

# Inferenz im multiplen Regressionsmodell

## Kapitel 4, Teil 2

Ökonometrie I  
Michael Hauser

# Inhalt

- ▶ ANOVA, analysis of variance
- ▶ korrigiertes  $R^2$ ,  $\bar{R}^2$
- ▶  $F$ -Test
- ▶  $F$ -Test bei linearen Restriktionen
- ▶ Erwartungstreue, Effizienz, Konsistenz
- ▶  $\chi^2$ -,  $F$ -Verteilung im Anhang

# **Varianzzerlegung, ANOVA, analysis of variance**

# Varianzzerlegung, ANOVA

$$y = \hat{y} + \hat{u}$$

$$TSS = ESS + RSS$$

Die Varianzzerlegungstabelle dient der Feststellung der FGe von TSS, ESS und RSS. (*df* degrees of freedom, FGe)

Quelle	SS	df's	SS/df
Gesamt	$TSS = S_{yy}$	$n - 1$	
Regression	$ESS = R^2 S_{yy}$	$k$	$R^2 S_{yy} / k = MS_1$
Residuen	$RSS = (1 - R^2) S_{yy}$	$n - k - 1$	$\frac{(1 - R^2) S_{yy}}{n - k - 1} = MS_2$

$MS \dots$  mittlere SS     $MS_2 = \hat{\sigma}^2 = \sum \hat{u}_i^2 / (n - k - 1)$

$TSS \dots$  total SS =  $\sum (y_i - \bar{y})^2$      $ESS \dots$  explained SS     $RSS \dots$  residual SS

# Varianzzerlegung, ANOVA

- ▶ Für die Berechnung der Varianz von  $y$  verliert man 1 FG, da das Mittel abgezogen wird.
- ▶ Die Regression besitzt  $k$  FGe, ie die Anzahl der eingeschlossen Variablen (ohne Konstante).
- ▶ Für die Berechnung der Varianz der Residuen verlieren wir  $(k + 1)$  FGe.

## F-Ratio

$MS_1$  und  $MS_2$  bezeichnen die *mittleren Quadratsummen*, vergleichbar den um FGe korrigierten Varianzen.

$$F = \frac{MS_1}{MS_2} = \frac{ESS/k}{RSS/(n-k-1)} = \frac{(TSS - RSS)/k}{RSS/(n-k-1)}$$

$F$  bezeichnet den Quotienten aus  $MS_1$  und  $MS_2$ , und wird *F - Ratio* genannt.

$$F = \frac{R^2 S_{yy}/k}{(1 - R^2) S_{yy}/(n - k - 1)}$$

bzw

$$F = \frac{[S_{yy} - (1 - R^2) S_{yy}]/k}{(1 - R^2) S_{yy}/(n - k - 1)}$$

# ***F*-Test**

# F-Test

Wir vergleichen 2 Modelle:

- ▶ Das **restringierte Modell**, [R]:

$$y = \beta_0 + u$$

$RSS_R$  ...  $RSS$  des restringierten Modells

Es gelten hier die Restriktionen:

$$\beta_1 = \dots = \beta_k = 0$$

Alle  $\beta$ -Koeffizienten ausgenommen der Interzept sind null.

- ▶ Das **unrestringierte Modell**, [UR]:

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + u$$

$RSS_{UR}$  ...  $RSS$  des unrestringierten Modells



## F-Test

Wir wollen wissen, welches der beiden Modelle das geeignetere ist. Dazu konstruieren wir einen Test. Die Hypothesen für den **F-Test** lauten

$$H_0 : \beta_1 = \dots = \beta_k = 0$$

$$H_1 : \text{mindestens ein } \beta_j \neq 0$$

Die Test-Statistik

$$F = \frac{(RSS_R - RSS_{UR})/k}{RSS_{UR}/(n - k - 1)} \sim F_{k, n-k-1}$$

ist unter Nullhypothese  $F$ -verteilt mit  $k$  und  $(n - k - 1)$  FGen.

Man schätzt beide Modelle  $R$  und  $UR$ , berechnet die entsprechenden  $RSS$ ,  $RSS_R$  und  $RSS_{UR}$ , und setzt ein.

Bem: Hier gilt als Spezialfall

$$F = \frac{ESS_{UR}/k}{RSS_{UR}/(n - k - 1)}$$

# F-Test

- ▶ Die Nullhypothese (restringiertes Modell) besagt, dass keine der unabhängigen Variablen einen Erklärungswert hat.
- ▶ Die Nullhypothese ist als *Ausschlussbedingung* für alle potenziell erklärenden Variablen formuliert.

Da gemeinsam auf den Ausschluss aller unabhängigen Variablen getestet wird, spricht man auch von einem Test auf die *Signifikanz der Regression*.

## F-Test und $R^2$

Modell:

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + u$$

Eine andere Formulierung der Teststatistik ist

$$F = \frac{R^2}{1 - R^2} \frac{n - k - 1}{k}$$

Der  $F$ -Test kann als *Test auf das multiple Bestimmtheitsmaß*,  $R^2$ , aufgefasst werden.

$$H_0 : R^2 = 0 \quad \text{und} \quad H_1 : R^2 > 0$$

Dazu berechnen wir nur das  $R^2$  des UR-Modells.

## Bsp: Okun's Gesetz

Wir wollen testen, ob die Variable  $\rho_t$  im Modell eingeschlossen werden soll, bzw ob  $R^2 \neq 0$  ist.

$$H_0 : \beta_1 = 0 \quad \text{bzw} \quad H_0 : R^2 = 0$$

$$H_1 : \beta_1 \neq 0 \quad \text{bzw} \quad H_1 : R^2 > 0$$

Wir schätzen das unrestringierte Modell (1977-2010) [UR]

$$\begin{array}{ll} \widehat{\Delta u}_t = 0.500 - 0.165 \rho_t & R^2 = 0.322 \\ (0.116) (0.042) & RSS = 5.575 \end{array}$$

und das restringierte Modell (1977-2010) [R]

$$\begin{array}{ll} \widehat{\Delta u}_t = 0.144 & R^2 = 0.000 \\ (0.086) & RSS = 8.224 \end{array}$$

## Bsp: Okun's Gesetz

Die  $F$ -Statistik ergibt sich als

$$F = \frac{(RSS_R - RSS_{UR})/k}{RSS_{UR}/(n - k - 1)} = \frac{(8.224 - 5.575)/1}{5.575/(34 - 1 - 1)} = 15.205$$

Der kritische Werte der  $F$ -Verteilung mit  $(1, 32)$  FGen und  $\alpha = 0.05$  ist 4.149.  
Damit wird die  $H_0$  abgelehnt. Der zugehörige  $p$ -Wert, 0.0005, ist kleiner als 0.05.

In EViews:

```
scalar p = 1 - @cfdist(15.205, 1, 32) ... p-Wert bzw  
scalar crit = @qfdist(1- $\alpha$ , 1, 32) ... kritischer Wert.
```

In R:

```
p <- 1 - pf(15.205, 1, 32) ... p-Wert, pf(...) =  $P(X \leq x)$ , bzw  
crit <- qf(1- $\alpha$ , 1, 32) ... kritischer Wert, qf(...) = Quantil.
```

## Okun's Gesetz: $F$ -Test

- ▶ Den gleichen  $F$ -Wert erhält man bei Verwendung von  $R^2$  aus dem unrestringierten Modell.
- ▶ Bei  $k = 1$  liefert der  $F$ -Test *und* der beidseitige  $t$ -Test zu  $\beta_1 = 0$  die selbe Entscheidung. Die  $p$ -Werte sind bei beiden Tests gleich.
- ▶ EViews und R geben in jedem Regressionsoutput den  $F$ -Wert und zugehörigen  $p$ -Wert zur Standard-Hypothese  $H_0 : \beta_1 = \dots = \beta_k = 0$  an. (Signifikanz der Regression)

## F-Test für eine allgemeine $H_0$

Der  $F$ -Test kann auch verwendet werden, wenn nur eine Auswahl der Variablen  $x_j$ ,  $j = 1, \dots, k$ , aus dem Modell ausgeschlossen werden soll.

Das restringierte Modell mit  $\ell < k$  sei

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_\ell x_\ell + u \quad [R]$$

Das unrestringierte Modell ist

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_\ell x_\ell + \beta_{\ell+1} x_{\ell+1} + \dots + \beta_k x_k + u \quad [UR]$$

Die Hypothese lautet

$$H_0 : \beta_{\ell+1} = \dots = \beta_k = 0 \quad \ell < k$$

$$H_1 : \text{mindestens ein } \beta_j \neq 0 \quad \ell < j \leq k$$

## F-Test für eine allgemeine $H_0$

Die  $F$ -Statistik ist

$$F = \frac{(RSS_R - RSS_{UR})/(k - \ell)}{RSS_{UR}/(n - k - 1)}$$

Sie ist  $F$ -verteilt mit  $(k - \ell)$  und  $(n - k - 1)$  FGen.

*Unter der  $H_0$  liegen  $(k - \ell)$  Null-Restriktionen vor.*



## ***F*-Test für lineare Restriktionen auf Parameter**

## F-Test für lineare Restriktionen auf Parameter

**Bsp:** Restriktionen für die Parameter treten zB bei der Cobb-Douglas Produktionsfunktion auf:

$$Q = AK^\alpha L^\beta v$$

Logarithmiert lautet sie

$$\log(Q) = \gamma + \alpha \log(K) + \beta \log(L) + u \quad [UR]$$

$\gamma \dots \log(A)$ ,  $u \dots \log(v)$ .

Unter der Hypothese konstanter Skalenerträge gilt

$$\alpha + \beta = 1$$

Das ist eine **lineare Restriktion** auf die Parameter.

## F-Test für lineare Restriktionen auf Parameter

Wir substituieren  $\alpha$  durch

$$\alpha = 1 - \beta$$

Das restringierte Modell lautet daher

$$\log(Q) = \gamma + 1 \log(K) + \beta[\log(L) - \log(K)] + u$$

Wir bringen  $\log(K)$  auf die linke Seite und erhalten das *Schätzmodell*

$$[\log(Q) - \log(K)] = \gamma + \beta[\log(L) - \log(K)] + u \quad [R]$$

## F-Test für lineare Restriktionen auf Parameter

Die Hypothesen lauten

$$H_0 : \alpha + \beta = 1 \quad \text{und} \quad H_1 : \alpha + \beta \neq 1$$

$$F = \frac{(RSS_R - RSS_{UR})/1}{RSS_{UR}/(n - 2 - 1)}$$

Im unrestringierten Modell ist  $k = 2$ . Die Differenz der Anzahl der Parameter des unrestringierten und restringierten Modell ist 1. Daher liegt eine  $F$ -Vtlg mit  $(1, n - 2 - 1)$  FGen vor.

Allgemein liegt bei  $m$  Restriktionen eine  $F$ -Vtlg mit  $(m, n - k - 1)$  FGen vor.

In EViews ist es möglich für das unrestringiert spezifizierte Modell allgemeine lineare Restriktionen als Formel einzugeben.

**Adjusted  $R^2$ ,  $\bar{R}^2$**

## Adjusted $R^2$ , $\bar{R}^2$

Das multiple Bestimmtheitsmaß,  $R^2$ , wird aus dem Quotienten von  $ESS / TSS$ , bestimmt. Verwendet man stattdessen die um FGe korrigierte SS, erhält man das **korrigierte** oder **adjusted**  $R^2$ ,  $\bar{R}^2$ .

Es ist definiert analog zu  $1 - R^2 = RSS / TSS$ :

$$1 - \bar{R}^2 = \frac{RSS / (n - k - 1)}{TSS / (n - 1)} = (1 - R^2) \frac{n - 1}{n - k - 1}$$

**Bem:** Es gilt  $RSS = (1 - R^2)TSS$

## Adjusted $R^2$ , $\bar{R}^2$

Seine Eigenschaften sind:

- ▶ Für kleine  $R^2$  kann  $\bar{R}^2$  negativ werden.

**Bsp:** Für  $R^2 < 0.1$ ,  $k = 2$  und  $n = 21$

$$\bar{R}^2 = 1 - \frac{n-1}{n-k-1}(1-R^2) = 1 - \frac{10}{9}(1-R^2) < 0$$

- ▶  $R^2$  wird mit dem Hinzufügen jeder Variablen etwas größer, unabhängig davon ob die Variable sinnvoll ist, oder nicht.
- ▶  $\bar{R}^2$  erfordert für jede zusätzliche Variable eine *Mindestreduktion in der RSS* - etwa einen Anteil von  $(1/n) \cdot (1 - \bar{R}^2)$  - damit es steigt.

**Statistische Konzepte:  
Erwartungstreue, Effizienz, Konsistenz**



# Theoretische Konzepte

Erwartungstreue, Effizienz und Konsistenz sind theoretische Eigenschaften von Schätzern. Dabei wird das Verhalten der Schätzer beschrieben, wenn das Modell für alle mögliche Stichproben der Länge  $n$  wiederholt geschätzt wird.

Für die Konsistenz treffen wir zusätzlich asymptotische Überlegungen, dh wir lassen  $n$  gegen unendlich gehen.

# Erwartungstreue

Ein Schätzer  $\hat{\theta}_n$  heißt **unverzerrt** für  $\theta$ , wenn  $E(\hat{\theta}_n) = \theta$ .

## **Bsp:** Stichprobenmittel

Wir ziehen aus einer Grundgesamtheit mit Erwartungswert  $\mu$  eine Stichprobe im Umfang  $n$ , und berechnen das Mittel,  $\bar{x}_n = \sum x_i/n$ . Für jede gezogene Stichprobe vom Umfang  $n$ , erhalten wir einen anderen Wert für  $\bar{x}_n$ . Berechnet man den Erwartungswert über alle  $\bar{x}_n$ , so erhält man  $\mu$ .

Das arithmetische Mittel ist daher ein erwartungstreuer (unverzerrter) Schätzer für den Erwartungswert  $\mu$  in der Grundgesamtheit.

# Erwartungstreue

Der **Bias** eines Schätzers ist

$$bias = E(\hat{\theta}_n) - \theta$$

Ist der Bias null heißt der Schätzer **unverzerrt**, bzw **erwartungstreu**, andernfalls **verzerrt**.

## ZGWS: Zentraler Grenzwertsatz

Angenommen sie ziehen  $n$  mal zufällig aus einer Grundgesamtheit,  $X$ , mit  $E(X) = \mu$ ,  $V(X) = \sigma^2$ . Einzelne Ziehungen werden mit  $X_i$  bezeichnet. Sie sind identisch verteilt wie  $X$ .

Dann sind Erwartungswert und Varianz des Stichprobenmittels  $\bar{X}_n = \frac{1}{n} \sum_i X_i$

$$E(\bar{X}_n) = \mu \quad \text{und} \quad V(\bar{X}_n) = \sigma^2/n$$

Weiters ist das Stichprobenmittel  $\bar{X}_n$  mit  $n \rightarrow \infty$

$$\bar{X}_n \stackrel{a}{\sim} N(\mu, \sigma^2/n)$$

**asymptotisch** normalverteilt.

Das Stichprobenmittel ist (für jedes  $n$ ) ein unverzerrter/erwartungstreuer Schätzer für  $\mu$ , dessen Varianz mit  $n \rightarrow \infty$  gegen null geht.

## MSE, mean square error

Wie wollen, dass unsere Schätzer eine kleine Varianz haben. Für das arithmetische Mittel ist sie  $(\sigma^2/n)$ .

Da nicht jeder Schätzer  $\hat{\theta}$  für  $\theta$  unverzerrt ist, definieren wir den **MSE, mean square error**, als

$$MSE(\hat{\theta}_n) = E[(\hat{\theta}_n - \theta)^2] = (\textit{bias})^2 + \textit{variance}$$

Dadurch erhält man einen Trade-off zwischen Bias und Varianz eines Schätzers. Dh ein größerer Bias könnte durch eine kleinere Varianz kompensiert werden, und umgekehrt.

Bem:  $MSE(\hat{\theta}_n) = E[(\hat{\theta}_n - \theta)^2] = E\{[\hat{\theta}_n - E(\hat{\theta}_n)]^2 + [E(\hat{\theta}_n) - \theta]^2\} = \textit{variance} + \textit{bias}^2$

# Effizienz

Effizienz eines Schätzers bezieht sich auf seine Varianz im Vergleich zu anderen Schätzern.

## **Bsp:**

Vergleichen sie die Fehlervarianz, wenn sie mit der Hand wiederholt eine Gerade in Punktwolken einzeichnen, mit der Fehlervarianz der OLS Lösung. OLS wird die kleinere Varianz liefern.

Der Schätzer, der die kleinste Varianz liefert, heißt **effizient**.

# Konsistenz, Bsp

**Bsp:** Für autoregressive Modelle, zB

$$y_t = \beta_0 + \beta_1 y_{t-1} + u_t, \quad |\beta_1| < 1$$

gibt es keinen Schätzer  $b_1$ , der in kleinen Stichproben unverzerrt wäre. Jedoch mit  $n \rightarrow \infty$  konvergiert der Erwartungswert des OLS Schätzers gegen  $\beta_1$ .

Das Konzept der Konsistenz fordert eine bestimmte Art der Konvergenz von  $b_1(n)$  gegen  $\beta_1$  mit  $n \rightarrow \infty$ .

$$b_1(n) \rightarrow \beta_1$$

# Konsistenz

Sei  $\hat{\theta}_n$  ein Schätzer basierend auf einer Stichprobe vom Umfang  $n$ . Dann heißt  $\hat{\theta}_n$  **konsistent**, wenn  $\forall \epsilon > 0$

$$\lim_{n \rightarrow \infty} \mathbf{P}(|\hat{\theta}_n - \theta| < \epsilon) = 1$$

Dh die Verteilung von  $\hat{\theta}_n$  konzentriert sich mit steigendem  $n$  immer mehr um die bzw an der Stelle  $\theta$ .

Dies entspricht einer *Konvergenz in der Wahrscheinlichkeit* und schreibt

$$\text{plim } \hat{\theta}_n = \theta$$

plim heißt auch *probability limit*.



# Konsistenz

*Graphik:*

*Verschiedene Normalverteilungen zentriert bei  $\theta$ ,  $N(\theta, \sigma^2/n)$ ,  $n = 10, 100, 1000$ .*

# Übungen und Referenzen

# Referenzen

Hackl 4.4, 5

Wooldridge 4

- 1 Wählen sie das Datenfile `dat_s_01.wf1` bzw `dat_s_01_22_R.txt`.  
Berechnen sie die Regression `Quick` → Estimate equation  
`pcr c pyr (wln(-1)/pcd(-1)) pcr(-1)`
- (a) Schreiben sie das Ergebnis der Schätzung an.
  - (b) Testen sie alle Parameter einzeln auf null.
  - (c) Welche Variable würden sie aus der Regression ausschließen?
  - (d) Testen sie alle Parameter (ohne  $\beta_0$ ) gemeinsam auf null. Zu welchem Schluß kommen sie?

# Übungen

- 2 Wählen sie das Datenfile `hprice1.wf1` bzw `hprice1_R.txt`. Darin befinden sich uA die Variablen `price` ... Verkaufspreis eines Einfamilienhauses, `assess` ... Schätzwert, `lotsize` ... Grundstücksfläche, `sqrft` ...  $m^2$ , `bdrms` ... Anzahl der Schlafzimmer

Zur Diskussion steht das Modell

$$\log(\text{price}) = \beta_0 + \beta_1 \log(\text{assess}) + \beta_2 \log(\text{lotsize}) + \beta_3 \log(\text{sqrft}) + \beta_4 \text{bdrms} + u$$

Wenn Preis- und Schätzwertbildung konsistent sind, sollte

$$\beta_1 = 1 \quad \text{und} \quad \beta_2 = \beta_3 = \beta_4 = 0$$

gelten. Alle Infos sollten in `assess` enthalten sein.

(a) Hier liegt eine einfache lineare Restriktion vor. Testen sie die Hypothese mittels  $F$ -Test, indem sie das restringierte Modell mit dem unrestringierten vergleichen. [Woo (4.47)]

- 2 (b) Was ist der Unterschied zwischen dem einfachen  $t$ -Test auf 1 für den Koeffizienten von  $\log(\text{assess})$  und dem Test in (a).  
Schreiben sie dazu die Modelle unter den beiden Nullhypothesen an.

- 3 Wählen sie das Datenfile `twoyear.wfl` bzw `twoyear_R.txt`. Darin befinden sich uA die Variablen  
`lwage` ...  $\log(\textit{wage})$ , `jc` ... Jahre in einem 2-year College,  
`univ` ... Jahre in einem 4-year College,  
`exper` ... Berufserfahrung in Monaten

Sie schätzen

$$\log(\textit{wage}) = \beta_0 + \beta_1 \textit{jc} + \beta_2 \textit{univ} + \beta_3 \textit{exper} + u$$

Ist der Grenzertrag eines Ausbildungsjahres eines 2-year College gleich dem Grenzertrag eines 4-year College?

Testen sie  $\beta_1 = \beta_2$ .

Formulieren sie das unrestringierte und das restringierte Modell, und wenden sie den  $F$ -Test an.

Testen von Modellen mit linearen Restriktionen auf den Parametern:

Schätzen des unrestringierten Modells,

im Estimation output → View Menü →

Coefficient Diagnostics →

Wald Test - Coefficient Restrictions

Restriktion nach Vorlage eingeben.

Unser Test ist die  $F$ -statistic.

Die default Modell-Notation in EViews ist

$$y = c(1) + c(2)*x1 + c(3)*x2$$