

# SE Geographie und Ökonomie

Einheit 6:  
Statistische Unsicherheit  
und Testen von Hypothesen  
im Linearen Regressionsmodell

Dieter Pennerstorfer  
dieter.pennerstorfer@jku.at

Institut für Volkswirtschaftslehre

WS 2021/22

# Motivation

- Bisher haben wir ein **einfaches lineare Regressionsmodell** betrachtet:

$$y_i = \alpha + \beta x_i + u_i$$

für  $i = 1, 2, \dots, n$  Beobachtungen

- Wenn wir die Methode der Kleinsten Quadrate anwenden, erhalten wir einen **Punktschätzer**  $\hat{\beta}$  (und  $\hat{\alpha}$ , den wir hier vernachlässigen).
- Der Punktschätzer für  $\beta$  gibt (für die Stichprobe) an, wie stark  $x_i$  im Durchschnitt  $y_i$  beeinflusst.
  - ▶ Eine Veränderung von  $x_i$  um eine Einheit führt zu einer Änderung von  $y_i$  um  $\hat{\beta}$  Einheiten.
- Wie groß ist die (statistische) Unsicherheit, wenn wir **Aussagen über die Grundgesamtheit** treffen möchten?
  - ▶ Der geschätzte Parameter  $\hat{\beta}$  ist immer mit Unsicherheit behaftet, da wir nie die gesamten Daten beobachten, sondern immer nur eine Stichprobe.
  - ▶ Die damit verbundene Unsicherheit sollte bei dessen Interpretation berücksichtigt werden.
- Wer bedienen uns Methoden der **schließenden Statistik** (siehe Einheit 3), um Rückschlüsse auf den Einfluss von  $x$  auf  $y$  in der **Grundgesamtheit** zu ziehen.

## Lernziele Einheit 6

- Sie verstehen, dass der Punktschätzer  $\hat{\beta}$  nur mit **Unsicherheit** geschätzt werden kann.
- Sie verstehen, wie Sie diese Unsicherheit in einem linearen Regressionsmodell bestimmen können.
- Sie können die Unsicherheit quantifizieren und interpretieren, indem Sie neben dem Punktschätzer auch einen **Intervallschätzer** für den Parameter  $\beta$  bestimmen können.
- Sie können **einfache Hypothesen** in Bezug auf den geschätzten Parameter  $\hat{\beta}$  aufstellen.
- Sie können diese **Hypothesen testen** und die Ergebnisse interpretieren.
- Abhängig von Ihrem Erkenntnisinteresse können Sie **zweiseitige** oder **einseitige Hypothesentests** durchführen.
- Sie können **in EXCEL** einen **Intervallschätzer** bestimmen und einen **Hypothesentest** durchführen.

# Anforderung an den geschätzten Parameter $\hat{\beta}$

Der geschätzte Parameter  $\hat{\beta}$  ist eine normalverteilte Zufallszahl und sollte folgende Anforderungen erfüllen:

- **Erwartungstreue:**

Der Erwartungswert des geschätzte Parameter  $\hat{\beta}$  entspricht dem gesuchten (allerdings unbekanntem und unbeobachtbaren) Parameter  $\beta$ :

$$E(\hat{\beta}) = \beta$$

- ▶ Unter unseren Annahmen A1-A3 (und B1) ist  $\hat{\beta}$  ein **erwartungstreuer** (auch: **unverzerrter**) **Schätzer** und bildet den **kausalen Wirkungszusammenhang** ab.

- **Konsistenz:**

Mit zunehmendem Stichprobenumfang wird die Varianz des Schätzers kleiner:

$$\lim_{n \rightarrow \infty} \text{var}(\hat{\beta}) \rightarrow 0$$

- **Effizienz:**

Der (unbekannte) Parameter  $\beta$  wird **sehr präzise** geschätzt, d.h.  $\text{var}(\hat{\beta})$  ist möglichst klein. Ein Schätzer gilt dann als **effizient**, wenn er innerhalb der unverzerrten Schätzer die **kleinste Streuung** aufweist.

# Punktschätzer & Unsicherheiten

- Der **Kleinstquadratschätzer erfüllt die Anforderungen** in Bezug auf Erwartungstreue, Konsistenz und Effizienz. Wir wollen hier die Unsicherheit quantifizieren, wenn wir Rückschlüsse auf die Grundgesamtheit ziehen.
- Wenn die A- und B-Annahmen erfüllt sind, kann die **Verteilung des geschätzten Parameters**  $\hat{\beta}$  bestimmt werden:

$$\hat{\beta} \sim N\left(\beta, \frac{\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2}\right)$$

- Wir wissen von letzter Einheit, dass  $u_i \sim N(0, \sigma^2)$ . Daher gibt:

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^n (u_i - \bar{u})^2 = \frac{1}{N} \sum_{i=1}^n u_i^2$$

- Wir können folgende Umformung vornehmen:

$$\sum_{i=1}^n (x_i - \bar{x})^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} (n-1) = \hat{s}_x^2 (n-1)$$

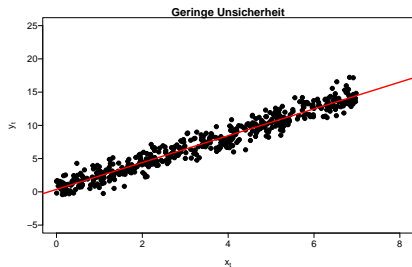
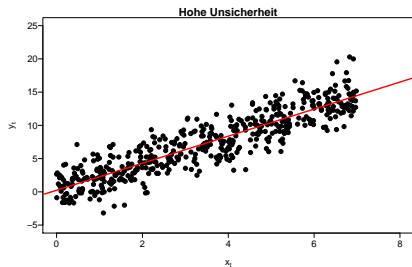
# Bestimmungsgründe der Unsicherheit

$$\hat{\beta} \sim N\left(\beta, \frac{\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2}\right) \text{ bzw. } \hat{\beta} \sim N\left(\beta, \frac{\sigma^2}{\hat{s}_x^2(n-1)}\right)$$

Die **Genauigkeit**, mit der der Parameter  $\beta$  geschätzt werden kann, hängt von 3 Faktoren ab.  $\beta$  kann umso genauer geschätzt werden

- 1 je kleiner die **Varianz der Störgrößen**  $\sigma^2$  ist.  
Das bedeutet, dass das Regressionsmodell den Zusammenhang zwischen den Variablen  $x$  und  $y$  gut erklärt, und nur ein kleiner (unsystematischer) Teil unerklärt bleibt. Die Störgrößen  $u_i$  sind daher für die meisten Beobachtungen (im Betrag) sehr klein.
- 2 je größer der **Stichprobenumfang**  $n$  ist.
- 3 je größer die **Streuung (Varianz) der erklärenden Variable**  $x$  ist.

# Punktschätzer & Unsicherheiten: Grafische Darstellung



- In beiden Datensätzen sind die Punktschätzer (die Lage der Regressionsgeraden) gleich und die Annahmen B1-B4 erfüllt.
- Im Datensatz der linken Grafik ist die Abweichung der Beobachtungen von der Regressionsgerade (also  $|u_i|$ ) im Durchschnitt größer → **größere Unsicherheit** in der Schätzung des Modells.
- Wichtigkeit von **Stichprobenumfang** und **Streuung der erklärenden Variable** ist ebenfalls ersichtlich.

## Intervallschätzer für $\beta$ wenn $\sigma^2$ bekannt ist

- Die Standardabweichung der Verteilung des geschätzten Parameters wird wieder **Standardfehler** genannt und mit **se** (*standard error*) bezeichnet. Der Standardfehler ist die Wurzel der Varianz der Verteilung des geschätzten Parameters  $\hat{\beta}$ :

$$se(\hat{\beta}) = \sqrt{\frac{\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}$$

- Da  $\hat{\beta}$  normalverteilt ist, können wir mittels Standardfehler  $se(\hat{\beta})$  die **Unsicherheit quantifizieren**. Der (unbeobachtete) Parameter  $\beta$  liegt mit einer Wahrscheinlichkeit  $(1 - \alpha)$  **in folgendem Bereich**:

$$\hat{\beta} - u_{1-\frac{\alpha}{2}} \cdot se \leq \beta \leq \hat{\beta} + u_{1-\frac{\alpha}{2}} \cdot se$$

- Der Bereich kann auch als **Intervall** angeschrieben werden:

$$\left[ \hat{\beta} - u_{1-\frac{\alpha}{2}} \cdot se; \hat{\beta} + u_{1-\frac{\alpha}{2}} \cdot se \right]$$

- Für  $\alpha = 0.05$  gilt (da  $u_{1-\frac{\alpha}{2}} = u_{0.975} = 1.96$ ):

$$\left[ \hat{\beta} - 1.96 \cdot se; \hat{\beta} + 1.96 \cdot se \right]$$



## Intervallschätzer für $\beta$ wenn $\sigma^2$ unbekannt ist

- $\sigma^2$  ist üblicherweise **unbeobachtbar**, kann aber unter den Annahmen A1-A3 und B1-B3 **geschätzt werden**.
- Da unser Modell die Wirklichkeit widerspiegelt (unter A1-A3), können wir die **Residuen (die geschätzten Störterme)** aus unserem Modell verwenden:

$$\hat{u}_i = y - \hat{\alpha} - \hat{\beta}x_i$$

- Unter den Annahmen B1-B3 sind die **aggregierten und durch  $(n - 2)$  geteilten Residuenquadrate** ein **konsistenter Schätzer** für die Varianz der Störgrößen  $\sigma^2$ :

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n (\hat{u}_i - \bar{u})^2}{n - 2} = \frac{\sum_{i=1}^n \hat{u}_i^2}{n - 2}$$

- Die  $(n - 2)$  im Nenner berücksichtigen, dass wir in unserem Modell zwei geschätzte Parameter haben ( $\hat{\alpha}$  und  $\hat{\beta}$ ) und deshalb nur  $(n - 2)$  **Freiheitsgrade** übrig bleiben.
- In einem Modell mit  $p$  Parameter müssten wir als folglich durch  $(n - p)$  teilen.

## Intervallschätzer für $\beta$ wenn $\sigma^2$ unbekannt ist

- Da die Varianz der Residuen der Stichprobe ( $\hat{\sigma}^2$ ) ein erwartungstreuer Schätzer für die (unbekannte) Varianz der Störgrößen der Grundgesamtheit ( $\sigma^2$ ) ist, **kann  $\sigma^2$  durch  $\hat{\sigma}^2$  ersetzt werden**. Daher wird der unbeobachtete Standardfehler  $se$  des geschätzten Parameters  $\hat{\beta}$  durch den **geschätzten Standardfehler  $\hat{se}$**  ersetzt:

$$\hat{se}(\hat{\beta}) = \sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}$$

- Da die Varianz nicht bekannt ist, sondern geschätzt werden muss, **erhöht sich die Unsicherheit** der Aussage über den geschätzten Parameter  $\hat{\beta}$ . Um diese Unsicherheit zu kompensieren, muss das **Konfidenzintervall etwas breiter** angelegt werden. Dies wird erreicht, indem man nicht die Perzentile der Standardnormalverteilung  $u_{1-\frac{\alpha}{2}}$  verwendet, sondern die Perzentile  $t_{n-2; 1-\frac{\alpha}{2}}$  der **t-Verteilung**.
- Der unbeobachtete Parameter  $\beta$  liegt daher mit einer **Wahrscheinlichkeit von  $(1 - \alpha)$**  in folgendem **Intervall**:

$$\left[ \hat{\beta} - t_{n-2; 1-\frac{\alpha}{2}} \cdot \hat{se}; \hat{\beta} + t_{n-2; 1-\frac{\alpha}{2}} \cdot \hat{se} \right]$$

- Beachte: Argumentation und Vorgehensweise sind analog zu Einheit 3!

# Intervallschätzer: Beispiel

- Wie hilft uns der Intervallschätzer, die **statistische Unsicherheit des geschätzten Parameters**  $\hat{\beta}$  zu bestimmen?
- In unserem Anwesenheit-Studienerfolg Modell erhalten wir folgende Werte für den geschätzten Parameter  $\hat{\beta}$  und den geschätzten Standardfehler des geschätzten Parameters,  $\hat{se}(\hat{\beta})$ :

$$\hat{\beta} = 4.88$$

$$\hat{se}(\hat{\beta}) = 0.88$$

- Basierend auf diesen Schätzergebnissen liegt der **(unbeobachtete) Parameter  $\beta$** 
  - ▶ in **90 %** aller Fälle (d.h.  $\alpha = 0.10$ ) im Intervall  $[4.88 - 1.86 \cdot 0.88; 4.88 + 1.86 \cdot 0.88] = [3.25; 6.52]$
  - ▶ in **95 %** aller Fälle (d.h.  $\alpha = 0.05$ ) im Intervall  $[4.88 - 2.31 \cdot 0.88; 4.88 + 2.31 \cdot 0.88] = [2.85; 6.91]$
  - ▶ in **99 %** aller Fälle (d.h.  $\alpha = 0.01$ ) im Intervall  $[4.88 - 3.36 \cdot 0.88; 4.88 + 3.36 \cdot 0.88] = [1.93; 7.84]$
  - ▶ **wenn** die Annahmen A1-A3 und B1-B4 erfüllt sind!!

# Intervallschätzer: Anmerkungen

- Die Größe  $\alpha$  legt die **Irrtumswahrscheinlichkeit** fest und wird auch als **Signifikanzniveau** bezeichnet.
- In der Praxis konzentrieren wir uns oft auf ein Signifikanzniveau von 5 %, d.h. in 95 % aller Fälle liegt der richtige Wert  $\beta$  innerhalb des entsprechenden Konfidenzintervalls.
- In **EXCEL** ist unter **Daten** → **Analyse** → **Datenanalyse** → **Regression** ein “Konfidenzniveau” von 95 % (also eine Irrtumswahrscheinlichkeit von  $\alpha = 0.05$ ) voreingestellt (kann aber verändert werden).
- Der Satz “In X % aller Fälle...” ist darauf zurückzuführen, dass wir bei der (theoretischen) Konstruktion des Konfidenzintervalls von unendlich vielen Zufallsstichproben ausgehen.
- Oft hört/liest man auch “Das Konfidenzintervall umschließt den wahren Wert mit Wahrscheinlichkeit von X %”. Diese Aussage erleichtert das Verständnis für Konfidenzintervalle, ist aber eigentlich nicht korrekt.
- Wenn die Anzahl der **Beobachtungen relative groß** ist, können die Perzentile der **t-Verteilung durch** die Perzentile einer **Standardnormalverteilung ersetzt** werden.

# Intervallschätzer: Zusammenfassung

- Die **Schätzungen der Modellparameter** sind mit **Unsicherheiten** verknüpft.
- Wir müssen diese **Unsicherheiten quantifizieren**, um relevante Aussagen treffen zu können.
- Eine Möglichkeit ist es, einen sogenannten **Intervallschätzer** zu konstruieren.
- Der Intervallschätzer gibt an, **in wie vielen Fällen der tatsächliche (aber unbeobachtete) Parameter in diesem Intervall** liegt.

# Testen von Hypothesen: Motivation

- Konfidenzintervalle sind sehr nützlich, aber manchmal möchten wir gewisse **Hypothesen** direkter **testen**.
- Sie als Studierende (aber auch mich als Lehrenden) könnten in unserem Anwesenheits-Klausur-Modell interessieren, ob
  - ▶ der geschätzter Effekt unterschiedlich von 0 ist ( $\beta \neq 0$ ).
  - ▶ höhere Anwesenheit zu einer Punktsteigerung führt ( $\beta > 0$ ).
  - ▶ die Anwesenheit bei einer zusätzlichen Einheit die Klausurpunkte um mehr als 3 Punkte steigert ( $\beta > 3$ ).
- Wie diese Fragen zeigen, kann man an **einseitigen oder zweiseitigen Hypothesentests** interessiert sein.
- Wir können solche Hypothesen mit unseren bisher erlernten Mitteln einfach formulieren und testen!

# Testen von Hypothesen: Wiederholung

Um Aussagen über die Grundgesamtheit zu treffen, wenn man nur eine Stichprobe kennt, gibt es neben der Angabe eines Konfidenzintervalls die Möglichkeit, statistische Tests durchzuführen. **Ein statistischer Test ist eine Regel zur Entscheidung unter Unsicherheit**, die sich daraus ergibt, weil man keine Kenntnisse über die Grundgesamtheit hat. Die Entscheidung ist zwischen zwei Behauptungen zu treffen, die als Hypothesen bezeichnet werden.

- **Nullhypothese ( $H_0$ ):** Behauptung über die **Grundgesamtheit**.
  - ▶ Hier: Behauptung über den Parameter  $\beta$  der Grundgesamtheit.
- **Alternativhypothese ( $H_1$ ):** Behauptung über die **Grundgesamtheit**
- Null- und Alternativhypothese **schließen einander aus** und **ergänzen sich** (d.h. eine der beiden Hypothesen wird angenommen).
- Beim Testen gibt es **zwei Möglichkeiten**:
  - i) Die Nullhypothese kann nicht abgelehnt werden. Man entscheidet sich daher für die Nullhypothese. (Achtung: Das bedeutet nicht, dass die Nullhypothese zutrifft!)
  - ii) Die Nullhypothese wird zugunsten der Alternativhypothese abgelehnt.

# Zweiseitiger Hypothesentest

- Vorbereitung: Allgemeine und spezielle **Voraussetzungen** des Tests **überprüfen** und **Irrtumswahrscheinlichkeit** ( $\alpha$ -Fehler) **festlegen** (z.B.  $\alpha = 0.01$  oder  $\alpha = 0.05$ ).
- Formulierung der Hypothesen:

Nullhypothese  $H_0 : \beta = q$

Alternativhypothese  $H_1 : \beta \neq q$

- Mit dem Hypothesentest wird überprüft, ob es **wahrscheinlich** ist, dass der **(unbeobachtete) Parameter**  $\beta$  in der Nähe bzw. nicht in der Nähe von  $q$  liegt. Oft wird  $q = 0$  gewählt und überprüft, ob der geschätzte Parameter  $\hat{\beta}$  signifikant von 0 verschieden ist (und sich  $\beta$  wahrscheinlich von 0 unterscheidet).
- $H_0$  wird abgelehnt, wenn  $\hat{\beta}$  signifikant **größer oder** signifikant **kleiner** als  $q$  ist  $\rightarrow$  daher die Bezeichnung **“zweiseitiger” Hypothesentest**.
- Als Nullhypothese wird eine möglichst unvoreingenommene Behauptung verwendet. Die **nachzuweisende Behauptung** wird als **Alternativhypothese** formuliert, weil die Nullhypothese abzulehnen ( $H_1$  zu akzeptieren) eine statistisch größere Aussagekraft hat.



# Zusammenhang Intervallschätzer und Hypothesentests

## 1. Hypothesen:

▶  $H_0 : \beta = q; H_1 : \beta \neq q$

## 2. Entscheidungsregel: mittels geschätztem Konfidenzintervall:

▶  $H_0$  wird nicht verworfen, wenn

$$q \in \left[ \hat{\beta} - t_{n-2; 1-\frac{\alpha}{2}} \cdot \widehat{se}; \hat{\beta} + t_{n-2; 1-\frac{\alpha}{2}} \cdot \widehat{se} \right]$$

▶  $H_0$  zugunsten von  $H_1$  verworfen, wenn  $q \notin [; \cdot]$

## Entscheidungsregel mittels standardisierter Teststatistik:

$H_0$  wird nicht verworfen, wenn:

$$\hat{\beta} - t_{n-2; 1-\frac{\alpha}{2}} \cdot \widehat{se} \leq q \leq \hat{\beta} + t_{n-2; 1-\frac{\alpha}{2}} \cdot \widehat{se}$$

$$-t_{n-2; 1-\frac{\alpha}{2}} \cdot \widehat{se} \leq q - \hat{\beta} \leq +t_{n-2; 1-\frac{\alpha}{2}} \cdot \widehat{se}$$

$$-t_{n-2; 1-\frac{\alpha}{2}} \leq \frac{q - \hat{\beta}}{\widehat{se}} \leq +t_{n-2; 1-\frac{\alpha}{2}}$$

$$\text{t-Wert} = \left| \frac{\hat{\beta} - q}{\widehat{se}(\hat{\beta})} \right| \leq t_{n-2; 1-\frac{\alpha}{2}}$$

# Standardisierte Teststatistik

## 1. Hypothesen:

- ▶  $H_0 : \beta = q; H_1 : \beta \neq q$

## 2. Entscheidungsregel: mittels standardisierten Teststatistik

- ▶  $H_0$  wird nicht verworfen, wenn t-Wert =  $\left| \frac{\hat{\beta} - q}{\widehat{se}(\hat{\beta})} \right| \leq t_{n-2; 1 - \frac{\alpha}{2}}$
- ▶  $H_0$  zugunsten von  $H_1$  verworfen, wenn t-Wert =  $\left| \frac{\hat{\beta} - q}{\widehat{se}(\hat{\beta})} \right| > t_{n-2; 1 - \frac{\alpha}{2}}$
- ▶ Vorgehensweise:

Berechnung einer **standardisierten Teststatistik (t-Wert)**:

$$\text{t-Wert} = \left| \frac{\hat{\beta} - q}{\widehat{se}(\hat{\beta})} \right|$$

Ermittlung des **kritischen Werts**  $c$  der Verteilung:  $c = t_{n-2; 1 - \frac{\alpha}{2}}$

$H_0$  **wird nicht verworfen**, wenn t-Wert  $\leq c$

$H_0$  **wird verworfen**, wenn t-Wert  $> c$

Hinweis: Oft wird der t-Wert (in Lehrbüchern oder etwa auch in EXCEL) als  $\frac{\hat{\beta} - q}{\widehat{se}(\hat{\beta})}$  berechnet und kann demnach auch negative Werte annehmen. Wichtig ist, dass Sie in diesem Fall den Betrag des ausgewiesenen t-Werts mit dem kritischen Wert  $c$  vergleichen!

# Zweiseitiger Hypothesentest: Anleitung

Um einen **zweiseitigen Hypothesentest** durchzuführen, müssen Sie folgende Schritte durchführen:

- Aufstellen von  $H_0$  ( $H_0 : \beta = q$ ) und  $H_1$  ( $H_1 : \beta \neq q$ )
- Festlegung des Signifikanzniveaus  $\alpha$
- Schätzung von  $\widehat{se}(\widehat{\beta})$
- Berechnung des t-Wertes:  $t = \left| \frac{\widehat{\beta} - q}{sd(\widehat{\beta})} \right|$
- Ermittlung des kritischen Wertes  $c$ :  $c = t_{n-2; 1 - \frac{\alpha}{2}}$
- Vergleichen von  $c$  und  $t$ , **falls**  $t > c$  **wird**  $H_0$  **verworfen**.

## Zweiseitiger Test: Beispiel

- Nehmen wir an, wir wurden beauftragt herauszufinden, ob es einen **Wirkungszusammenhang zwischen Anwesenheit und Prüfungserfolg** gibt. Wir wollen wissen, ob  $\beta \neq 0$ .
- Die Aussage ( $\beta \neq 0$ ) wird daher als Alternativhypothese  $H_1$  formuliert. Daher haben wir folgende Hypothesen:  $H_0 : \beta = 0$  und  $H_1 : \beta \neq 0$ .
- Mit der Kleinstquadratschätzung haben wir für den Parameter  $\beta$  und den Standardfehler  $se(\hat{\beta})$  folgende Werte geschätzt:

$$\hat{\beta} = 4.88$$

$$\hat{se}(\hat{\beta}) = 0.88$$

- Hinweis: Damit unser **Test valide** ist, müssen alle **Annahmen A1-A3 und B1-B4 erfüllt** sein. Wir nehmen an dieser Stelle an, dass diese Annahmen erfüllt sind (siehe aber die Diskussion dazu in Einheit 5).

## Zweiseitiger Test: Beispiel

Wir testen die Hypothese, ob es einen **Wirkungszusammenhang zwischen Anwesenheit und Prüfungserfolg** gibt, in den folgenden Schritten:

- Aufstellen von  $H_0$  ( $H_0 : \beta = 0$ ) und  $H_1$  ( $H_1 : \beta \neq 0$ )
- Festlegung des Signifikanzniveaus  $\alpha$ :  $\alpha = 0.10$
- Schätzung von  $se(\hat{\beta})$ :  $\widehat{se}(\hat{\beta}) = 0.88$
- Berechnung des t-Wertes:  $t = \left| \frac{\hat{\beta} - q}{sd(\hat{\beta})} \right| = \left| \frac{4.88 - 0}{0.88} \right| = 5.54$
- Ermittlung des kritischen Wertes  $c$  für eine t-Verteilung mit 8 Freiheitsgraden:  $c = t_{8;0.95} = 1.86$
- Vergleichen von  $c$  und  $t$ :  $H_0$  wird verworfen, da  $|5.54| > 1.86$ .
- **Interpretation:**
  - ▶ Unser Test ergibt, dass auf einem 10 % Signifikanzniveau die **Anwesenheit einen Effekt auf den Studienerfolg** hat. Das bedeutet, dass die Wahrscheinlichkeit, dass es auch in der **Grundgesamtheit** einen Effekt zwischen Anwesenheit und Studienerfolg gibt, **zumindest 90 %** beträgt.
  - ▶ Der Hypothesentest ist sehr ähnlich zu der Berechnung von Konfidenzintervallen.

# Einseitiger Hypothesentest

- Oft interessiert uns nicht nur, ob  $\beta$  unterschiedlich von  $q$  ist, sondern auch, ob der Parameter **größer oder kleiner** als  $q$  ist. Wir sprechen dann von einem **einseitigen Hypothesentest**.
- Wenn wir an der (*rechtsseitigen*) Hypothese, ob  $\beta > q$  interessiert sind,
  - ▶ formulieren wir die Hypothesen so:  $H_0 : \beta \leq q$ ,  $H_1 : \beta > q$
  - ▶ sprechen wir von einem **rechtsseitigen Hypothesentest**.
- Wenn wir an der (*linksseitigen*) Hypothese, ob  $\beta < q$  interessiert sind,
  - ▶ formulieren wir die Hypothesen so:  $H_0 : \beta \geq q$ ,  $H_1 : \beta < q$
  - ▶ sprechen wir von einem **linksseitigen Hypothesentest**.
- **Beachten Sie:**
  - ▶ Die Aussage, die wir zeigen möchten (an der wir interessiert sind), formulieren wir immer als Alternativhypothese  $H_1$ .
  - ▶  $\beta = q$  ist immer Teil der Nullhypothese  $H_0$ .
  - ▶ Der kritische Wert  $c = t_{n-2; 1-\alpha}$  ist kleiner als beim zweiseitigen Test (wo gilt:  $c = t_{n-2; 1-\frac{\alpha}{2}}$ ). Um die Nullhypothese  $H_0$  zu verwerfen, muss beim einseitigen Test neben  $t > c$  auch noch gelten:
    - beim rechtsseitigen Test:  $\hat{\beta} > q$
    - beim linksseitigen Test:  $\hat{\beta} < q$

# Einseitiger (rechtsseitiger) Hypothesentest: Anleitung

Um einen **einseitigen (rechtsseitiger) Hypothesentest** durchzuführen (wenn Sie interessiert sind, ob  $\beta > q$ ), müssen Sie folgende Schritte durchführen:

- Aufstellen von  $H_0$  ( $H_0 : \beta \leq q$ ) und  $H_1$  ( $H_1 : \beta > q$ )
- Festlegung des Signifikanzniveaus  $\alpha$
- Schätzung von  $\widehat{se}(\widehat{\beta})$
- Berechnung des t-Wertes:  $t = \left| \frac{\widehat{\beta} - q}{sd(\widehat{\beta})} \right|$
- Ermittlung des kritischen Wertes  $c$ :  $c = t_{n-2; 1-\alpha}$
- Vergleichen von  $c$  und  $t$ : falls  $t > c$  und  $\widehat{\beta} > q$  wird  $H_0$  verworfen

# Einseitiger (linksseitiger) Hypothesentest: Anleitung

Um einen **einseitigen (linksseitigen) Hypothesentest** durchzuführen (wenn Sie interessiert sind, ob  $\beta < q$ ), müssen Sie folgende Schritte durchführen:

- Aufstellen von  $H_0$  ( $H_0 : \beta \geq q$ ) und  $H_1$  ( $H_1 : \beta < q$ )
- Festlegung des Signifikanzniveaus  $\alpha$
- Schätzung von  $\widehat{se}(\widehat{\beta})$
- Berechnung des t-Wertes:  $t = \left| \frac{\widehat{\beta} - q}{sd(\widehat{\beta})} \right|$
- Ermittlung des kritischen Wertes  $c$ :  $c = t_{n-2; 1-\alpha}$
- Vergleichen von  $c$  und  $t$ : falls  $t > c$  und  $\widehat{\beta} < q$  wird  $H_0$  verworfen



# Einseitiger (rechtsseitiger) Hypothesentest: Beispiel I

- Es wird überlegt, eine **Anwesenheitspflicht** einzuführen. Dies sollte aber nur geschehen, wenn es genug **Evidenz** dafür gibt, dass **Anwesenheit die Klausurergebnisse verbessert**. Wir sind daher interessiert, ob es sehr wahrscheinlich ist, dass  $\beta > 0$ .
- Um eine **Empfehlung** abgeben zu können, ob eine Anwesenheitspflicht die Klausurergebnisse verbessert, müssen wir die entsprechenden Hypothesen formulieren:

$$H_0 : \beta \leq 0; \quad H_1 : \beta > 0$$

- Von unserer Hypothese sehen wir, dass wir einen **einseitiger Hypothesentest** durchführen müssen. Wir definieren die interessante Hypothese als  $H_1$  (zur Erinnerung:  $H_0$  abzulehnen ist immer ein stärkeres Signal).
- Um einen Hypothesentest durchführen zu können, müssen wir außerdem wieder ein **Signifikanzniveau** festlegen. Da die Einführung einer Anwesenheitspflicht einen starken Eingriff in die Autonomie der Studierenden darstellt, wollen wir erst dann eine Empfehlung (für eine Anwesenheitspflicht) abgeben, wenn wir uns sehr sicher sind, dass die Anwesenheitspflicht auch tatsächlich (in der Grundgesamtheit) einen positiven Effekt hat. Wir legen daher eine **sehr niedrige Irrtumswahrscheinlichkeit** von 1% fest:  $\alpha = 0.01$

# Einseitiger (rechtsseitiger) Hypothesentest: Beispiel I

Um einen **einseitigen (rechtsseitiger) Hypothesentest** durchzuführen, müssen Sie folgende Schritte durchführen:

- Aufstellen von  $H_0$  ( $H_0 : \beta \leq 0$ ) und  $H_1$  ( $H_1 : \beta > 0$ )
- Festlegung des Signifikanzniveaus  $\alpha$ :  $\alpha = 0.01$
- Schätzung von  $se(\hat{\beta})$ :  $\widehat{se}(\hat{\beta}) = 0.88$
- Berechnung des t-Wertes:  $t = \left| \frac{\hat{\beta} - q}{sd(\hat{\beta})} \right| = \left| \frac{4.88 - 0}{0.88} \right| = 5.54$
- Ermittlung des kritischen Wertes  $c$ :  $c = t_{n-2; 1-\alpha} = t_{8; 0.99} = 2.90$
- Vergleichen von  $c$  und  $t$ : Da  $t > c$  und  $\hat{\beta} > 0$  wird  $H_0$  verworfen.

**Schlussfolgerung:** Es ist **sehr wahrscheinlich** (zu mehr als 99%), dass es in der Grundgesamtheit einen **positiven Effekt** von Anwesenheit auf den Klausurerfolg gibt. Man könnte daher eine **entsprechende Empfehlung abgeben**.

## Einseitiger (rechtsseitiger) Hypothesentest: Beispiel II

Sie überlegen, ob Sie das Fachwissenschaftliche Seminar besuchen, oder ein Fußballspiel im Fernsehen ansehen sollen. Sie wollen eigentlich das Spiel sehen, aber auch ein gutes Klausurergebnis erzielen. Sie beschließen, die Lehrveranstaltung nur dann zu besuchen, wenn Sie zu 90 % sicher sein können, dass die Anwesenheit Ihr Klausurergebnis um zumindest 4 Punkte verbessert.

- Aufstellen von  $H_0$  ( $H_0 : \beta \leq 4$ ) und  $H_1$  ( $H_1 : \beta > 4$ )
- Festlegung des Signifikanzniveaus  $\alpha$ :  $\alpha = 0.10$
- Schätzung von  $se(\hat{\beta})$ :  $\widehat{se}(\hat{\beta}) = 0.88$
- Berechnung des t-Wertes:  $t = \left| \frac{\hat{\beta} - q}{sd(\hat{\beta})} \right| = \left| \frac{4.88 - 4}{0.88} \right| = 1.00$
- Ermittlung des kritischen Wertes  $c$ :  $c = t_{n-2; 1-\alpha} = t_{8; 0.90} = 1.40$
- Vergleichen von  $c$  und  $t$ : Da  $t < c$  (obwohl  $\hat{\beta} > 4$ ) kann  $H_0$  nicht verworfen werden.

**Schlussfolgerung:** Sie sollten sich das Fußballspiel anschauen.

# Umsetzung in EXCEL

- **Daten** → **Analyse** → **Datenanalyse** → **Regression**

	<i>Koeff.</i>	<i>Stand.-f.</i>	<i>t-Stat.</i>	<i>P-Wert</i>	<i>Untere 95%</i>	<i>Obere 95%</i>
Schnittpunkt	5.233	12.117	0.432	0.6772	-22.710	33.175
Anwesenheit	4.883	0.881	5.544	0.0005	2.852	6.914

- **EXCEL** berechnet **automatisch** folgende Maßzahlen:

- ▶ *Koeffizient* =  $\hat{\beta}$  (Punktschätzer)
- ▶ *Standardfehler* =  $\widehat{se}(\hat{\beta})$
- ▶ *t-Statistik* =  $\frac{\hat{\beta}-0}{\widehat{se}} = \frac{\hat{\beta}}{\widehat{se}}$  ist der t-Wert für den zweiseitigen Hypothesentest mit  $H_0 : \beta = 0$ . (Falls t-Statistik < 0 muss der Betrag genommen werden.)
- ▶ *P-Wert* ist die dazugehörige Wahrscheinlichkeit.
- ▶ *Untere* bzw. *Obere 95 %* gibt den Intervallschätzer für das 95 %-Konfidenzintervall an. Es kann wahlweise noch ein weiteres (z.B. 90 % oder 99 %) Intervall ausgewiesen werden.
- ▶ Die Nullhypothese  $H_0 : \beta = 0$  wird verworfen wenn der *P-Wert* kleiner als das Signifikanzniveau  $\alpha$  ist.

# Umsetzung in EXCEL

- **Daten** → **Analyse** → **Datenanalyse** → **Regression**

	<i>Koeff.</i>	<i>Stand.-f.</i>	<i>t-Stat.</i>	<i>P-Wert</i>	<i>Untere 95%</i>	<i>Obere 95%</i>
Schnittpunkt	5.233	12.117	0.432	0.6772	-22.710	33.175
Anwesenheit	4.883	0.881	5.544	0.0005	2.852	6.914

- Um **andere Hypothesen** zu testen, muss:

- ▶ selbständig der t-Wert mit  $t = \left| \frac{\hat{\beta} - q}{\widehat{se}} \right|$  berechnet werden.
- ▶ der kritische Wert  $c$  der t-Verteilung bestimmt werden:
  - ★ mit T.INV( $1 - \alpha/2$ ;  $n - 2$ ) oder T.INV.2S( $\alpha$ ;  $n - 2$ ) bei zweiseitigem Hypothesentest
  - ★ mit T.INV( $1 - \alpha$ ;  $n - 2$ ) bei einseitigem Hypothesentest
- ▶ Alternativ dazu kann direkt der p-Wert berechnet werden:
  - ★ T.VERT.2S(|t-Wert|;  $n - 2$ ): p-Wert bei zweiseitigem Hypothesentest; wenn p-Wert <  $\alpha$  wird  $H_0$  verworfen
  - ★ T.VERT.RE(|t-Wert|;  $n - 2$ ): p-Wert bei einseitigem Hypothesentest; wenn p-Wert <  $\alpha$  wird  $H_0$  verworfen (und  $\hat{\beta} > q$  bei einem rechtsseitigen Test bzw.  $\hat{\beta} < q$  bei einem linksseitigen Test)

# Hypothesentest: Zusammenfassung

- Um **Hypothesen für die Grundgesamtheit** zu testen, muss die **statistische Unsicherheit**, die mit der **Schätzungen der Modellparameter** verknüpft ist, berücksichtigt werden.
- Die Hypothese, an der wir interessiert sind, wird immer als Alternativhypothese  $H_1$  formuliert.
- Wenn man daran interessiert ist, ob sich der Parameter von einem bestimmten Wert unterscheidet (d.h.  $H_1 : \beta \neq q$ ), muss ein **zweiseitiger Hypothesentest** durchgeführt werden.
- Wenn man daran interessiert ist, ob der Parameter größer als ein bestimmter Wert ist (d.h.  $H_1 : \beta > q$ ), muss ein **einseitiger (rechtsseitiger) Hypothesentest** durchgeführt werden.
- Wenn man daran interessiert ist, ob der Parameter kleiner als ein bestimmter Wert ist (d.h.  $H_1 : \beta < q$ ), muss ein **einseitiger (linksseitiger) Hypothesentest** durchgeführt werden.
- Die Hypothesentests liefern nur dann gültige Ergebnisse, wenn alle **Annahmen A1-A3 und B1-B4 erfüllt** sind.