
       -0.1         0.7
```

$\hat{\beta}_0$ $\hat{\beta}_1$

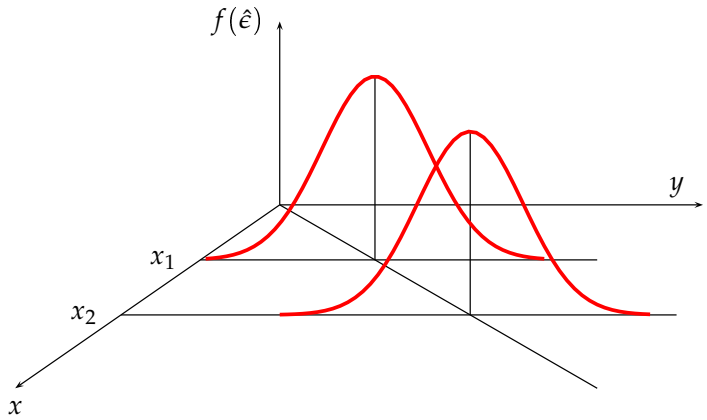
Wahrscheinlichkeitsverteilung des Fehlerterms

Regression // Modellierungsschritte

1. Hypothese über die deterministische Komponente.
2. Schätzen die unbekannten Parameter im Modell.
- **Spezifizieren die Wahrscheinlichkeitsverteilung des Fehlerterms.**
  - Schätzen die Standardabweichung des Fehlers.
4. Überprüfen das Modell.
5. Verwenden das Modell für die Prognose.

Lineare Regression // Annahmen

1. Das Mittel der Wahrscheinlichkeitsverteilung des Fehlers ist Null.
2. Die Wahrscheinlichkeitsverteilung des Fehlers hat eine konstante Varianz. (Unabhängig von x)
3. Der Fehler ist normalverteilt.
4. Die Fehler sind unabhängig von einander.

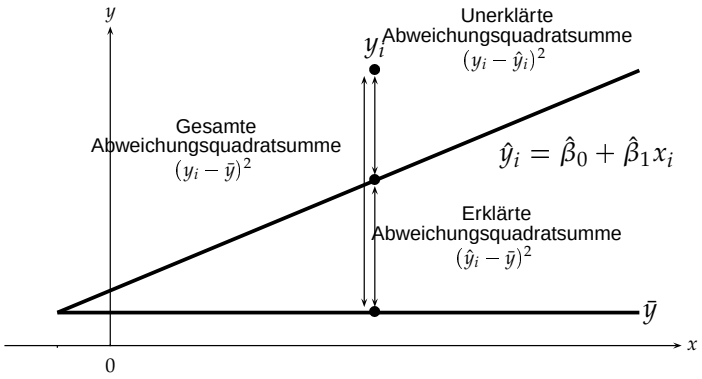


Variation im (Zufalls-) Fehler

1. Die Varianz des Fehlers ist die Varianz der Differenz zwischen aktuellem y und modelliertem $\hat{y}$.
2. Die Variation im Fehler wird gemessen durch die Stichprobenstandardabweichung von $\hat{\epsilon}$, s_{ϵ}.
3. Die Größe der Variation im Fehler beeinflusst:
   - Parametersignifikanz
   - Vorhersagegenauigkeit
4. Die Variation im Fehler steht für die Unsicherheit im Modell. Sie entsteht durch Effekte, die nicht durch das Modell erfasst werden.

Maße für die Variation in der Regression

1. Gesamtsumme der Abweichungsquadrate (in y)
   - Variation der beobachteten y-Werte um dessen Mittel $\bar{y}$.
   - Dt.: Summe der quadratischen Abweichungen total, SQT.
   - Engl.: Total sum of squares, SS_{yy}.
2. Erklärte Variation
   - Variation, die durch die Beziehung von x und y entsteht.
   - Dt.: Erklärte Abweichungsquadratsumme, SQE.
   - Engl.: Sum of squares in the regression, SSR.
3. Unerklärte Variation
   - Variation, die nicht in der Regression erklärt wird.
   - Dt.: Abweichungsquadratsumme der Residuen, SQR.
   - Engl.: Sum of squares in the errors, SSE.



SQ-Gesamt = SQ-Erklärt + SQ-Unerklärt

$$SS_{yy} = SSR + SSE$$

$$\sum (y_i - \bar{y})^2 = \sum (\hat{y}_i - \bar{y})^2 + \sum (y_i - \hat{y}_i)^2$$

Bestimmtheitsmaß

- Anteil der Variation, die durch die Beziehung zwischen x und y „erklärt“ wird. $0 \leq r^2 \leq 1$

$$r^2 = \frac{\text{erklärte Variation}}{\text{gesamte Variation}} = \frac{SQE}{SQT}$$

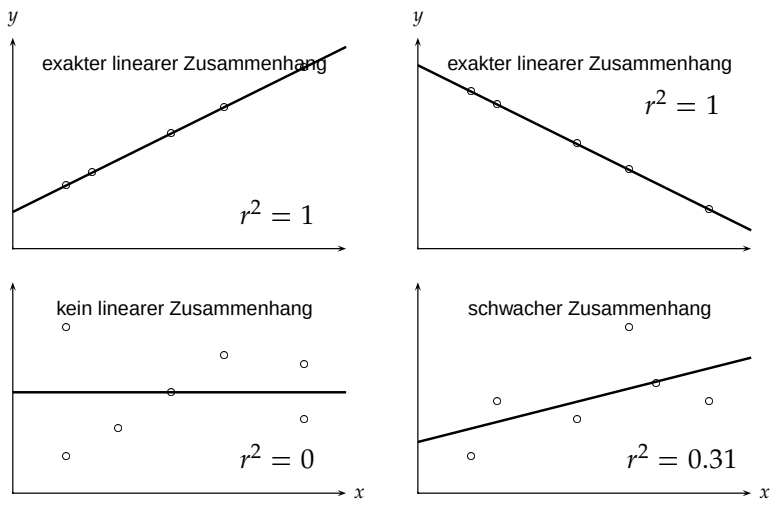
$$= \frac{SSR}{SS_{yy}} = \frac{\sum (\hat{y}_i - \bar{y})^2}{\sum (y_i - \bar{y})^2}$$

Eine andere Darstellung:

$$r^2 = 1 - \frac{\text{unerklärte Variation}}{\text{gesamte Variation}} = 1 - \frac{SQR}{SQT} = 1 - \frac{SSE}{SS_{yy}} = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$$

$$r^2 = 1 - \frac{\sum y_i^2 - \hat{\beta}_0 \sum y_i - \hat{\beta}_1 \sum x_i y_i}{\sum y_i^2 - (1/n)(\sum y_i)^2}$$

Bestimmtheitsmaß // Beispiele



Sie sind ein Marketinganalyst für Haribo. Sie schätzen

$$\hat{y}_i = -0.1 + 0.7\, x_i$$

wobei y ... Verkäufe und x ... Werbeausgaben in 1000 € sind.
Die Daten dazu sind:

Monat	Werbeausgaben (1000 €)	Verkäufe (Einheiten)
1	1	1
2	2	1
3	3	2
4	4	2
5	5	4

Berechnungstabelle: Beispiel

x	y	x^2	y^2	$x\,y$
1	1	1	1	1
2	1	4	1	2
3	2	9	4	6
4	2	16	4	8
5	4	25	16	20
15	10	55	26	37

$$\begin{aligned}\hat{y}_i &= \hat{\beta}_0 + \hat{\beta}_1 x_i \\ &= -0.10 + 0.70 x_i\end{aligned}$$

$$\begin{aligned}r^2 &= 1 - \frac{\sum y_i^2 - \hat{\beta}_0 \sum y_i - \hat{\beta}_1 \sum x_i y_i}{\sum y_i^2 - (1/n)(\sum y_i)^2} \\ &= 1 - \frac{26 - (-0.10)\cdot 10 - 0.70 \cdot 37}{26 - (1/5)\cdot 10^2} \\ &= 0.817\end{aligned}$$

Wie ist das Bestimmtheitsmaß von $r^2 = 0.817$ zu interpretieren?

Computer Output // R

```
> summary(lm(y ~ x))
Call:
lm(formula = y ~ x)

Residuals:
       1          2          3          4          5
 4.000e-01 -3.000e-01 -3.237e-17 -7.000e-01  6.000e-01

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.1000     0.6351   0.157   0.8849
x             0.7000     0.1915   3.656   0.0354 *
--
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.6055 on 3 degrees of freedom
Multiple R-Squared:  0.8167 , Adjusted R-squared: 0.7556
F-statistic: 13.36 on 1 and 3 DF,  p-value: 0.03535
```

r^2

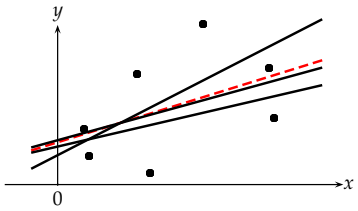
Modellüberprüfung

Regression // Modellierungsschritte

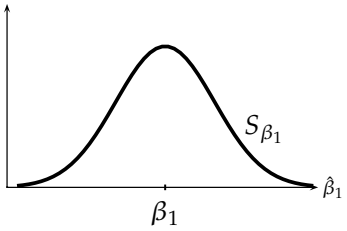
1. Hypothese über die deterministische Komponente.
2. Schätzen die unbekannten Parameter im Modell.
3. Spezifizieren die Wahrscheinlichkeitsverteilung des Fehlerterms.
   - Schätzen die Standardabweichung des Fehlers.
4. **Überprüfen das Modell.**
5. Verwenden das Modell für die Prognose.

Test des Steigungsparameters

1. $\hat{\beta}_1$ ist ein Indikator für den linearen Zusammenhang zwischen x und y.
2. Hängt von der Steigung, β_1, in der Grundgesamtheit (wahrem Modell) ab.
3. Hypothesen
   - H_0: $\beta_1 = 0$ (Kein linearer Zusammenhang)
   - H_A: $\beta_1 \neq 0$ (Linearer Zusammenhang)
4. Theoretische Basis ist die Stichprobenverteilung der Steigung.



Stichprobenverteilung



Alle möglichen
Stichproben-Steigungen

Stichpr. 1	0.51
Stichpr. 2	0.27
Stichpr. 3	0.24
...	...

Sehr große Anzahl von

Steigungen.

Teststatistik für die Steigung

$$t_{n-2} = \frac{\hat{\beta}_1 - \beta_1}{s_{\hat{\beta}_1}}$$

wobei

$$s^2_{\hat{\beta}_1} = \frac{s^2_\epsilon}{\sum (x_i - \bar{x})^2} \quad \text{und} \quad s^2_\epsilon = \frac{1}{n-2} \sum \hat{\epsilon}_i^2$$

t-Verteilung, Freiheitsgrade: $\nu = n - 2$

Zur Berechnung:

$$s^2_\epsilon = \frac{1}{n-2} \sum \hat{\epsilon}_i^2 = \frac{1}{n-2} \left[\sum y_i^2 - \hat{\beta}_0 \sum y_i - \hat{\beta}_1 \sum x_i y_i \right]$$

$$\sum (x_i - \bar{x})^2 = \sum x_i^2 - \frac{1}{n} \left(\sum x_i \right)^2$$

Test auf den Steigungsparameter // Beispiel

Sie sind ein Marketinganalyst für Haribo. Sie schätzen

$$\hat{y}_i = -0.1 + 0.7\, x_i \quad \text{und} \quad s_\epsilon = 0.6055$$

wobei y ... Verkäufe und x ... Werbeausgaben in 1000 € sind.
Gibt es eine Beziehung zwischen Werbeausgaben und Verkäufen?
Testen Sie bei einem Signifikanzniveau von 10%.

Monat	Werbeausgaben	Verkäufe
1	1	1
2	2	1
3	3	2
4	4	2
5	5	4

x	y	x^2	y^2	xy
1	1	1	1	1
2	1	4	1	2
3	2	9	4	6
4	2	16	4	8
5	4	25	16	20
15	10	55	26	37

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i = -0.10 + 0.70 x_i$$

$$\begin{aligned} s_\epsilon^2 &= \frac{1}{n-2} \sum \epsilon_i^2 = \\ &= \frac{1}{n-2} [\sum y_i^2 - \hat{\beta}_0 \sum y_i - \hat{\beta}_1 \sum x_i y_i] \\ &= \frac{1}{5-2} [26 - (-0.10) \cdot 10 - 0.70 \cdot 37] \\ &= 0.3666 \end{aligned}$$

$$\begin{aligned} \sum (x_i - \bar{x})^2 &= \sum x_i^2 - \frac{1}{n} (\sum x_i)^2 \\ &= 55 - \frac{1}{5} (15)^2 = 10 \end{aligned}$$

$$s_{\hat{\beta}_1}^2 = \frac{s_\epsilon^2}{\sum (x_i - \bar{x})^2} = \frac{0.3666}{10} = 0.03666$$

$$t_{n-2} = \frac{\hat{\beta}_1 - \beta_1}{s_{\hat{\beta}_1}} = \frac{0.70 - 0}{0.1915} = 3.655$$

Test auf den Steigungsparameter // Lösung

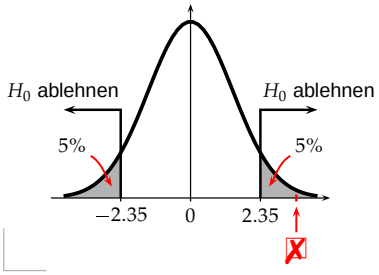
$H_0: \beta_1 = 0$

$H_A: \beta_1 \neq 0$

$\alpha = 0.10$

$\nu = 5 - 2 = 3$

Kritische Werte: ± 2.353



Teststatistik:

$$t = \frac{\hat{\beta}_1 - \beta_1}{s_{\hat{\beta}_1}} = \frac{0.70 - 0}{0.1915} = 3.655$$

Entscheidung:

H_0 wird abgelehnt.

Interpretation:

Es gibt Evidenz für eine Beziehung zwischen Werbeausgaben und Verkäufen.

Computer Output // R

```
> summary(lm(y x))
Call:
lm(formula = y ~ x)

Residuals:
       1          2          3          4          5
 4.000e-01 -3.000e-01 -3.237e-17 -7.000e-01  6.000e-01

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.1000     0.6351   0.157   0.8849
x             0.7000     0.1915   3.656   0.0354 *
...
...
```

$s_{\hat{\beta}_1}$ $t = \frac{\hat{\beta}_1}{s_{\hat{\beta}_1}}$

Verwenden das Modell für die Prognose

Regression // Modellierungsschritte

1. Hypothese über die deterministische Komponente.
2. Schätzen die unbekannten Parameter im Modell.
3. Spezifizieren die Wahrscheinlichkeitsverteilung des Fehlerterms.
   - Schätzen die Standardabweichung des Fehlers.
4. Überprüfen das Modell.
- **Verwenden das Modell für die Prognose.**

Prognose mit Regressionsmodellen

Arten von Prognosen

- Punktprognose
- Intervallprognose

Wahres Modell:

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

Geschätztes Modell:

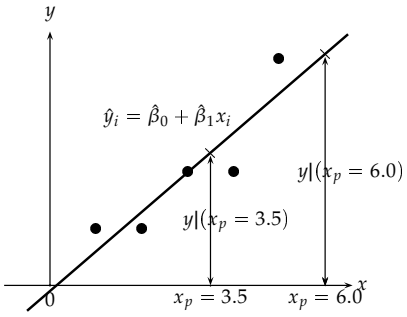
$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

Prognose: $\hat{y}_p = y|x_p$

$$\hat{y}_p = \hat{\beta}_0 + \hat{\beta}_1 x_p$$

Punktprognose // Beispiel

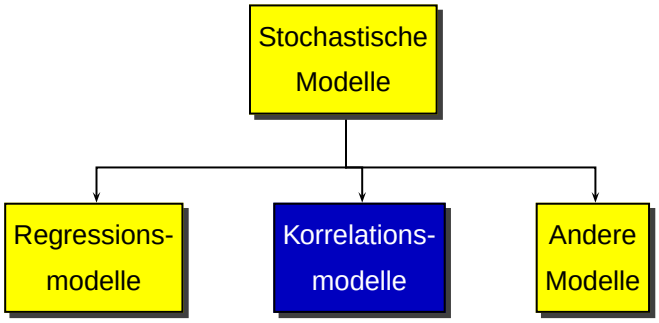
Marketinganalyst bei Haribo ($\hat{y}_i = -0.10 + 0.70 x_i$)



$$y|(x_p = 3.5) = -0.10 + 0.70 \cdot 3.5 = 2.35$$

$$y|(x_p = 6.0) = -0.10 + 0.70 \cdot 6.0 = 4.10$$

Korrelationsmodelle



Korrelationsmodelle

1. Beantworten die Frage: „Wie stark ist der lineare Zusammenhang zwischen 2 Variablen?“
2. Korrelationskoeffizient
   - Korrelationskoeffizient in der Grundgesamtheit wird als ρ (rho) bezeichnet.
   - Er liegt immer zwischen -1 und $+1$: $-1 \leq \rho \leq 1$.
   - Er misst den linearen Zusammenhang.

Stichprobenkorrelationskoeffizient

Der Stichprobenkorrelationskoeffizient, r, zwischen x und y ist der Pearson'sche Produktmoment Koeffizient der Korrelation:

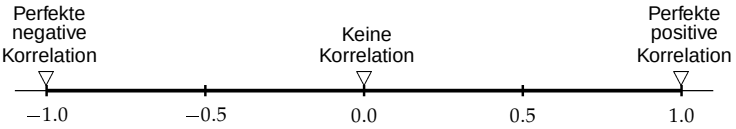
$$r = \sqrt{\text{Bestimmtheitsmaß}}$$

aber mit einem **Vorzeichen** versehen.

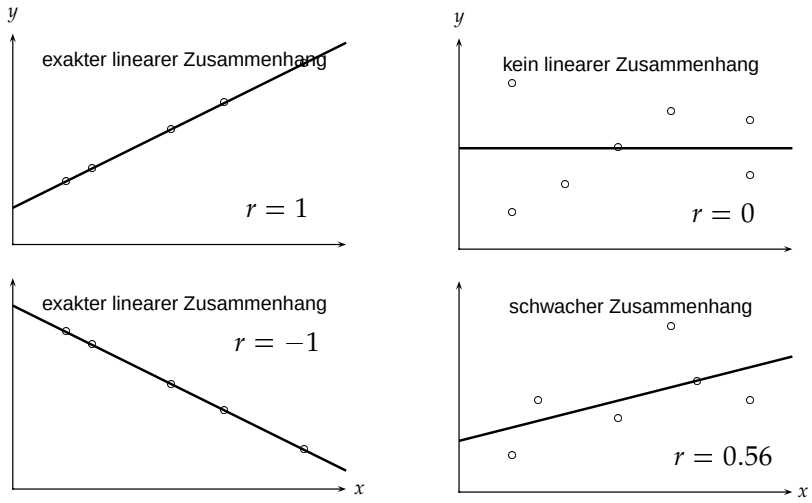
$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2}\sqrt{\sum(y_i - \bar{y})^2}}$$

Berechnung über die Arbeitstabelle:

$$r = \frac{\sum x_i y_i - n\,\bar{x}\,\bar{y}}{\sqrt{\sum x_i^2 - n\bar{x}^2}\sqrt{\sum y_i^2 - n\bar{y}^2}}$$



Korrelationskoeffizient: Beispiele



Test auf den Korrelationskoeffizienten

1. Wir testen auf einen linearen Zusammenhang zwischen zwei numerischen (quantitativen) Variablen.
2. Das ist das gleiche Problem, wie der Test auf die Steigung in der Grundgesamtheit, β_1, auf Null.
3. Hypothesen
   - H_0: $\rho = 0$ (Keine Korrelation)
   - H_A: $\rho \neq 0$ (Korrelation)

1. Beschrieben lineares Regressionsmodell.
2. Gaben die einzelnen Modellierungsschritte an.
3. Erklärten die (gewöhnliche) Methode der kleinsten Quadrate.
4. Berechneten und testeten Regressionskoeffizienten.
5. Prognostizieren die „Antwort“-Variable.
6. Interpretieren einen Computerausdruck.
7. Zeigten den Zusammenhang zwischen der Steigung der Regressionsgeraden und dem Korrelationskoeffizienten auf.