GML

Ökonometrie: Lin. Reg. Inkl. Max. Lik.

All models are wrong but some are useful, turn data into assets

Prof. Dr. Frank Lehrbass

Copyright

Quellen

- Wir verwenden die genannten Bücher und geben die Quellen im Text wie folgt an:

- Verbeek, 2017, Seite xx
- Schröder, 2002, dito
- Danielsson, 2011, dito

Was man üblicherweise lernt …

Bei der linearen Regression wird versucht, eine beobachtete abhängige Variable durch eine oder mehrere unabhängige Variablen zu erklären.

Zwei (oder mehr) metrische Merkmale, X,Y, wobei angenommen wird, dass X unabhängig ist, der Wert von Y aber von X abhängt.

Das erklärte oder abhängige Merkmal Y ist häufig eine Zielgröße (z.B. Gewinn, Umsatz), das erklärende oder unabhängige Merkmal X häufig eine Instrumentgröße (Absatzpreis, Werbeetat).

Mathematisch lautet die Regressionsfunktion (engl.: regression function):

$$y = f(x)$$

Am häufigsten verwendet wird ein lineares Regressionsmodell (engl.: linear regression model):

Koeffizient

$$y = a + b \cdot x$$

abhängige Variable

Konstante

unabhängige Variable

Quelle: Früheres Skript Big Data Analytics der FOM

Was man üblicherweise lernt …

- Im Allgemeinen liegen die Punkte nicht exakt auf einer Geraden, das Regressionsmodell beinhaltet also für jede Beobachtung *i* einen Fehlerterm e_i (Residuum):

$$y_i = a + b \cdot x_i + e_i$$

- Die „beste" Regressionsgerade minimiert die Quadratsumme dieser Residuen:

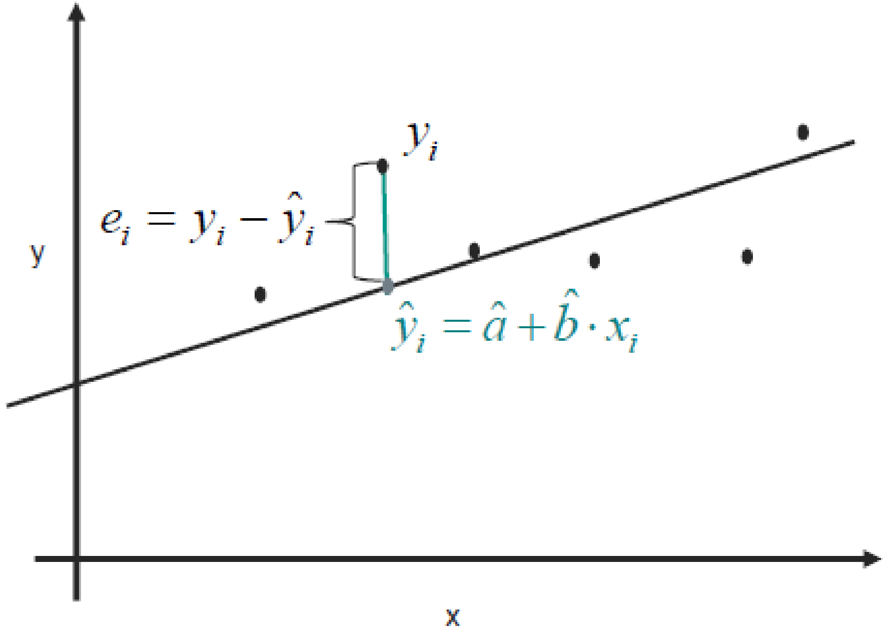
$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - (a + bx_i))^2 =: q(a,b)$$

- Die optimalen Schätzwerte für *a* und *b* können jetzt mit Hilfe der partiellen Ableitungen von *q(a,b)* nach *a* und *b* gefunden werden: Überprüfen der notwendigen und hinreichenden Bedingungen für die Existenz von Extremwerten...

Quelle: Früheres Skript Big Data Analytics der FOM

Was man üblicherweise lernt …

Methode der kleinsten Quadrate



$$\min \sum_{i=1}^{n} e_i^2$$

Quelle: Früheres Skript Big Data Analytics der FOM

Was man üblicherweise lernt …

- Der optimale Schätzer für b ist die Kovarianz von x und y geteilt durch die Varianz von x: $\hat{b} = \frac{s_{xy}}{s_x^2}$

  Der optimale Schätzer für a ist der Mittelwert von y minus b mal den Mittelwert von x:

  - $\hat{a} = \bar{y} - \hat{b} \cdot \bar{x}$

  ⇨ Kovarianz, Varianz und Mittelwert können mit dem bisher Gelernten berechnet werden.

- Die Punktprognose für y_0 gegeben einen gegeben (neuen) Wert x_0 ist die Summe aus dem Schätzer für *a* und x_0 mal dem Schätzer für *b*. x_0 wird dabei entweder gesteuert (z.B. Marketingetat) oder ist aus anderen Gründen bekannt (z.B. Nachfrage).

$$\hat{y}_0 = \hat{a} + \hat{b} \cdot x_0$$

Quelle: Früheres Skript Big Data Analytics der FOM

CAPM

- Wir verwenden das Bsp aus Verbeek, 2017, Seite 40ff
- Das Coding heisst: GML_02_01 capm.R
- Wir schätzen das Beta für Unternehmen aus dem Bausektor (constructions)
- Mit einer einfachen lin. Reg. Monatlicher „Xcess return for a sub-index on xcess for broad index”

- Was erwarten wir als Ergebnis der Schätzung aus Sicht des CAPM?
- Mit welchem “Lernalgorithmus” // Schätzverfahren bestimmen wir die Modellparameter?

CAPM

- Führen Sie Zeilen 1-42 aus!

- Werden die Erwartungen aus Sicht des CAPM erfüllt?
- Wieviel Prozent der Varianz des Regressanden erklärt das Modell?
- Welchen Anteil hat das idiosynkratische Risiko?
- Wie hoch ist das Beta?
- Erzeugen Sie eine Grafik: plot(X,Y) und abline(reg)!

```
> summary(reg)

Call:
lm(formula = Y ~ X)

Residuals:
    Min      1Q  Median      3Q     Max 
-12.879  -1.641  -0.115   1.520  11.725 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) -0.07259    0.11542  -0.629     0.53    
X            1.16817    0.02549  45.837   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.837 on 608 degrees of freedom
Multiple R-squared:  0.7756,	Adjusted R-squared:  0.7752 
F-statistic:  2101 on 1 and 608 DF,  p-value: < 2.2e-16
```

CAPM

- Nun machen wir ein Experiment, indem wir eine Aktienrendite künstlich gemäss CAPM erzeugen

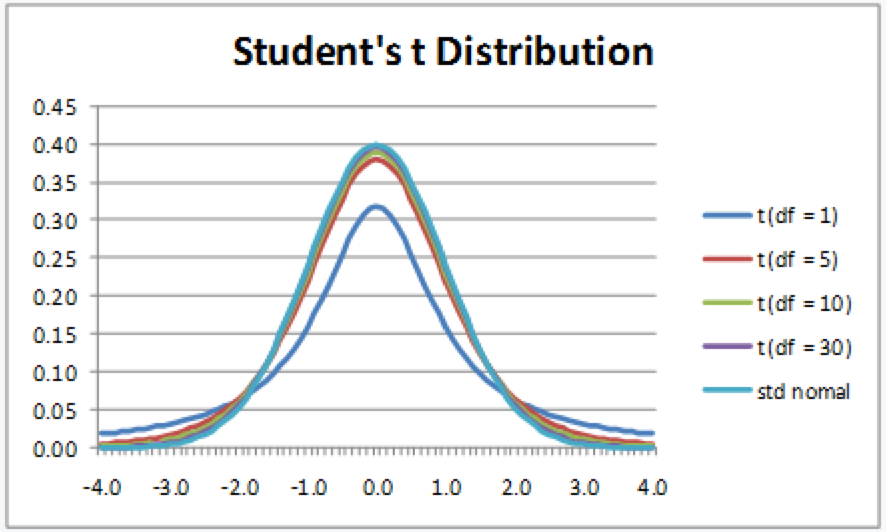
  set.seed(2023)
  eps = rnorm(length(X),0,2.5)
  *Y_made_up = 0.5*X+eps*

- Welches Beta hat diese Aktie?
- Was und warum machen wir es so, insbes. Seed?
- Führen Sie Zeilen 47-54 aus!
- Was entdecken Sie wieder, insbes. sigma?
- Ändern Sie den Seed von 2023 auf Ihr Geburtsjahr!
- Was ändert sich?

```
> summary(reg_hide_n_seek)

Call:
lm(formula = Y_made_up ~ X)

Residuals:
    Min      1Q  Median      3Q     Max 
-8.124 -1.722   0.016   1.673  11.436 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  0.02729    0.10056   0.271    0.786    
X            0.47541    0.02220  21.412   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.472 on 608 degrees of freedom
Multiple R-squared:  0.4299,	Adjusted R-squared:  0.429 
F-statistic: 458.5 on 1 and 608 DF,  p-value: < 2.2e-16
```

CAPM

- Zusammenfassung (Chart von real-statistics.com)



- Wir haben df = 608
- Welche Glocke beschreibt die Verteilung der geschätzten Betas?
- Wo liegt Ihr Schwerpunkt?
- Die Schätzwerte für Beta sind Student-t verteilt. Teilt man sie durch ihren SE = Stdev der Schätzwerte, so gehorcht diese Teststatistik einer …. Verteilung. Wie sieht die Grafik aus?

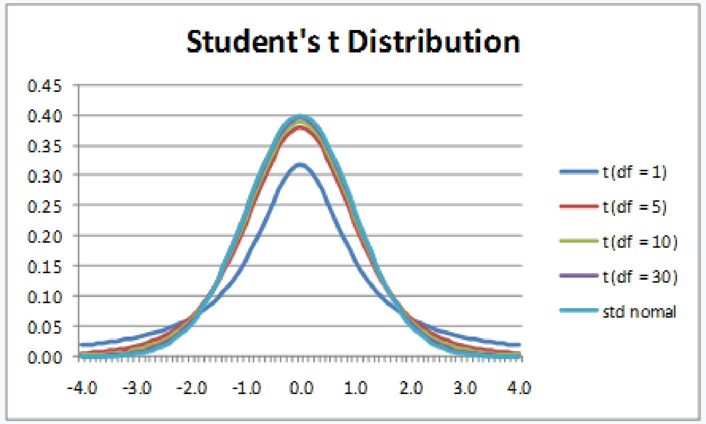
CAPM

- Hypothesentest Rezeptur

Grundlagen Statistik

WH: Ablauf statistischer Tests

1. Formulierung der zu überprüfenden Hypothesen (Nullhypothese H_0 und Alternativhypothese H_1).
2. Festlegen der zulässigen Fehlerwahrscheinlichkeit (Testniveau α) für das spätere Testergebnis (α = 0,05; 0,01; …).
3. Berechnung einer Prüfgröße (Teststatistik), die sich aus der Stichprobe ermitteln lässt, und Bestimmung der Verteilung der Prüfgröße.
4. Bestimmung eines kritischen Wertes aus der Verteilung der Prüfgröße (und aus dem Testniveau α), dessen Unter- oder Überschreiten zur Ablehnung der Nullhypothese führt.
5. Vergleich von Prüfgröße und kritischem Wert und Entscheidung.



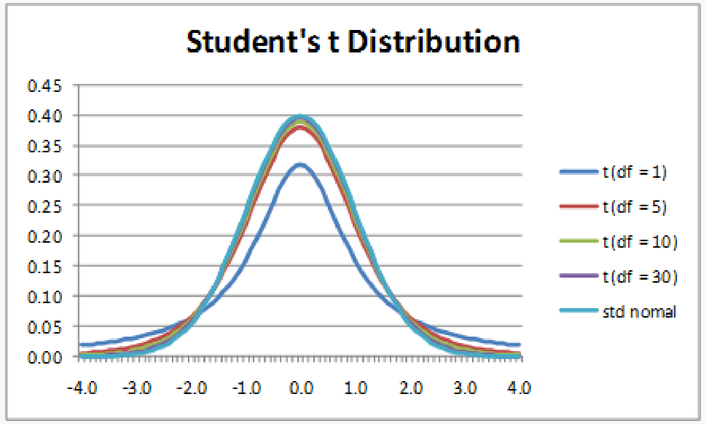
- Wir wollen die H0: Beta = 0 testen.

- Was tun?

CAPM

- Hypothesentest
- Wir wollen die H0: Beta = 0 testen.
- Was ist unsere Prüfgrösse?
- Signifikanzniveau sei 5%.
- Bei welchem krit. Wert verwerfen wir H0?
- Welche (mind. 3) H0 können mit dem R Output getestet werden?
- Welche Teststatistik ergibt sich durch Quadrieren aus einer Anderen?
- IaW Welcher Test ist bei einfacher lin. Reg. redundant?

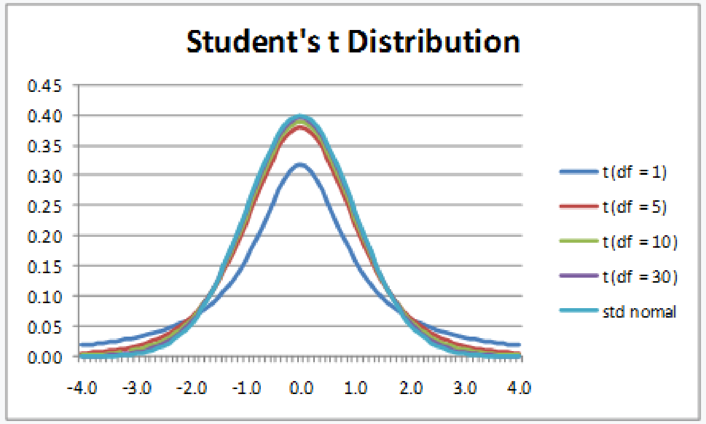
```
> summary(reg_hide_n_seek)

Call:
lm(formula = Y_made_up ~ X)

Residuals:
   Min      1Q  Median      3Q     Max 
-8.124 -1.722   0.016   1.673  11.436 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  0.02729    0.10056   0.271    0.786    
X            0.47541    0.02220  21.412   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.472 on 608 degrees of freedom
Multiple R-squared:  0.4299,	Adjusted R-squared:  0.429 
F-statistic: 458.5 on 1 and 608 DF,  p-value: < 2.2e-16
```



CAPM

- Hypothesentest
- Wir wollen die H0: Beta = 0 testen.
- Was ist unsere Prüfgrösse?
- Signifikanzniveau sei 5%.
- Bei welchem krit. Wert verwerfen wir H0?
- Welche (mind. 3) H0 können mit dem R Output getestet werden?
- Welche Teststatistik ergibt sich durch Quadrieren aus einer Anderen?
- IaW Welcher Test ist bei einfacher lin. Reg. redundant?

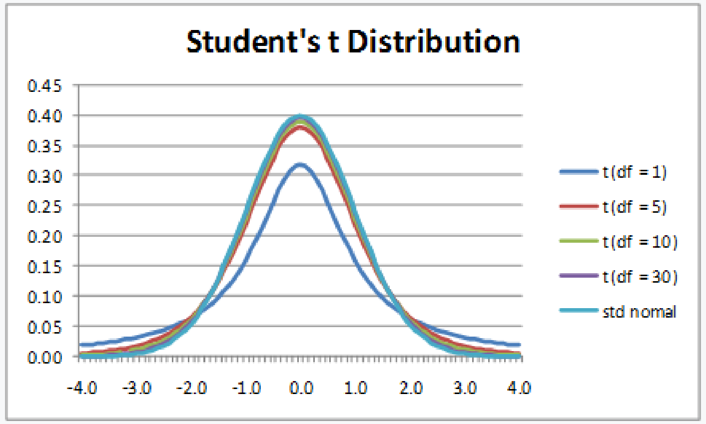
```
> summary(reg_hide_n_seek)

Call:
lm(formula = Y_made_up ~ X)

Residuals:
    Min      1Q  Median      3Q     Max
-8.124 -1.722   0.016   1.673 11.436

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.02729    0.10056   0.271    0.786
X            0.47541    0.02220  21.412   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.472 on 608 degrees of freedom
Multiple R-squared:  0.4299,    Adjusted R-squared:  0.429
F-statistic: 458.5 on 1 and 608 DF,  p-value: < 2.2e-16
```



CAPM

- Hypothesentest
- Wir wollen die H0: Beta = q = hier 0.5 testen.
- Signifikanzniveau sei 5%.
- Teststatistik ist: t = (Beta - q) / SE = -1.107567
- Bei welchem krit. Wert verwerfen wir H0?
- Hinweis: Mit denselben t-Test-Statistiken kann man auch H0: Beta>q oder Beta<q testen. Mehr dazu in der Literatur.

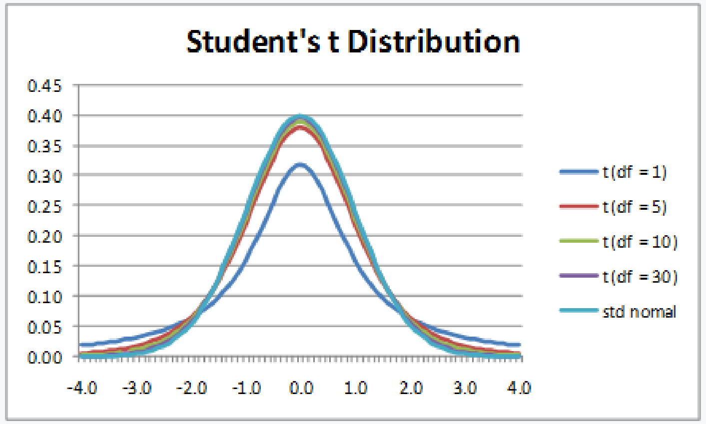
```
> summary(reg_hide_n_seek)

Call:
lm(formula = Y_made_up ~ X)

Residuals:
    Min      1Q  Median      3Q     Max 
-8.124 -1.722   0.016   1.673  11.436 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  0.02729    0.10056   0.271    0.786    
X            0.47541    0.02220  21.412   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.472 on 608 degrees of freedom
Multiple R-squared:  0.4299,	Adjusted R-squared:  0.429 
F-statistic: 458.5 on 1 and 608 DF,  p-value: < 2.2e-16
```



CAPM

- Hypothesentest
- H0: Beta = 0.5 behalten wir bei.
- Was ist mit H0: Beta = q = nun 0.45?
- Signifikanzniveau sei 5%.
- Teststatistik ist: t = (Beta - q) / SE = 1.144685
- Bei welchem krit. Wert verwerfen wir H0?
- Bedeutet das Nicht-Verwerfen einer H0 ihren Beweis?
- Zu welcher Kuriosität würde dies hier führen?

```
> summary(reg_hide_n_seek)

Call:
lm(formula = Y_made_up ~ X)

Residuals:
   Min     1Q Median     3Q    Max 
-8.124 -1.722  0.016  1.673 11.436 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  0.02729    0.10056   0.271    0.786    
X            0.47541    0.02220  21.412   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.472 on 608 degrees of freedom
Multiple R-squared:  0.4299,	Adjusted R-squared:  0.429 
F-statistic: 458.5 on 1 and 608 DF,  p-value: < 2.2e-16
```



CAPM

- Nun machen wir ein weiteres Experiment, indem wir eine Aktienrendite künstlich gemäss CAPM erzeugen
- Diesmal gibt es einen Strukturbruch

```
set.seed(2023)
Y_made_up_break = Y_made_up
Y_made_up_break[300:length(X)] = 1.5*X[300:length(X)]+eps[300:length(X)]
```

- Worin besteht er?

- Was wird bei dieser lin. Reg. rauskommen?

CAPM

- Ausführen 65-72
- Interpretieren Sie!

```
> summary(reg_hide_n_seek_break)

Call:
lm(formula = Y_made_up_break ~ X)

Residuals:
     Min       1Q   Median       3Q      Max
-12.4397  -2.0743   0.2054   2.2157  11.8803

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.09565    0.13625   0.702    0.483
X            1.01340    0.03008  33.686   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.349 on 608 degrees of freedom
Multiple R-squared:  0.6511,	Adjusted R-squared:  0.6506
F-statistic:  1135 on 1 and 608 DF,  p-value: < 2.2e-16
```

- Wir führen einen Dummy als weiteren Regressor ein, der Null bis zum Bruch ist und danach wie die Marktportfolioüberrendite
- Was wird bei dieser lin. Reg. rauskommen?

CAPM

- Ausführen 74-78
- Interpretieren Sie!
- Was schätzt der Koeff von D_period?

```
> summary(reg_hide_n_seek_break_d)

Call:
lm(formula = Y_made_up_break ~ X + D_period)

Residuals:
    Min      1Q  Median      3Q     Max 
-8.1242 -1.7219  0.0159  1.6729 11.4358 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  0.02729    0.10069   0.271    0.786    
X            0.47547    0.03262  14.575   <2e-16 ***
D_period     0.99990    0.04439  22.523   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.474 on 607 degrees of freedom
Multiple R-squared:   0.81,	Adjusted R-squared:  0.8093 
F-statistic:  1294 on 2 and 607 DF,  p-value: < 2.2e-16
```

- Wir führen einen Dummy ein, der im Januar 1 ist und 0 sonst
- Was wird bei dieser lin. Reg. rauskommen?

CAPM

- Ausführen 84-86
- Generiert der Januar alpha?

```
> summary(reg_jan)

Call:
lm(formula = Y ~ X + D_january)

Residuals:
     Min       1Q   Median       3Q      Max 
-12.8203  -1.6622  -0.0673   1.5535  11.7772 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) -0.12246    0.12031  -1.018    0.309    
X            1.16683    0.02548  45.797   <2e-16 ***
D_january    0.60354    0.41494   1.455    0.146    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.835 on 607 degrees of freedom
Multiple R-squared:  0.7763,	Adjusted R-squared:  0.7756 
F-statistic:  1054 on 2 and 607 DF,  p-value: < 2.2e-16
```

- Wichtig: Haben nun zwei Ansätze Strukturbrüche bei alpha und beta zu ermitteln
- Wenn wir nicht Wissen, wo genau Bruch liegt, empfiehlt sich rekursive Paraschätzung

CAPM

- Hypothesentest
- Übliche Niveaus nach Schröder, 2002, 68
- Welche Zahl erkennen Sie wieder?

	rechtsseitig	rechtsseitig	zweiseitig	zweiseitig
Signifikanzniveau	FG 30	FG 100	FG 30	FG 100
1%	2,46	2,36	2,75	2,63
5%	1,70	1,66	2,04	1,98

- Auffällig war, dass alle unsere Schätzergebnisse nahe der Wahrheit waren
- Versteckten Modellparametern & ihren Änderungen (Strukturbruch) kamen wir auf die Schliche
- Ist dies Zufall?

Gauss Markov

- Es ist sehr wichtig, zwischen Residuen und Störtermen zu unterscheiden!

- Wir illustrieren dies an der einfachen Regression, die wir mit der Methode der KQ/OLS schätzen

- Annahme (a0)
  - (w) Das wahre Modell sei: $y = a + bx + e$
  - (g) Das geschätzte Modell sei: $y = \alpha + \beta x + u$

- Der Störterm ist die Zufallsvariable e und das Residuum ist u
- Wir haben n Beobachtungen und indizieren diese mit i = 1, …, n.

- *Welche der beiden Gleichungen generiert die Daten der Grundgesamtheit (population)?*

- *Welche liegt hinter/steuert die Wirklichkeit?*

- *Welche bedient sich/basiert auf der Stichprobe (sample)?*

Gauss Markov

Es gilt der folgende **Satz von Gauss und Markov** (Verbeek, 2017, 15)

Wenn für die N Störterme e_i die folgenden Gauss-Markov Bedingungen gelten:

- (a1) $E[e_i]$=0, i=1,…,n (Im Mittel heben sich die Störungen weg, ignorierte Faktoren heben sich ggs auf)
- (a2) e und x sind unabhängig (Exogenität von x: Nur y durch Modell erklärt, x von ausserhalb)
- (a3) $Var[e_i]$=σ^2 (Homoskedastizität)
- (a4) $Cov[e_i, e_j]$=0, $i \neq j$ (Keine Autokorrelation im Störterm)
- (a5) Im Falle einer **multiplen** Regression: Keine lineare Abhängigkeit in x

Dann sind die **Schätzer** $\boldsymbol{\alpha}, \boldsymbol{\beta}$ die

- **B**esten (Sie haben die kleinste Varianz …)
- **L**inearen (unter allen linearen … Schätzern)
- **U**nverzerrten Schätzer (Erwartungstreue) (E[ß]=b und E[α]=a, d.h. wir schätzen im Mittel richtig)

für die wahren Parameter a, b. Das nennt man auch die **BLUE Eigenschaft der Kleinste Quadrate Schätzer** $\boldsymbol{\alpha}, \boldsymbol{\beta}$.

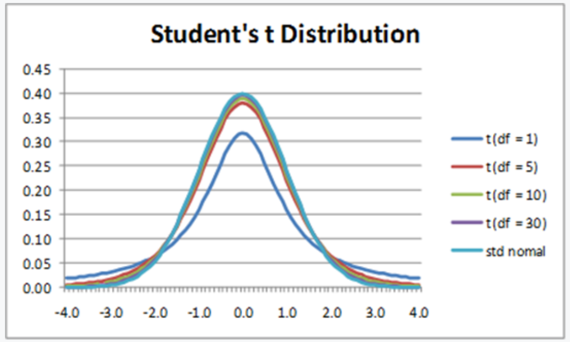


Illustration von B und U anhand Glocke

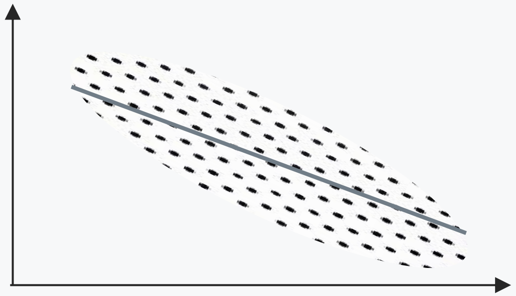
CAPM & Regressionsdiagnostik

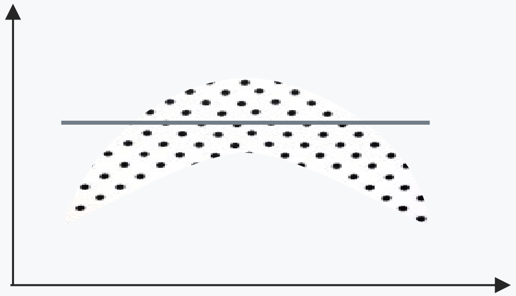
- Für die Beta-Schätzung des Bausektors wollen wir nun prüfen, ob die GM Annahmen erfüllt sind. Wir fangen mit a0 an.

- Denn im Fall von Nichtlinearität liefert die Regressionsgerade nicht mehr die besten Schätzer (d. h. sie minimieren nicht mehr den Abstand zwischen tatsächlichen und geschätzten Werten).

- Die Folge ist eine Verzerrung der Schätzwerte

- Für eine einfache lineare Regression (eine unabhängige Variable) kann die Prüfung über ein Streudiagramm der abhängigen (y-Achse) gegen die unabhängige (x-Achse) Variable erfolgen. Sind in der Punktewolke Muster zu beobachten, so spricht dies gegen einen linearen Zusammenhang von unabhängiger und abhängiger Variable.





- Mit welcher Variablentrafo könnte man das rechte Problem linearisieren?

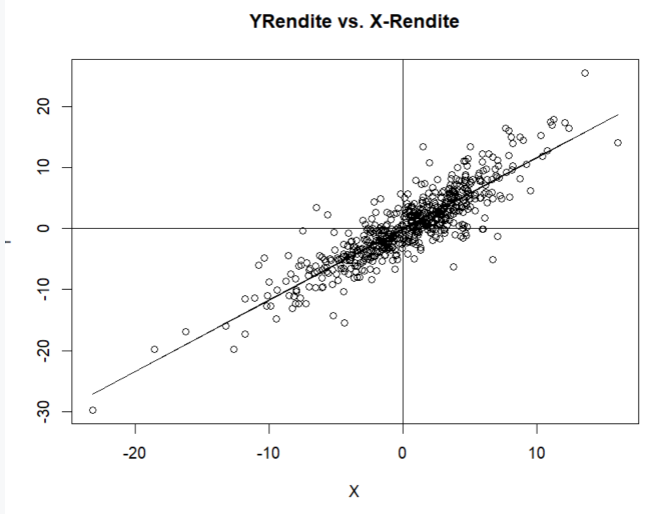
CAPM & Regressionsdiagnostik

- Der REgression Specification Error Test (RESET-Test) von Ramsey testet allgemein auf eine Fehlspezifikation eines Regressionsmodells. Ursachen für eine Fehlspezifikation können entweder fehlende wichtige erklärende Variablen oder ein nichtlinearer Zusammenhang zwischen abhängiger und unabhängigen Variablen sein.

- Ramsey konnte zeigen, dass die fitted values in potenzierter Form eine geeignete Näherung für eine Fehlspezifikation darstellen, wobei eine Potenzierung bis zur vierten Potenz als ausreichend angesehen wird.

- Ausführen 90-96

```
> resettest(reg)#H0 No spec error

        RESET test

data:  reg
RESET = 4.5204, df1 = 2, df2 = 606, p-value = 0.01125
```



- Was machen Sie mit H0?

CAPM & Regressionsdiagnostik

- Die Diagnostik wird auch Residualanalyse genannt, weil wir aus den Egenschaften der Residuen auf Störtermeigenschaften schliessen können.
- Das haben wir schon getan, als das Sigma des Störterms als Summe Residuenquadrate / FG geschätzt wurde!
- Kennt man die Mathe hinter dem was wir tun, so ist klar, dass bei einer lin. Reg. mit Achsenabschnitt das Mittel der Residuen nahe 0 ist
- Zudem steht der Residualvektor senkrecht auf ßX (Theo. Beste Approx.)
- Deshalb helfen uns die Residuen bei a1 und a2 nicht, obwohl a2 sehr wichtig ist! Das müssen wir durch unseren Ansatz sicher stellen. So sollte die Aktie nicht allzu prominent im Marktportfolio sein, weil sonst endogen. Weites Feld!
- Bei a3 und a4 lohnt eine Analyse der Residuen

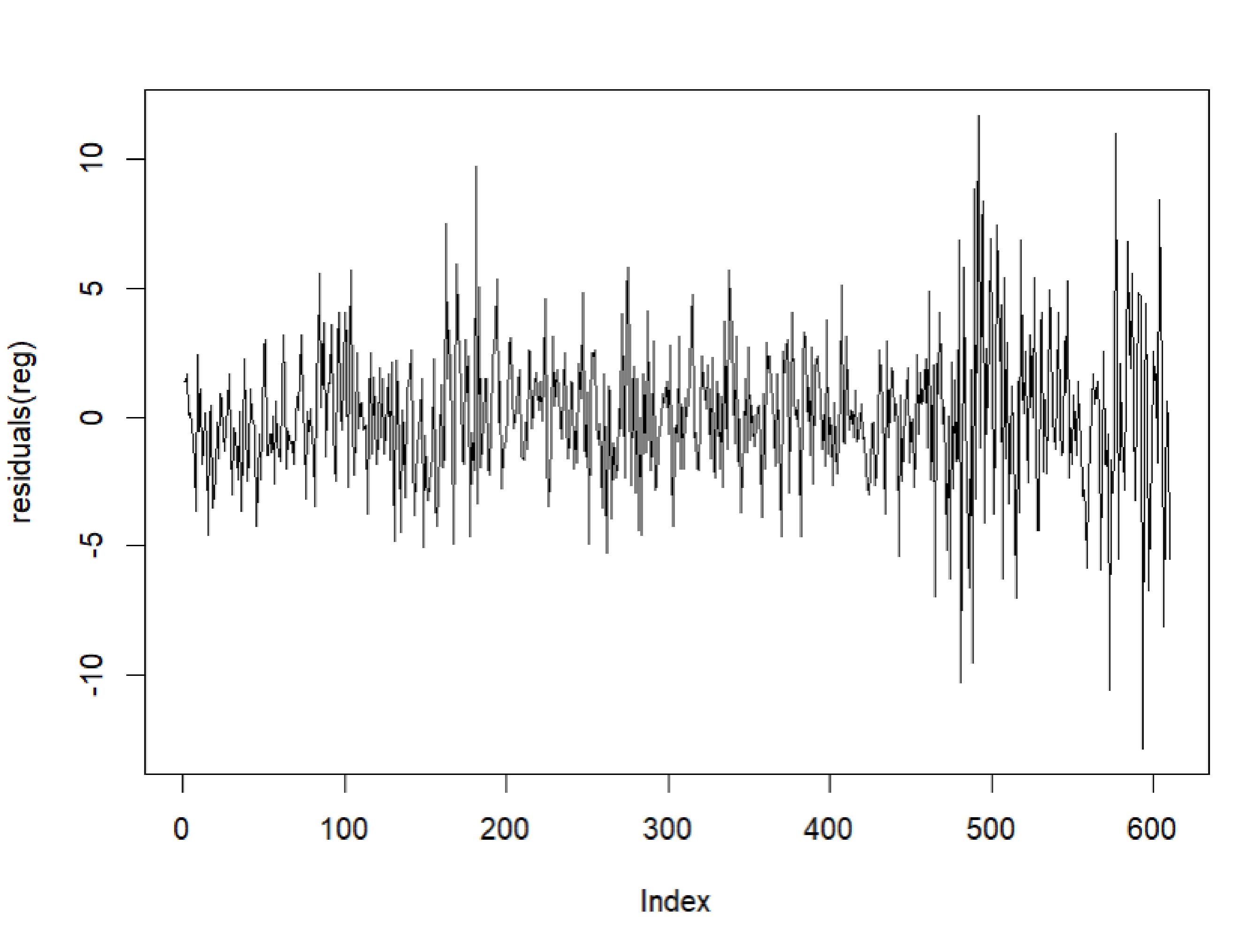
CAPM & Regressionsdiagnostik

- Bei a3 lohnt eine Analyse der Residuen
- Ausführen 101-106
- Der White-Test kann auch bei nicht-normalverteilten Residuen angewendet werden.
- Analog zum Breusch-Pagan-Test wird die Residuenvarianz über eine Hilfsregression modelliert.
- Als Nullhypothese wird bei beiden formuliert: Homoskedastizität.
- Interpretieren Sie!

```
> bptest(reg)#H0: Homoscedasticity!

	studentized Breusch-Pagan test

data:  reg
BP = 4.5821, df = 1, p-value = 0.03231

>
> #white, Gehrke S 71
> bptest(reg, ~I(X^2)) #H0: Homoscedasticity!

	studentized Breusch-Pagan test

data:  reg
BP = 11.794, df = 1, p-value = 0.0005941
```

CAPM & Regressionsdiagnostik

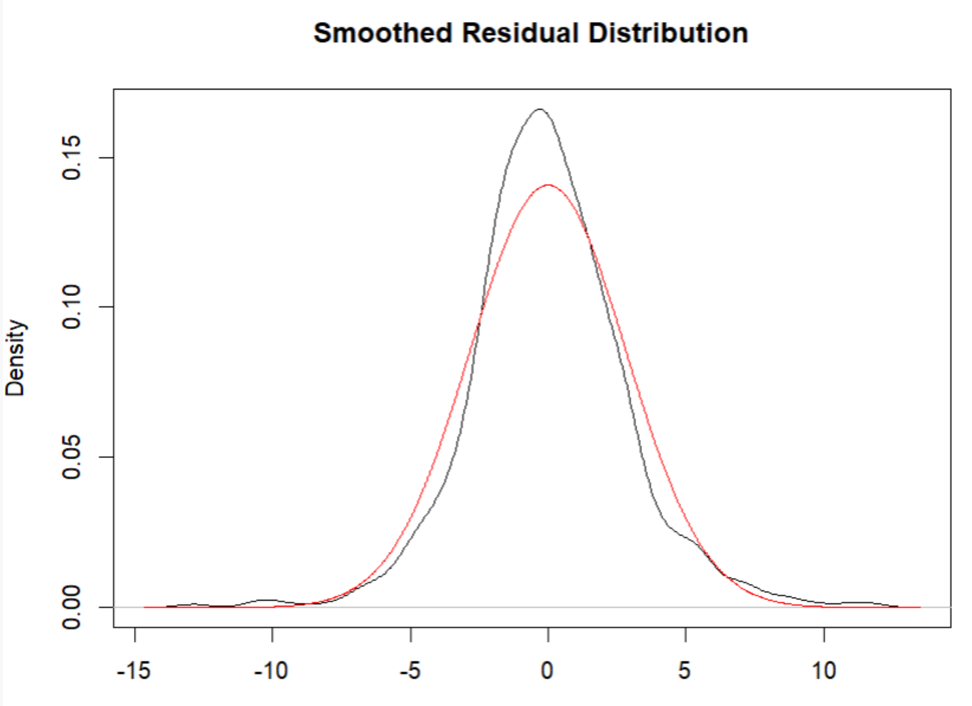
- Bei a3 lohnt eine Analyse der Residuen
- Wie wirkt die Grafik auf Sie?
- Selbe Urne jeden Monat?



- Stylized Facts of Returns:
- Hetero & leptokurtic

CAPM & Regressionsdiagnostik

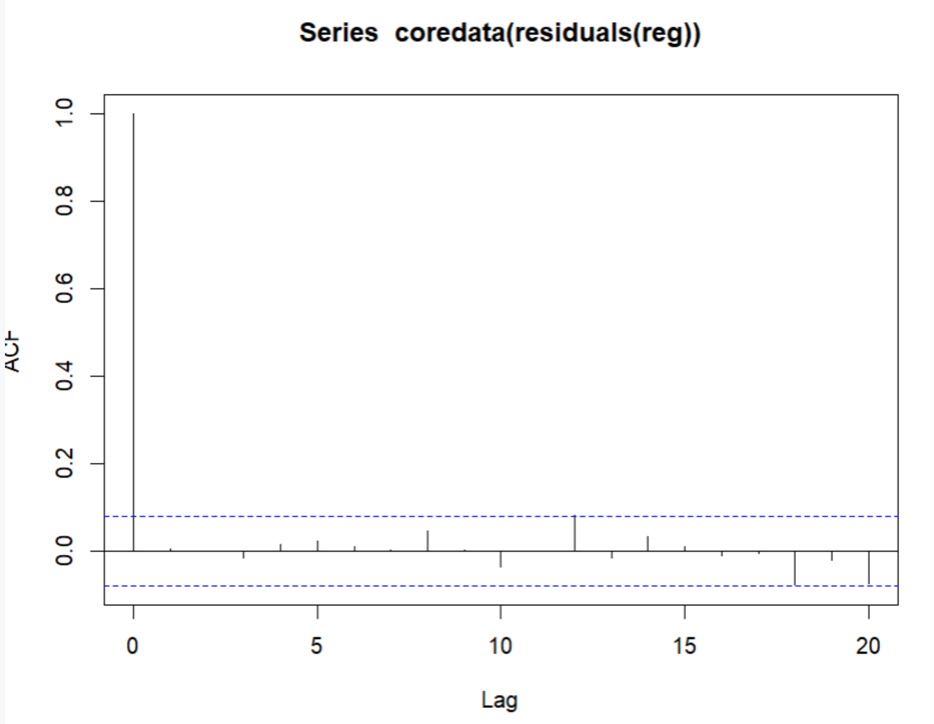
- Ausführen 108-118
- Stylized Facts of Returns:
- Hetero & leptokurtic
- Was wird hier bestätigt?
- Wie sinnvoll ist das Entfernen von Ausreissern?
- Wie heissen die an Finanzmärkten?
- Wie sinnvoll ist Risikomessung mit normalverteilten Returns?
- Was meint hier “black swans”?




```
> shapiro.test(residuals(reg))#H0: Is normal

        Shapiro-Wilk normality test

data:  residuals(reg)
W = 0.97474, p-value = 9.315e-09

> jarque.bera.test(residuals(reg))#same

        Jarque Bera Test

data:  residuals(reg)
X-squared = 118.89, df = 2, p-value < 2.2e-16
```

CAPM & Regressionsdiagnostik

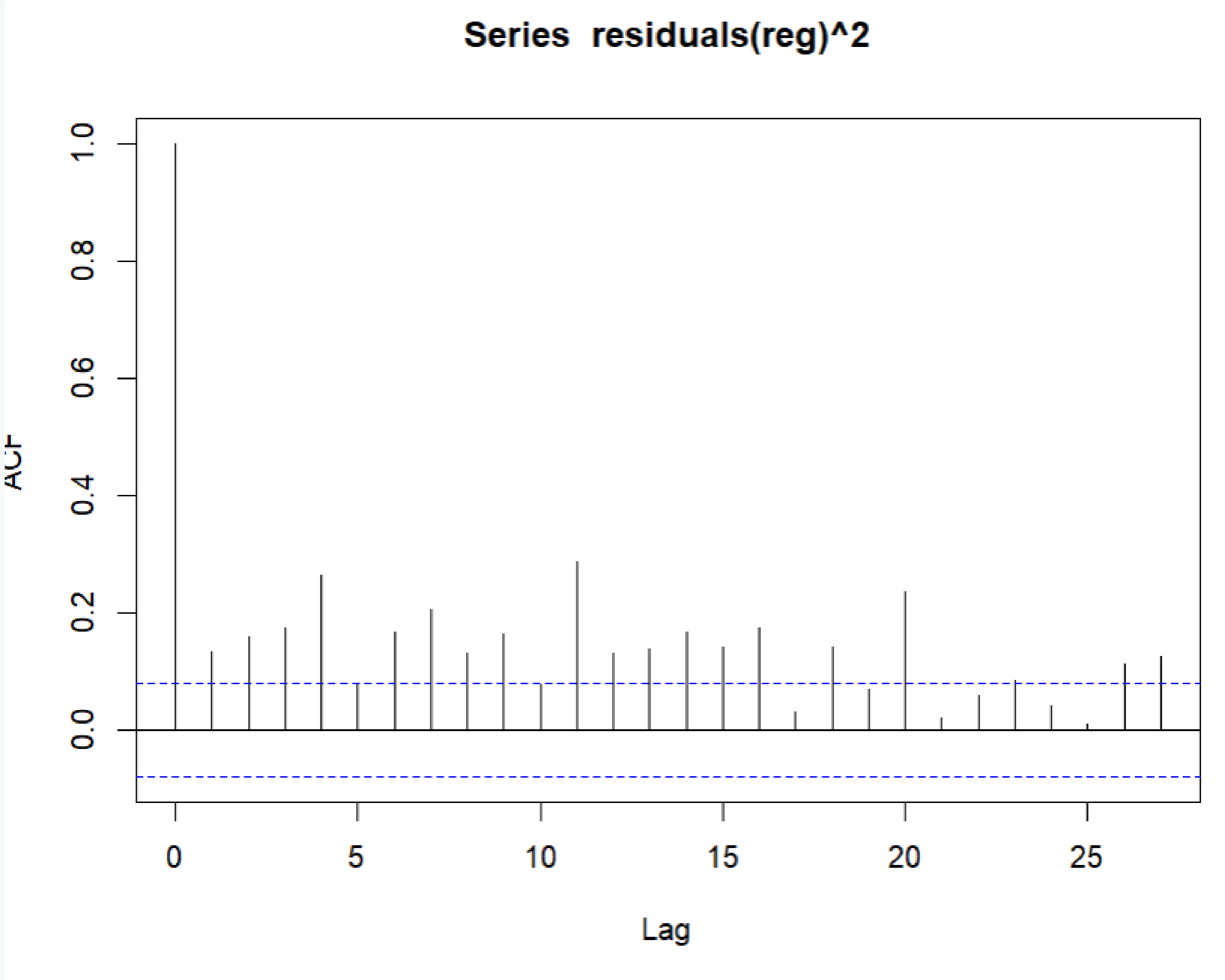
- Auch bei a4 lohnt eine Analyse der Residuen
- Ausführen 121-126
- DW testet nur auf Lag 1
- ACF sagt schon viel
- Interpretieren Sie!
- Weiteres stylized fact: Lohnt technisches Trading Day to Day /Month same?




```
> dwtest(reg)#H0: No autocorr

	Durbin-Watson test

data:  reg
DW = 1.9829, p-value = 0.4149
alternative hypothesis: true autocorrelation is greater than 0
```

CAPM & Regressionsdiagnostik

- Auch bei a4 lohnt eine Analyse der Residuen
- Ausführen 128
- Was hat sich geändert?
- Weiteres stylized fact: Vola Cluster
- Ziehen Sie das Fazit über die GM Annahmen!



Noch ein Satz

Es gilt bei large samples, Details bei Verbeek, 2017, 34 + 128:

Wenn für die N Störterme e_i die folgenden Bedingungen gelten:

- (a1) $E[e_i]$=0, i=1,…,n (Im Mittel heben sich Störungen* weg, ignorierte Faktoren heben sich ggs auf)
- (a2) e und x sind unabhängig (Exogenität von x: Nur y durch Modell erklärt, x von ausserhalb)
- (a3) $Var[e_i]$ variiert, d.h. Heteroskedastizität liegt vor, oder Homoskedastizität
- (a4) Autokorrelation im Störterm geht nach „wenigen" Lags auf Null
- (a5) Im Falle einer multiplen Regression: Keine lineare Abhängigkeit in x

Dann sind die **Schätzer** $\boldsymbol{\alpha, \beta}$

- **Konsistente** Schätzer (d.h. ß konvergiert in WS gg b, analog α gg a;
  formal lim $_{n\text{->}inf}$ P(|b-ß$_n$|>d) = 0 für d > 0 (beliebig klein!)
- Die t-Werte können auf Basis der Heteroscedasticity & Autocorrelation Consistent Standard Errors (**HAC SE**) als **normalverteilt** interpretiert werden

Best gilt nicht mehr, weil es bessere Modellansätze gibt, z B Cochrane … GARCH ….

CAPM

- Führen Sie Zeilen 132-136 aus!
- Wie hoch ist das Beta? Ist es nun anders?
- Wie verlässlich ist es?
- Ihr Chef möchte ein 95% Intervall um den Punktschätzer. Wie rechnen Sie?
- Was haben Sie dabei vom Satz benutzt?
- Was machen Sie nun mit H0: Beta = 0?
- Und H0: Alpha = 0?

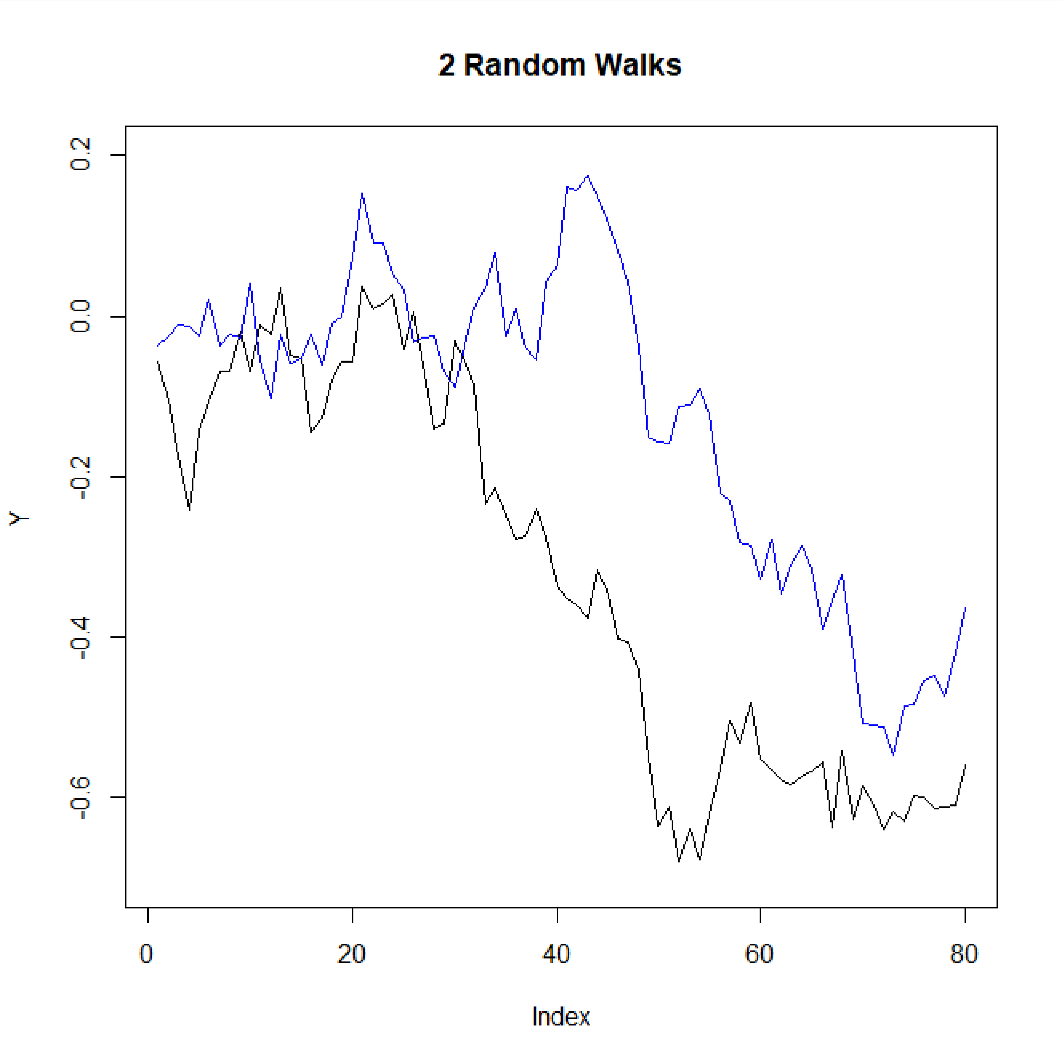
```
> #Newey-West-Korrektur demzufolge sinnvoll
> coeftest(reg)

t test of coefficients:

             Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.072585   0.115424 -0.6289   0.5297
X            1.168168   0.025485 45.8372   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

> #vs
> coeftest(reg,vcov=NeweyWest(reg))#now use heteroskedasticity and autocorrelation

t test of coefficients:

             Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.072585   0.115715 -0.6273   0.5307
X            1.168168   0.034186 34.1706   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

> #consistent (HAC) covariance matrix estimators. Note the decreased t value!
```

Spuriosities

- Zeit für ein weiteres Experiment
- Wir erzeugen Y und X wie folgt

  set.seed(129) #für Optik
  Y=cumsum(rnorm(80,0,0.05))
  X=cumsum(rnorm(80,0,0.05))

- Wieviel hat Y mit X zu tun?
- Gilt “Y ist Fkt von X”, d.h. Y(X)?
- Was suggeriert die Grafik (X ist blau)?
- Was wird bei lin. Reg. rauskommen?



Spuriosities

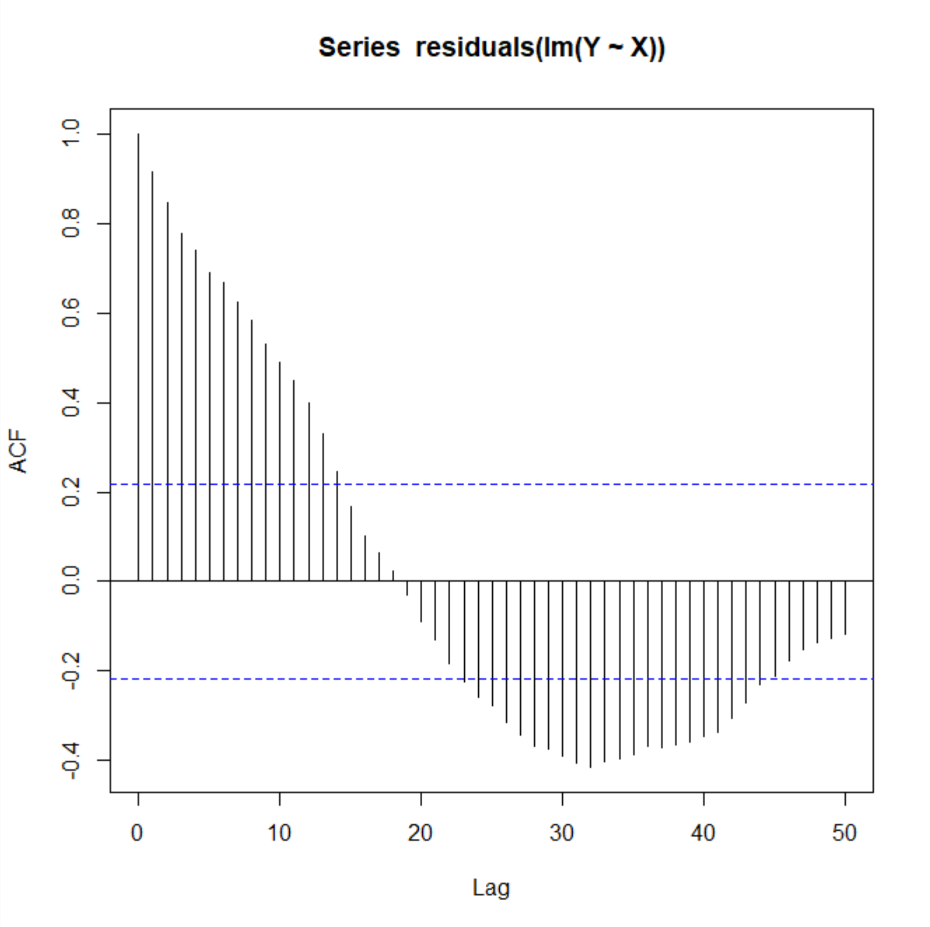
- Interpretieren Sie!
- Was ist noch zu tun?

```
> summary(lm(Y~X))

Call:
lm(formula = Y ~ X)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.37932 -0.09732  0.01488  0.14310  0.28491 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) -0.21704    0.02263  -9.590 7.79e-15 ***
X            0.89103    0.09820   9.073 7.78e-14 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1714 on 78 degrees of freedom
Multiple R-squared:  0.5135,	Adjusted R-squared:  0.5073 
F-statistic: 82.33 on 1 and 78 DF,  p-value: 7.776e-14
```

Spuriosities

- Interpretieren Sie!
- Helfen hier HAC SE? Verschwinden die Autokorr.?
- Und warum mache ich eine Reg auf Diff?

```
> summary(lm(diff(Y)~diff(X)))

Call:
lm(formula = diff(Y) ~ diff(X))

Residuals:
      Min        1Q    Median        3Q       Max 
-0.141399 -0.029254  0.002105  0.039332  0.110149 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.006306   0.005871  -1.074    0.286
diff(X)      0.015653   0.119916   0.131    0.896

Residual standard error: 0.05199 on 77 degrees of freedom
Multiple R-squared:  0.0002212, Adjusted R-squared:  -0.01276 
F-statistic: 0.01704 on 1 and 77 DF,  p-value: 0.8965
```



- Nachlesen bei Gehrke (2022, 360ff)

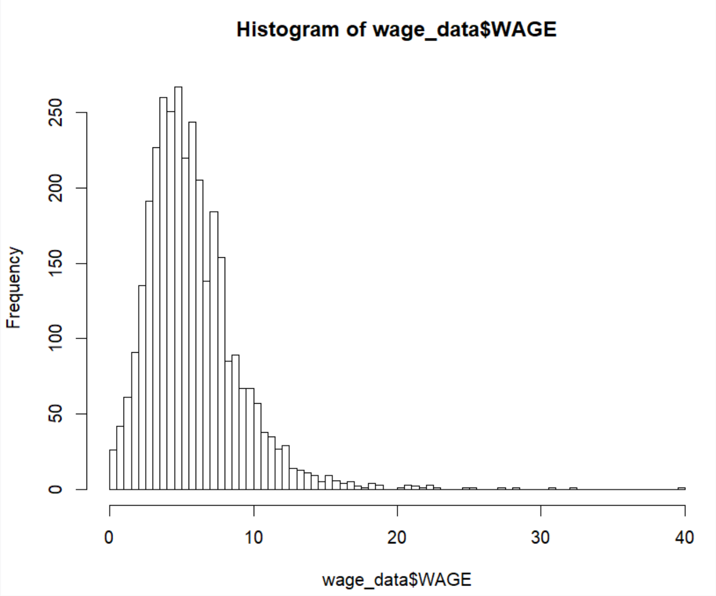
Multiple lineare Regression

- Wording: Wenn Y von mehreren X abhängt, ist es multipel. Wenn Y mehrdimensional wird, ist es multivariate Regression z B SURE
- Wir besprechen ein Bsp. aus Verbeek, 2017, 28
- Coding heisst GML_02_02 wages.R
- Ausführen 1-30 ergibt:
- Y = wage/hr in USD
- Beschreiben Sie den Datensatz!
- Angenommen die Firma hat 100 Tsd Beschäftigte und wir haben diese Stipo an mehreren Werkstoren genommen
- Diskriminiert diese Firma Frauen? Wie würden Sie bei Gericht argumentieren?

```
> head(wage_data)
  EXPER MALE SCHOOL     WAGE
1     9    0     13 6.315296
2    12    0     12 5.479770
3    11    0     11 3.642170
4     9    0     14 4.593337
5     8    0     14 2.418157
6     9    0     14 2.094058
> dim(wage_data)
[1] 3294    4
> summary(wage_data)
     EXPER             MALE            SCHOOL           WAGE
 Min.   : 1.000   Min.   :0.0000   Min.   : 3.00   Min.   : 0.07656
 1st Qu.: 7.000   1st Qu.:0.0000   1st Qu.:11.00   1st Qu.: 3.62157
 Median : 8.000   Median :1.0000   Median :12.00   Median : 5.20578
 Mean   : 8.043   Mean   :0.5237   Mean   :11.63   Mean   : 5.75759
 3rd Qu.: 9.000   3rd Qu.:1.0000   3rd Qu.:12.00   3rd Qu.: 7.30451
 Max.   :18.000   Max.   :1.0000   Max.   :16.00   Max.   :39.80892
```

Multiple lineare Regression

- Schrittweiser Analyseprozess
- Visualisierungen als warm up
- Führen Sie Schritte 2-3 aus
- Was sehen Sie?



Multiple lineare Regression

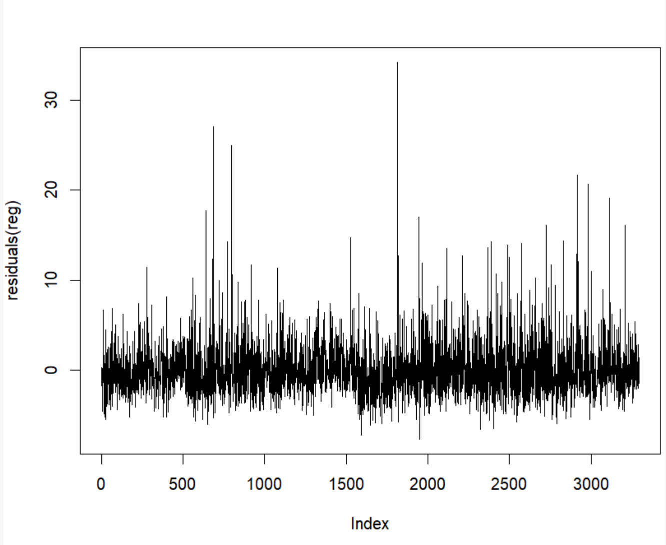
- Schrittweiser Analyseprozess
- Führen Sie Schritt 4 aus
- Was ergibt eine naive ad hoc Betrachtung?
- Führen Sie Schritte 5-7 aus
- Wie ändert sich das Bild, wenn Sie für andere Faktoren kontrollieren
- Was ist noch zu tun?

```
> mean(wages_males)-mean(wages_females)
[1] 1.166097
```

```
> summary(regMult)

Call:
lm(formula = wage_data$WAGE ~ wage_data$SCHOOL + wage_data$MALE +
    wage_data$EXPER)

Residuals:
   Min     1Q Median     3Q    Max
-7.654 -1.967 -0.457  1.444 34.194

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)      -3.38002    0.46498  -7.269 4.50e-13 ***
wage_data$SCHOOL  0.63880    0.03280  19.478  < 2e-16 ***
wage_data$MALE    1.34437    0.10768  12.485  < 2e-16 ***
wage_data$EXPER   0.12483    0.02376   5.253 1.59e-07 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.046 on 3290 degrees of freedom
Multiple R-squared:  0.1326,	Adjusted R-squared:  0.1318
F-statistic: 167.6 on 3 and 3290 DF,  p-value: < 2.2e-16
```

Multiple lineare Regression

- Führen Sie Schritt 8 aus
- Nun haben wir Querschnittsdaten, 1 Zeitpunkt 5 Werktore. Kein Platz für serielle / AutoKorrelation!
- Welche Korrelationen im Störterm könnte es geben? Denken Sie daran, was alles in eps steckt!
- Wie beurteilen Sie H0: Homoskeda.?
- Welche Tabelle würden Sie dem Gericht vorstellen?
- Was kann das Gericht noch monieren?

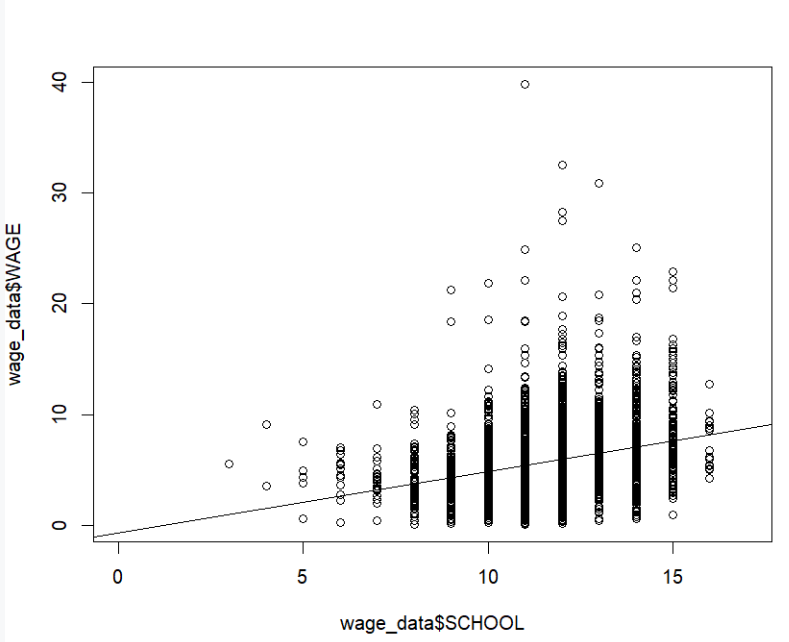



```
> coeftest(reg)

t test of coefficients:

                   Estimate Std. Error t value  Pr(>|t|)
(Intercept)       -3.380018   0.464977 -7.2692 4.496e-13 ***
wage_data$SCHOOL   0.638798   0.032796 19.4780 < 2.2e-16 ***
wage_data$MALE     1.344369   0.107676 12.4853 < 2.2e-16 ***
wage_data$EXPER    0.124825   0.023763  5.2530 1.592e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> #vs
> coeftest(reg,vcov=NeweyWest(reg))#now use heteroskedasticity and autocorrelation

t test of coefficients:

                   Estimate Std. Error t value  Pr(>|t|)
(Intercept)       -3.380018   0.557540 -6.0624 1.493e-09 ***
wage_data$SCHOOL   0.638798   0.037951 16.8323 < 2.2e-16 ***
wage_data$MALE     1.344369   0.134770  9.9753 < 2.2e-16 ***
wage_data$EXPER    0.124825   0.024799  5.0335 5.076e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Multiple lineare Regression

- Führen Sie Schritt 6 aus

```
Call:
lm(formula = wage_data$WAGE ~ wage_data$SCHOOL)

Residuals:
   Min     1Q Median     3Q    Max 
-6.744 -2.024 -0.482  1.443 34.403 

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)    
(Intercept)      -0.72251    0.38739  -1.865   0.0623 .  
wage_data$SCHOOL  0.55716    0.03298  16.896   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.137 on 3292 degrees of freedom
Multiple R-squared:  0.0798,	Adjusted R-squared:  0.07952 
F-statistic: 285.5 on 1 and 3292 DF,  p-value: < 2.2e-16
```



- Personen mit höherer Intelligenz iwS werden wohl länger auf der Schule gewesen sein
- Was vermuten Sie bzgl der Lohnhöhe?
- Das heisst u U, dass der Mehrverdienst durch ein weiteres Schuljahr (Grafik) nicht so stark wirkt wie geschätzt. Kausale Interpretation ist also schwierig. Wenn Du ein Jahr länger …
- Hin zur Lösung: Nimm IQ o.a. mit auf. Würde auch R^2 erhöhen ….

Multiple lineare Regression

- Lesen Sie Verbeek, 2017, 146 ff. sowie 162ff. für die Nachbereitung

- Das im Jahr 2020 erschienene Paper von Karsten Lübke, Matthias Gehrke, Jörg Horst und Gero Szepannek über Kausale Inferenz als Bestandteil der Data-Literacy Ausbildung (Why We Should Teach Causal Inference: Examples in Linear Regression With Simulated Data, https://doi.org/10.1080/10691898.2020.1752859) gehört als meistgelesenes Paper des Jahrgangs 28 des Journal of Statistics and Data Science Education zur "Most Read Collection" der American Statistical Association (https://think.taylorandfrancis.com/2021-asa-most-read/)

- „We find … ‚correlation does not mean causation' … to be grossly misleading. Causation manifests itself in correlation" (Cohen*, 1983, 15).

*Cohen,Jacob & Patricia, „Applied Regression/Correlation Analysis for the Behavioral Sciences", 2nd ed, 1983

Multiple lineare Regression

- Zuweilen vereinfacht*:

Arbeiten Sie Ch. 2-4 in Verbeek (2017) und

Kap. 2 in Gehrke (2022) nach. D.h. Lesen S. 8-101 und begleitend R ausführen.

Zu Beginn der nächsten V. geben Sie mir Feedback.

Insbes. Q&A.

Von einem kausalen Zusammenhang zwischen einer unabhängigen und einer abhängigen Variable kann man ausgehen, wenn
a) zwischen der unabhängigen Variable und der abhängigen Variable ein statistischer Zusammenhang besteht,
b) die Veränderung der unabhängigen Variable der der abhängigen Variable zeitlich vorausgeht, und
c) alternative Erklärungen für den statistischen Zusammenhang ausgeschlossen werden können.

Wird zum Beispiel vermutet, dass die wahrgenommene Produktqualität ein Treiber der Kundenzufriedenheit ist, so
a) müssen Produktqualität und Zufriedenheit miteinander korrelieren,
b) muss sich die Produktqualität verändert haben, bevor sich die Zufriedenheit verändert, und
c) dürfen Veränderungen von Produktqualität und Zufriedenheit nicht auf andere (dritte) Variablen zurückzuführen sein,

damit der kausale Zusammenhang zwischen den beiden Variablen bestätigt wird.

* Lüken/Schimmelpfennig, „Einführung in Kausalanalysen“, planung & analyse 01/2013

Multiple lineare Regression

- Wie könnte eine Klausuraufgabe aussehen?

3 Thema Lineares Regressionsmodell (20 Punkte)

Gegeben ist ein lineares Regressionsmodell in üblicher Notation wie folgt $y = \alpha+\beta*x+\varepsilon$

Für den letztgenannten Störterm werden die folgenden Annahmen gemacht:

- Der Erwartungswert der Störgröße ist Null.
- Die Störgrößen dürfen nicht untereinander korrelieren (keineAutokorrelation).
- Die Varianz der Störgrößen erfüllt das Kriterium der Homoskedastizität, sie ist konstant und endlich.
- Störgrößen und Regressoren sind voneinander unabhängig.

a. Wie nennt man diese Annahmen? (2 Punkte)

b. Was bedeutet es explizit, das die OLS (KQ) Schätzfunktionen die BLUE Eigenschaft haben? Was kürzt dies ab? Und wofür steht jede der abgekürzten Eigenschaften? (8 Punkte)

Multiple lineare Regression

- Wie könnte eine Klausuraufgabe aussehen?

Sie erhalten einen Datensatz mit Informationen über 1000 Beschäftigte in einem Betrieb, die alle einer ähnlichen Tätigkeit nachgehen. 500 Beschäftigte sind weiblich. Der Datensatz enthält Jahreslohn, Alter, Gewicht und Geschlecht (1=männlich, 0 sonst) je Beschäftigten. Sie regressieren in einem linearen Regressionsmodell den Jahreslohn auf die übrigen Grössen sowie eine Konstante und erhalten folgende t-Werte:

Achsenabschnitt (Konstante)	2,5
Alter	3
Gewicht	-0,3
Geschlecht	2,4

c. Kommentieren Sie je Regressor (ohne Achsenabschnitt) seine Signifikanz (3 Punkte)

d. Wie beurteilen Sie die Behauptung des Managements, dass in diesem Betrieb keine Geschlechtergruppe bei der Bezahlung bevorzugt wird? (2 Punkte)

e. Welche Diagnostik ist für eine Abrundung Ihrer Untersuchung erforderlich? (5 Punkte)

Multiple lineare Regression

- Wie könnte eine Klausuraufgabe aussehen?

Sie erhalten einen Datensatz mit Informationen über 1000 Beschäftigte in einem Betrieb, die alle einer ähnlichen Tätigkeit nachgehen. 500 Beschäftigte sind weiblich. Der Datensatz enthält Jahreslohn, Alter, Gewicht und Geschlecht (1=männlich, 0 sonst) je Beschäftigten. Sie regressieren in einem linearen Regressionsmodell den Jahreslohn auf die übrigen Grössen sowie eine Konstante und erhalten folgende t-Werte:

Achsenabschnitt (Konstante)	2,5
Alter	3
Gewicht	-0,3
Geschlecht	2,4

c. Kommentieren Sie je Regressor (ohne Achsenabschnitt) seine Signifikanz (3 Punkte)

d. Wie beurteilen Sie die Behauptung des Managements, dass in diesem Betrieb keine Geschlechtergruppe bei der Bezahlung bevorzugt wird? (2 Punkte)

e. Welche Diagnostik ist für eine Abrundung Ihrer Untersuchung erforderlich? (5 Punkte)

Multiple lineare Regression

- Versuch eines wrap ups
- Details bei
  - Verbeek (2017)
  - Gehrke (2022)

Voraussetzung	Voraussetzungs-verletzung	Konsequenzen	Was bleibt?	Was könnte man tun?
Linearität der Parameter	Nichtlinearität	Verzerrung der Schätzwerte		Andere funktionale Form
Vollständigkeit des Modells	Unvollständigkeit	Verzerrung der Schätzwerte		Mehr Regressoren
Homoskedastizität der Störterme	Heteroskedastizität	• Ineffizienz: Es gibt bessere Schätzfunktion! • SE falsch!	LU // Konsistenz	• HAC SE nutzen • Heteroskedastizität explizit modellieren z B GARCH …
Unabhängigkeit der Störgrößen	Autokorrelation	• Ineffizienz: Es gibt bessere Schätzfunktion! • SE falsch!	LU // Konsistenz	• HAC SE nutzen • Autokorrelation explizit modellieren z B Cochrane …
Keine lineare Abhängigkeit zwischen den unabhängigen Variablen	Multikollinearität	Präzision der Schätzwerte geringer als nötig, d.h. SE korrekt aber zu hoch	BLU	Reduktion der Anzahl der Regressoren gemäss ökon. Vorverständnis oder Zielkriterium
Normalverteilung der Störgrößen	nicht normalverteilt	Ungültigkeit der Signifikanztests, wenn Stichprobe klein ist. Sonst Heilung durch ZGWS.	BLU	Stipo vergrössern. Wenn Stipo gross, kann man ML nutzen und dann mit anders verteiltem Störterm schätzen (s.o. z B GARCH).

Multiple lineare Regression

Zusammenhang zur Korrelation

- Bislang Schätzverfahren / Lernverfahren KQ = Min Squared Error

- Wir bezeichnen die Kovarianz mit Cov(.,.), die Standardabweichung mit SD(.) und die Varianz mit Var(.). Durchrechnen führt zu:
- Die Schätzfunktion(Daten) für den Parmater b ist die Kovarianz von x und y geteilt durch die Varianz von x:

$$\beta = \frac{Cov(y,x)}{Var(x)}$$

Für den Parameter a ist es Mittelwert von y minus b mal den Mittelwert von x:

$$\alpha = \bar{y}\text{-}\beta\bar{x}$$

Welcher Zusammenhang besteht zwischen ß und der Korrelation von y und x?

Was sagen Sie wenn Ihr Chef moniert, dass Sie nur eine lin. Reg. Und keine Korr. Analyse gemacht haben?

Multiple lineare Regression

Sie haben ein Modell geschätzt, Diagnostik betrieben und wollen die Präzision Ihrer Schätzer ß erhöhen

- Der Standardfehler von ß setzt sich aus drei Komponenten zusammen (Details bei Davidson & MacKinnon, 2009, 101). Formal gilt für die Varianz von ß:

$$Var(\beta) = \frac{\sigma^2}{X'X}$$

- Wie ist der SE von Beta definiert? Was gibt R in der zweiten Spalte aus?

- Explizit sieht man die Varianz des Störterms im Zähler
- Im Nenner versteckt sich die Varianz der Regressoren. Streuen diese so gut wie gar nicht, so vermindert sich die Präzision des Schätzers bzw man kann ihn nicht mehr ausrechnen.
  - -> Eventdummies mit vielen Events sind also besser. D.h. statt je Analystenempfehlung einen Dummy, fasse alle Kaufempfehlungen zusammen. Maximale Varianz bei 50:50, aber nicht zwingend!
- Zwei Sachverhalte sind oben nicht erkennbar, gelten aber im Regelfall:

- Var(ß) sinkt, wenn die Stichprobe vergrössert wird (Diskussion -> Davidson & MacKinnon, 2009, 101)
- Var(ß) kann steigen bei Multikollinearität (Diskussion -> Maddala & Lahiri, 2009, 282)

* R. Davidson & J. G. MacKinnon (2009), Econometric Theory and Methods, Oxford University Press // G. S. Maddala & K. Lahiri (2009), Introduction to Econometrics, Wiley

Multiple lineare Regression

Motivation

Bsp. Aus Verbeek (2017, 188)

Urne mit roten und gelben Kugeln

Gesucht: Anteil an roten Kugeln

Stichprobe von N Kugeln

Definiere ZV yi = 1 falls Kugel i rot ist, 0 sonst.

Wenn p die wahre (unbekannte) Wahrscheinlichkeit ist, rot zu ziehen, dann gilt für diese Bernoulli ZV P{ yi = 1 } = p.

Mit N1 roten und N-N1 gelben Kugeln in der Stichprobe mit Zurücklegen gilt für die Wahrscheinlichkeit genau dieser Ziehung:

$$P\{N_1 \text{ red balls, } N - N_1 \text{ yellow balls}\} = p^{N_1}(1-p)^{N-N_1}.$$

Multiple lineare Regression

Motivation

$$P\{N_1 \text{ red balls, } N - N_1 \text{ yellow balls}\} = p^{N_1}(1-p)^{N-N_1}.$$

Man kann diese Wahrscheinlichkeit als Funktion von p lesen

Man spricht dann von “likelihood function” (lf)

Um das gesuchte p zu schätzen, wird nun der Wert für p genommen, der die Wahrscheinlichkeit maximiert, genau diese Ziehung als Stichprobe zu erhalten

Konkret: Maximiere lf bzgl. P

Es ist bequemer die Produkte in Summen zu verwandeln. Logarithmieren leistet dies. Da log eine monoton steigende Funktion ist, ist es egal ob man log(ll) – also llf - oder lf via p maximiert. Die Lösung ist identisch.

$$\log L(p) = N_1 \log(p) + (N - N_1)\log(1-p).$$

Multiple lineare Regression

Motivation

Maximiere llf bzgl. P

Bspw. Mit Stipo = 4 rote Kugeln in der Stichprobe vom Umfang 100

$$\frac{d \log L(p)}{dp} = \frac{N_1}{p} - \frac{N - N_1}{1 - p} = 0,$$

$$\hat{p} = N_1 / N$$

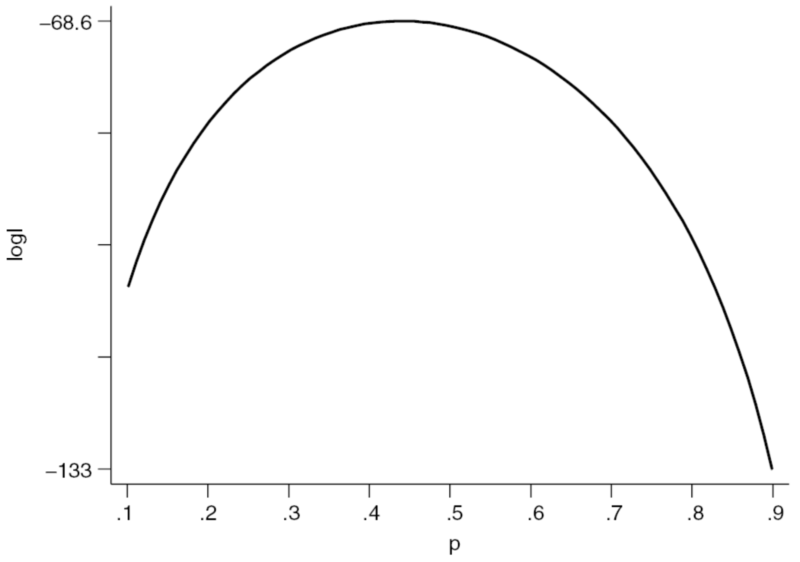


Figure 6.1 Sample loglikelihood function for $N = 100$ and $N_1 = 44$

Beschreiben Sie die Lösung in Ihren Worten

Der ML Schätzer ist …

Multiple lineare Regression

ML Schätzer für das lineare Regressionsmodell

Das Modell mit normalverteiltem Störterm (zusätzlich zu Gauss-Markov-Annahmen a0-a4):

$$y_i = \beta_1 + \beta_2 x_i + \varepsilon_i,$$

Die abhängige Variable ist nun stetig, weil der Störterm stetig ist. Wir betrachten die Dichte für Beobachtung yi, gegeben xi und die Modellparameter

$$f(y_i | x_i; \beta, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2}\frac{(y_i - \beta_1 - \beta_2 x_i)^2}{\sigma^2}\right\}$$

Da die Beobachtungen unabhängig sind, ist die gemeinsame Dichte das Produkt aus vielen Dichten über alle Beobachtungen i=1,…,N. Logarithmieren macht aus dem Produkt eine Summe und es ergibt sich …

Multiple lineare Regression

ML Schätzer für das lineare Regressionsmodell

Die folgende Log-likelihood Funktion (llf)

$$\log L(\beta, \sigma^2) = -\frac{N}{2}\log(2\pi\sigma^2) - \frac{1}{2}\sum_{i=1}^{N}\frac{(y_i - \beta_1 - \beta_2 x_i)^2}{\sigma^2}$$

Damit selbe Logik wie zuvor (aber nun drei Parameter zu schätzen)

Maximierung über die Parameter ß ist analytisch lösbar und Ergebnis wie „Kleinste Quadrate“ Methode – geg. einen sample y,x

Man sieht dies auch am Zähler.

Was steht dort ausgedrückt in Residuen?

Was gilt für max - f(x) und min f(x)?

Multiple lineare Regression

ML Schätzer für das lineare Regressionsmodell

Führen Sie in GML_02_01 capm.R Zeilen 169-187 aus

Numerisch maximiert wird *0.5*n*log(2*pi*sigma^2)-.5*(t(e)%*%e)/sigma^2*

Kommt Ihnen das bekannt vor?

$$\log L(\beta, \sigma^2) = -\frac{N}{2}\log(2\pi\sigma^2) - \frac{1}{2}\sum_{i=1}^{N}\frac{(y_i - \beta_1 - \beta_2 x_i)^2}{\sigma^2}$$

Interpretieren Sie das Ergebnis

```
> coef(reg)
(Intercept)           X
-0.07258514  1.16816828
> MLestimates[2]
$estimate
(Intercept)           X
-0.07259094  1.16816818 -2.83279913
```

Aber warum ist der Punktschätzer für Sigma bei OLS 2.837 und hier 2.833? ->FG vs n