

SE Geographie und Ökonomie

Einheit 3:

Univariate Datenanalyse: Schließende Statistik

Dieter Pennerstorfer

dieter.pennerstorfer@jku.at

Institut für Volkswirtschaftslehre
Johannes Kepler Universität Linz Linz



JOHANNES KEPLER
UNIVERSITÄT LINZ

Motivation

In der Praxis ist es oft nicht möglich, alle interessierenden Objekte der Grundgesamtheit zu untersuchen, da **Vollerhebungen teuer und zeitaufwendig** sind. Man muss sich dann mit einer Auswahl aus dieser Grundgesamtheit, der Stichprobe, begnügen. Das Interesse gilt aber üblicherweise der Grundgesamtheit. Die schließende (induktive) Statistik stellt Methoden bereit, **um Rückschlüsse von einer (Zufalls-) Stichprobe auf die Grundgesamtheit** zu ermöglichen. Diese Rückschlüsse sind mit einer **Unsicherheit** verbunden, die aber (unter gewissen Voraussetzungen) quantifiziert werden kann.

- Beispiele: Konsumerhebung, Befragungen zu Schattenwirtschaft, Wahlumfragen
- Interessen an Grundgesamtheit: Bei einer Wahlumfrage unter 1,000 WählerInnen geben 23 % der Befragten an, die Partei X wählen zu wollen. Man ist aber üblicherweise nicht daran interessiert, wie groß der Stimmenanteil der Partei X innerhalb der Stichprobe (den 1,000 Befragten) ist, sondern wie groß der Stimmenanteil der Partei X in der Grundgesamtheit ist.

Lernziele der Einheit 3

- Sie kennen den Unterschied zwischen der **Verteilung einer Zufallsstichprobe** und der **Verteilung der Mittelwerte von Zufallsstichproben**.
- Sie wissen (ausgehend von der Grundgesamtheit), wie die Mittelwerte von Zufallsstichproben verteilt sind.
- Sie können (ausgehend von der Grundgesamtheit) ein Intervall bestimmen, sodass der **Mittelwert einer Zufallsstichprobe** mit einer gewissen Wahrscheinlichkeit **innerhalb des Intervalls** liegt.
- Sie können (ausgehend von einer Stichprobe) ein Intervall bestimmen, sodass der **Mittelwert einer Grundgesamtheit** mit einer gewissen Wahrscheinlichkeit **innerhalb des Intervalls** liegt, sowohl
 - ▶ wenn die Standardabweichung der Grundgesamtheit bekannt ist, bzw.
 - ▶ wenn die Standardabweichung der Grundgesamtheit unbekannt ist.
- Sie erkennen den Zusammenhang zwischen Streuung eines Merkmals (einer Variable) und dem Stichprobenumfang einerseits, sowie der Größe des Konfidenzintervalls bzw. der statistischen Unsicherheit andererseits.
- Sie können einen **einseitigen** und einen **zweiseitigen Hypothesentest über Mittelwerte** durchführen.

Statistische Unsicherheit und die Rolle des Zufalls:

Der Untergang der Titanic

Tabelle: Häufigkeitsverteilung der Grundgesamtheit

Klasse	Ausprägung	Absolute Häufigkeiten	Relative Häufigkeiten	Relative Häufigkeiten (in Prozent)
m	x_m	h_m	p_m	P_m (in %)
1	verstorben	1,490	0.677	67.7
($r =$) 2	überlebt	711	0.323	32.3
Summe	($N =$)	2,201	1	100

Tabelle: Häufigkeitsverteilung der (Zufalls-) Stichprobe

m	x_m	h_m	p_m	P_m (in %)
1	verstorben	69	0.690	69.0
($r =$) 2	überlebt	31	0.310	31.0
Summe	($n =$)	100	1	100

Warum haben genaue 711 Personen überlebt?

- Ob eine Person überlebt hat, **hängt** von den Eigenschaften der Person (Alter, Körpergröße, Geschlecht, Kabinenkategorie, ...), aber **auch vom Zufall ab**. Unterstellen wir kurz, dass die Personeneigenschaften keine Rolle spielen, und jede Person eine Überlebenschance von 33 % hat. Das Merkmal x ($= 1$ wenn "überlebt", und 0 sonst) ist daher eine Zufallsvariable, die einer Verteilungsfunktion folgt, die durch den Erwartungswert $\mu = 0.330$ und eine Standardabweichung von $\sigma = 0.470$ (bzw. eine Varianz von $\sigma^2 = 0.221$) gekennzeichnet ist. Diese zugrunde liegende Verteilung **beobachten wir nicht** (ich habe mir die Zahlen nur ausgedacht).
- Zufall 1: Wir beobachten lediglich die konkreten Realisationen eines Zufallsexperimentes: von 2,201 Passagieren haben 711 ($\bar{x} = 0.323$) überlebt. Die Standardabweichung $s = 0.219$. Würden wir die Titanic nochmals losschicken und untergehen lassen, dann **würden vermutlich nicht exakt 711 Leute überleben**, sondern wahrscheinlich etwas mehr oder etwas weniger. Wie viel mehr oder weniger überleben, hängt vom **Zufall** ab.
- Zufall 2: Wenn wir eine Zufallsstichprobe ziehen, dann wird der Mittelwert \bar{x}_{100} und die Standardabweichung \hat{s} von der Grundgesamtheit abweichen, weil der Anteil der Überlebenden in der Stichprobe **zufälligerweise** höher oder niedriger als in der Grundgesamtheit (= alle Passagiere) ist.

Schließende Statistik

- Mittels **kausaler Inferenz** kann man ursächliche Mechanismen hinter beobachteten Zusammenhängen untersuchen.
Beispiel: Ist der größere Anteil an Überlebenden bei Gästen der 1. Klasse, der in der Grundgesamtheit beobachtet wird ($\bar{x}_{1.Klasse} = 0.62 > \bar{x}_{3.Klasse} = 0.25$), auf eine höhere Überlebenswahrscheinlichkeit zurückzuführen (d.h. $\mu_{1.Klasse} > \mu_{3.Klasse}$), oder kann der Unterschied lediglich durch Zufall erklärt werden, obwohl beide Gruppen die gleiche Überlebenswahrscheinlichkeit hatten ($\mu_{1.Klasse} = \mu_{3.Klasse}$)?
- Mittels **deskriptiver Inferenz** kann man auf Basis einer Stichprobe Rückschlüsse auf die Grundgesamtheit ziehen. Beispiel: Wenn in der Stichprobe der Anteil an Überlebenden bei Gästen der 1. Klasse größer ist als bei Gästen der 3. Klasse ($\bar{x}_{100;1.Klasse} = 0.60 > \bar{x}_{100;3.Klasse} = 0.27$), kann daraus geschlossen werden, dass auch in der Grundgesamtheit der Anteil der Überlebenden bei Gästen der 1. Klasse größer ist?

Wir versuchen hier vor allem, **Rückschlüsse von einer Stichprobe auf die Grundgesamtheit zu ziehen** (deskriptive Inferenz). Grundlage der Analyse ist eine Zufallsstichprobe aus der Grundgesamtheit. Aus Gründen der Übersichtlichkeit bezeichnen wir den Mittelwert der Grundgesamtheit mit μ und die Standardabweichung mit σ .

Anmerkung: Da die Grundgesamtheit üblicherweise sehr groß ist (etwa: 8.9 Mio. EinwohnerInnen in Ö), ist der Unterschied zwischen dem Mittelwert (eigentlich: \bar{x}) und Standardabweichung (eigentlich: s) der Grundgesamtheit und Mittelwert und Standardabweichung des Zufallsprozesses, der das Merkmal x erzeugt (nämlich μ und σ) vernachlässigbar klein.

Grundgesamtheit vs. Stichprobe

Die deskriptive Statistik beschäftigt sich mit der Verteilung einzelner Merkmale. Entsprechende Datensätze können entweder alle interessierenden Objekte umfassen (Vollerhebung einer **Grundgesamtheit**), oder nur einen Auszug aus dieser Grundgesamtheit (**Zufallsstichprobe**). In der Praxis wird der Datensatz meistens eine Stichprobe beinhalten. Zwei Aspekte sind hier interessant:

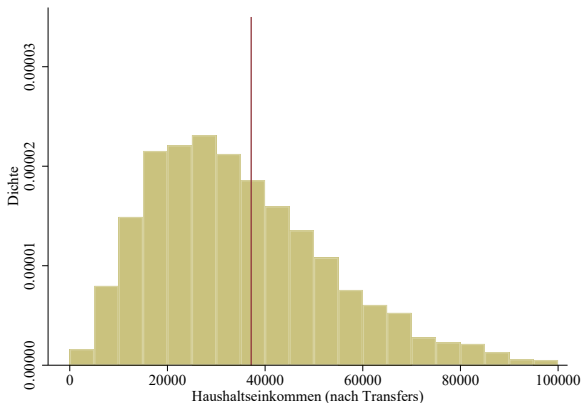
- i) Wenn wir die **Grundgesamtheit** kennen, können wir daraus etwas über die **Stichprobe** lernen.
- ii) Wenn wir die **Stichprobe** kennen, können wir daraus etwas über die **Grundgesamtheit** lernen. Die **schließende (induktive) Statistik** stellt Methoden bereit, um Rückschlüsse von einer Stichprobe auf die Grundgesamtheit zu ermöglichen.

Der Aspekt i) ist von geringer Relevanz, wir werden aber aus didaktischen Gründen damit starten.

Wir unterstellen für diese Einheit, dass **unser GSOEP-Datensatz alle Personen der Grundgesamtheit** (mit N Personen) umfasst, aus der wir zufällig eine Stichprobe vom Umfang n ziehen.

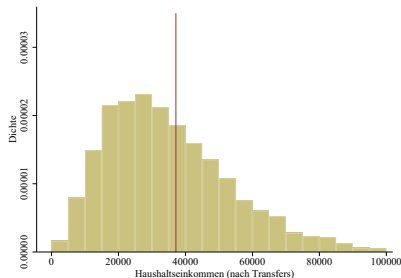
“Grundgesamtheit”

Wir beobachten das Haushaltseinkommen von $N = 5,407$ Personen (= Grundgesamtheit). Der Mittelwert $\mu = 37,150$ Euro, und die Standardabweichung $\sigma = 26,728$. Das Histogramm (Intervallbreite $d_i = 5,000$ Euro) sieht folgendermaßen aus: (vertikale rote Linie = Mittelwert)



Grundgesamtheit vs. Stichprobe: Beispiel (1)

Histogramm Grundgesamtheit

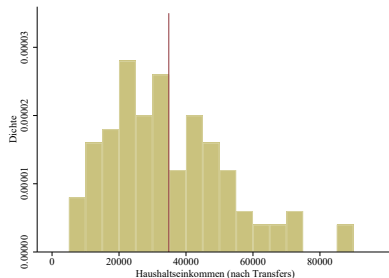


$$N = 5,407$$

$$\mu = 37,150$$

$$\sigma = 26,728$$

Histogramm Stichprobe



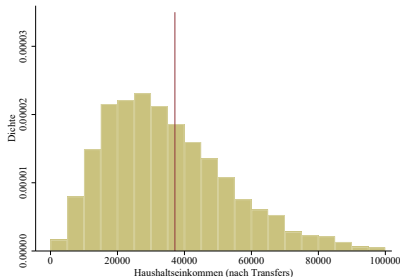
$$n = 100$$

$$\bar{x}'_{100} = 34,864$$

$$\hat{s}'_{100} = 17,637$$

Grundgesamtheit vs. Stichprobe: Beispiel (2)

Histogramm Grundgesamtheit

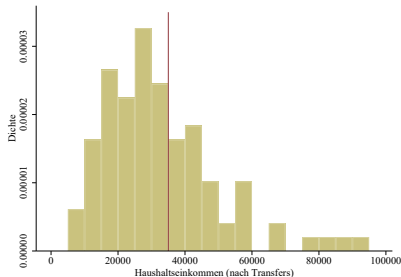


$$N = 5,407$$

$$\mu = 37,150$$

$$\sigma = 26,728$$

Histogramm Stichprobe



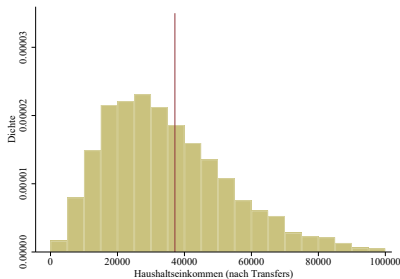
$$n = 100$$

$$\bar{x}''_{100} = 35,019$$

$$\hat{s}''_{100} = 24,460$$

Grundgesamtheit vs. Stichprobe: Beispiel (3)

Histogramm Grundgesamtheit

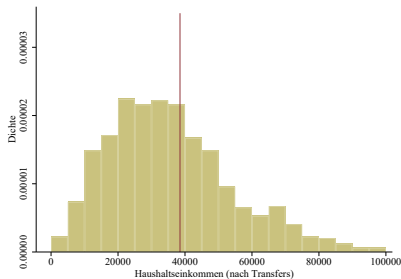


$$N = 5,407$$

$$\mu = 37,150$$

$$\sigma = 26,728$$

Histogramm Stichprobe



$$n = 1,000$$

$$\bar{x}'_{1000} = 38,509$$

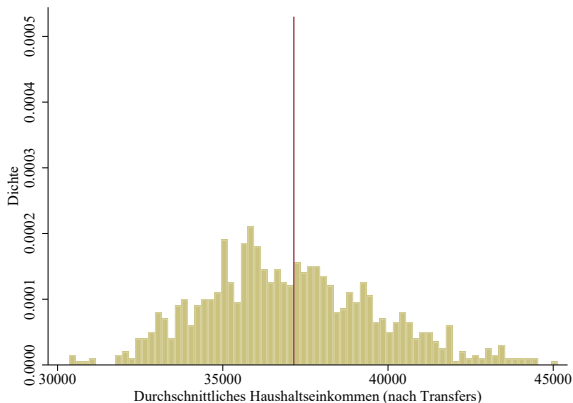
$$\hat{s}'_{1000} = 32,111$$

Erkenntnisse: Intuition 1 + 2

- Erkenntnis 1: Wenn ich eine Zufallsstichprobe ziehe, sieht die Verteilung dieser Stichprobe sehr ähnlich aus wie die Verteilung der Grundgesamtheit.
- Erkenntnis 2: Die Ähnlichkeit ist umso größer, je größer der Stichprobenumfang n ist.

Häufigkeitsverteilung der Mittelwerte (1)

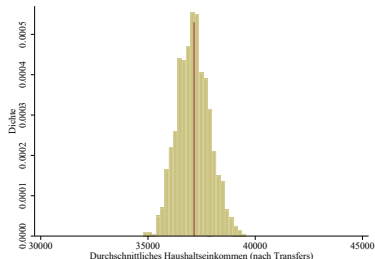
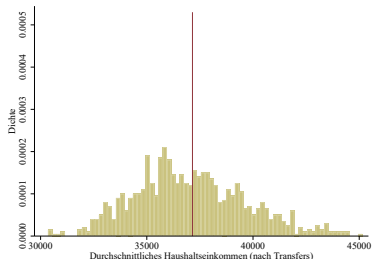
Ich ziehe eine Zufallsstichprobe von $n = 100$ Personen und berechne den Mittelwert der Stichprobe, \bar{x}_{100} . Ich wiederhole den Vorgang 1,000 Mal und erhalte daher 1,000 verschiedene Stichprobenmittelwerte (also \bar{x}'_{100} , \bar{x}''_{100} , \bar{x}'''_{100} , usw.). Die Verteilung der Stichprobenmittelwerte kann als Histogramm dargestellt werden:



(vertikale rote Gerade markiert den "wahren" Mittelwert $\mu = 37,150$)

Häufigkeitsverteilung der Mittelwerte (2)

Würde ich Stichproben vom Umfang $n = 1,000$ ziehen, würden die Stichprobenmittelwerte weniger stark streuen (rechte Grafik). Der Mittelwert einer Zufallsstichprobe mit $n = 1,000$ ist daher höchstwahrscheinlich "in der Nähe" vom Mittelwert μ der Grundgesamtheit. Bei einer kleinen Stichprobe ist dies weniger wahrscheinlich:



	N	Mittelwert	Std. Abw.	min	max
x (HH-Einkommen)	5,407	37,150	26,728	583	507,369
\bar{x} ($n = 100$)	1,000	37,060	2,677	30,053	47,483
\bar{x} ($n = 1,000$)	1,000	37,125	755	34,573	39,920

Erkenntnisse: Intuition 3 + 4

- Erkenntnis 3: Wenn ich mehrmals eine Stichprobe ziehe, streuen die Stichprobenmittelwerte um den Mittelwert der Grundgesamtheit.
- Erkenntnis 4: Die Streuung nimmt ab, je größer der Stichprobenumfang n ist.

Erkenntnisse: Formal

- x sei eine diskrete oder stetige (Zufalls-) Variable, die eine beliebige Verteilung aufweist, und den Erwartungswert μ und die Varianz σ^2 besitzt. (Wir nehmen an, dass die Variable x in der Grundgesamtheit diese Eigenschaften erfüllt.) Wenn man daraus eine Zufallsstichprobe vom Umfang n zieht, **weist diese Stichprobe (im Erwartungswert) dieselbe Verteilung** auf. Das bedeutet, dass:

- i) der Mittelwert der Stichprobe:

$$\bar{x}_n = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{n} (x_1 + x_2 + \cdots + x_n)$$

im Erwartungswert μ entspricht (d.h. $E(\bar{x}_n) = \mu$).

- ii) die Varianz der Stichprobe:

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

im Erwartungswert σ^2 entspricht (d.h. $E(\hat{s}^2) = \sigma^2$).

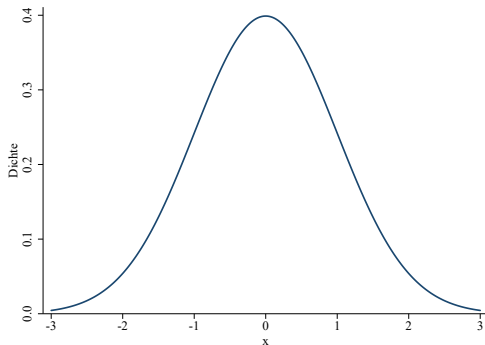
- \bar{x}_n ist selbst eine Zufallsvariable, die einer **Normalverteilung** mit Mittelwert μ und Varianz σ^2/n folgt (**zentraler Grenzwertsatz**): $\bar{x}_n \sim N(\mu, \sigma^2/n)$
- Die Standardabweichung der Verteilung der Mittelwerte wird auch **Standardfehler** (des Mittelwertes) genannt, und mit *se* (*standard error*) bezeichnet.

Was ist eine Normalverteilung? (1)

Ein **Normalverteilung** ist durch **Mittelwert** μ und **Varianz** σ^2 **gekennzeichnet** und wird als $N(\mu, \sigma^2)$ bzw. $NV(\mu, \sigma^2)$ beschrieben. Eine stetige Zufallsvariable x heißt normalverteilt, wenn sie eine Dichte der Form

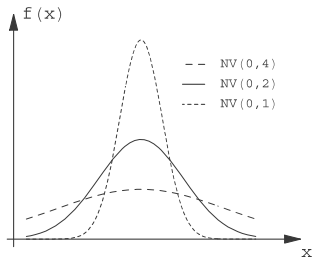
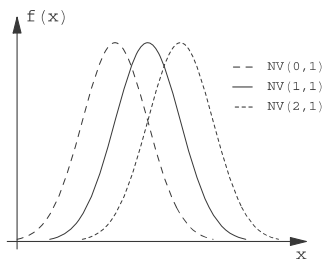
$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)} \quad \text{mit } \sigma \geq 0$$

besitzt. (Brauchen Sie nicht wissen.) Eine Normalverteilung $N(\mu = 0, \sigma^2 = 1)$ heißt **Standardnormalverteilung** und sieht so aus:



Was ist eine Normalverteilung? (2)

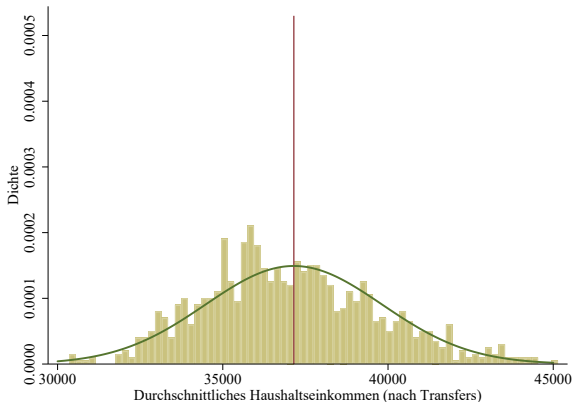
Normalverteilungen mit unterschiedlichen Lagemaßen (links) und unterschiedlichen Streuungsmaßen (rechts).



(siehe Duller, 2019, Abbildung 11.5)

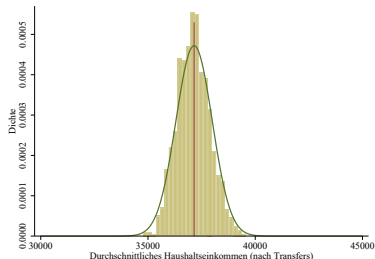
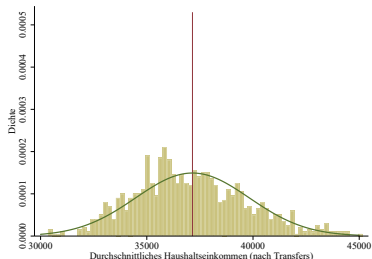
Häufigkeitsverteilung der Mittelwerte (1)

Kann durch Normalverteilung mit $\mu = 37,150$ und Standardabweichung $\sigma/\sqrt{n} = 26,728/\sqrt{100} = 2,672.8$ approximiert werden:



Häufigkeitsverteilung der Mittelwerte (2)

Würde ich Stichproben vom Umfang $n = 1,000$ ziehen, würden die Stichprobenmittelwerte weniger stark streuen (rechte Grafik). Der Mittelwert einer Zufallsstichprobe mit $n = 1,000$ ist daher höchstwahrscheinlich "in der Nähe" vom Mittelwert μ der Grundgesamtheit. Bei einer kleinen Stichprobe ist dies weniger wahrscheinlich:



	N	Mittelwert	Std. Abw.	min	max
x (HH-Einkommen)	5,407	37,150	26,728	583	507,369
\bar{x} ($n = 100$)	1,000	37,060	2,677	30,053	47,483
\bar{x} ($n = 1,000$)	1,000	37,125	755	34,573	39,920

Zusammenfassung und Schlussfolgerung

$\bar{x}_n \sim N(\mu, \sigma^2/n)$ bedeutet, dass der Stichprobenmittelwert \bar{x}_n ein **erwartungstreuer und konsistenter Schätzer** des Mittelwerts der Grundgesamtheit μ ist.

Erklärung:

- **Erwartungstreue:**

Der Erwartungswert des Schätzers (des Stichprobenmittelwerts) entspricht dem gesuchten Parameter (hier: dem Mittelwerts der Grundgesamtheit):

$$E(\bar{x}_n) = \mu$$

- **Konsistenz:**

Mit zunehmendem Stichprobenumfang wird die Varianz des Schätzers kleiner:

$$\lim_{n \rightarrow \infty} \text{Var}(\bar{x}_n) = \lim_{n \rightarrow \infty} \frac{\sigma^2}{n} \rightarrow 0$$

Was kann ich von der Grundgesamtheit über die Stichprobe lernen?

Ich kenne:

- den wahren Mittelwert der Grundgesamtheit: $\mu = 37,150$
- die Standardabweichung der Grundgesamtheit: $\sigma = 26,728$

Ich weiß:

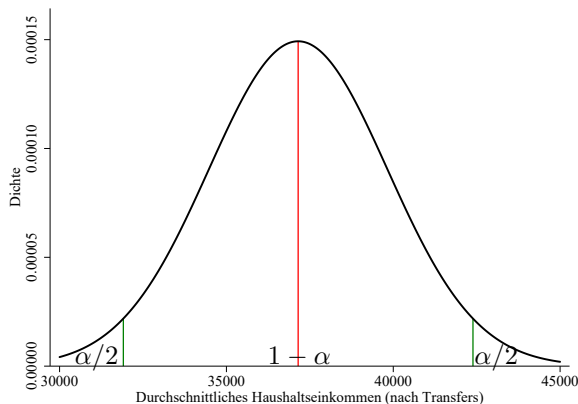
- dass der Mittelwert des Merkmals x einer Zufallsstichprobe vom Umfang n , \bar{x}_n , normalverteilt ist, mit Mittelwert μ und Standardfehler $se = \sigma/\sqrt{n}$.
Kurz: $\bar{x}_n \sim N(\mu, \sigma^2/n)$

Ich möchte:

- ein Intervall bestimmen, sodass der Mittelwert des Merkmals der Zufallsstichprobe, \bar{x}_n , mit einer bestimmten Wahrscheinlichkeit $1 - \alpha$ innerhalb des Intervalls (bzw. mit einer Wahrscheinlichkeit α außerhalb des Intervalls) liegt.

Intervall: Intuition

Verteilung der Stichprobenmittelwerte eines Merkmals x mit dem Stichprobenumfang $n = 100$, \bar{x}_{100} , mit $\mu = 37,150$ und Standardfehler $se = \sigma/\sqrt{n} = 26,728/\sqrt{100} \approx 2,673$ für $\alpha = 0.05$ (5 %):



Intervall: Berechnung

- $u_{1-\frac{\alpha}{2}}$ bezeichnet das $(1 - \frac{\alpha}{2})$ -Quantil der Standardnormalverteilung
- für $\alpha = 0.05$ ist $u_{1-\frac{\alpha}{2}} = u_{0.975} = 1.96$
- das bedeutet, dass eine standardnormalverteilte Zufallszahl zu 2.5 % kleiner als -1.96 ist, zu 2.5 % größer als +1.96, und zu 95 % zwischen -1.96 und +1.96 liegt.
→ siehe **EXCEL-Befehl** `NORM.INV(0.025;0;1)` und `NORM.INV(0.975;0;1)` zur Berechnung von Quantilen der Standardnormalverteilung.
- da \bar{x}_n normalverteilt, aber nicht standardnormalverteilt ist ($\bar{x}_n \sim N(\mu, se^2 = \sigma^2/n)$), müssen die Intervallgrenzen angepasst werden:
Untergrenze: $\mu - u_{1-\frac{\alpha}{2}} \cdot se = 37,150 - 1.96 \cdot 2,673 = 31,911$
Obergrenze: $\mu + u_{1-\frac{\alpha}{2}} \cdot se = 37,150 + 1.96 \cdot 2,673 = 42,389$
→ in `NORM.INV` kann auch gleich direkt Mittelwert und Standardfehler eingesetzt werden, um die Intervallgrenzen zu berechnen!

Intervall: Zusammenfassung

Allgemein gilt:

- für $(1 - \alpha)$ der gezogenen Stichproben:

$$\mu - u_{1-\frac{\alpha}{2}} \cdot se \leq \bar{x}_n \leq \mu + u_{1-\frac{\alpha}{2}} \cdot se$$

oder

$$\mu - u_{1-\frac{\alpha}{2}} \cdot se < \bar{x}_n < \mu + u_{1-\frac{\alpha}{2}} \cdot se$$

- für 95 % der gezogenen Stichproben (d.h. bei $\alpha = 0.05$):

$$\mu - 1.96 \cdot se \leq \bar{x}_n \leq \mu + 1.96 \cdot se$$

Im konkreten Beispiel gilt:

- Wenn ich eine Stichprobe vom Umfang $n = 100$ ziehe, dann liegt der Stichproben-Mittelwert der Ausprägung "Haushaltseinkommen (nach Transfers)" zu 95 % zwischen 31,911 Euro und 42,389 Euro.

Was kann ich von der Stichprobe über die Grundgesamtheit lernen?

Ich kenne:

- den wahren Mittelwert der Grundgesamtheit: $\mu = 37,150$
- die Standardabweichung der Grundgesamtheit: $\sigma = 26,728$

Ich weiß:

- dass der Mittelwert des Merkmals x einer Zufallsstichprobe vom Umfang n , \bar{x}_n , normalverteilt ist, mit Mittelwert μ und Standardfehler $se = \sigma/\sqrt{n}$.
Kurz: $\bar{x}_n \sim N(\mu, \sigma^2/n)$
- dass \bar{x}_n ein **erwartungstreuer Schätzer** von μ ist: $E(\bar{x}_n) = \mu$
- dass \bar{x}_n ein **konsistenter Schätzer** von μ ist:
$$\lim_{n \rightarrow \infty} Var(\bar{x}_n) = \lim_{n \rightarrow \infty} \frac{\sigma^2}{n} \rightarrow 0$$

Ich möchte:

- ein Intervall bestimmen, sodass der (unbeobachtete) Mittelwert der Grundgesamtheit μ mit einer bestimmten Wahrscheinlichkeit $1 - \alpha$ innerhalb des Intervalls (bzw. mit einer Wahrscheinlichkeit α außerhalb des Intervalls) liegt.

Logische Schlussfolgerung 1

Für $\alpha = 0.05$ gilt:

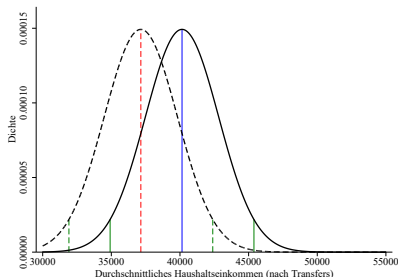
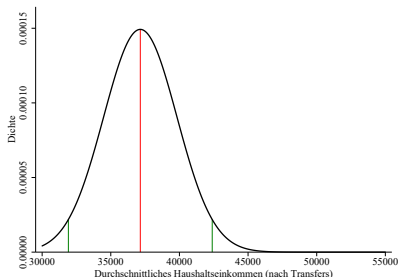
Für 95 % der gezogenen Stichproben liegen die Stichprobenmittelwerte im Intervall

$$\mu - 1.96 \cdot se \leq \bar{x}_n \leq \mu + 1.96 \cdot se$$

Daraus folgt (durch Äquivalenzumformung):

Für 95 % der gezogenen Stichproben **liegt der Mittelwert der Grundgesamtheit im Intervall**

$$\bar{x}_n - 1.96 \cdot se \leq \mu \leq \bar{x}_n + 1.96 \cdot se$$



Logische Schlussfolgerung 2

Für $\alpha = 0.05$ gilt:

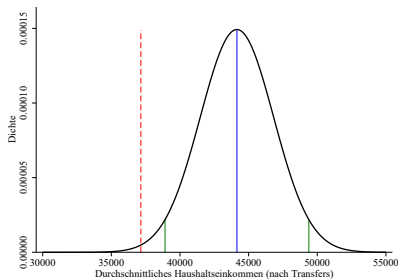
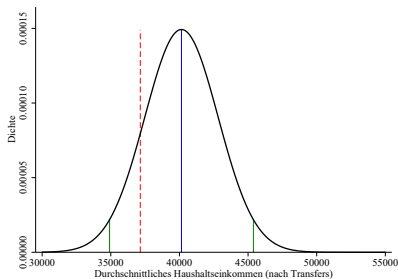
Für 5% der gezogenen Stichproben liegen die Stichprobenmittelwerte **nicht** im Intervall

$$\mu - 1.96 \cdot se \leq \bar{x}_n \leq \mu + 1.96 \cdot se$$

Daraus folgt (durch Äquivalenzumformung):

Für 5% der gezogenen Stichproben **liegt der Mittelwert der Grundgesamtheit außerhalb des Intervalls** (siehe rechte Grafik)

$$\bar{x}_n - 1.96 \cdot se \leq \mu \leq \bar{x}_n + 1.96 \cdot se$$



Was kann ich von der Stichprobe über die Grundgesamtheit lernen?

Ich kenne:

- **nichts über die Grundgesamtheit** (Standardfall!)
- den wahren Mittelwert der Grundgesamtheit: $\mu = 37,150$
- die Standardabweichung der Grundgesamtheit: $\sigma = 26,728$

Ich weiß:

- dass $\bar{x}_n \sim N(\mu, \sigma^2/n)$
 - ▶ \bar{x}_n ist ein **erwartungstreuer Schätzer** von μ ($E(\bar{x}_n) = \mu$) und ein **konsistenter Schätzer** von μ ($\lim_{n \rightarrow \infty} \text{Var}(\bar{x}_n) \rightarrow 0$)
- dass $\hat{s}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$ ein **erwartungstreuer Schätzer der (unbekannten) Varianz der Grundgesamtheit** μ ist: $E(\hat{s}^2) = \sigma^2$

Ich möchte:

- ein Intervall bestimmen, sodass der (unbeobachtete) Mittelwert der Grundgesamtheit μ mit einer bestimmten Wahrscheinlichkeit $1 - \alpha$ innerhalb des Intervalls (bzw. mit einer Wahrscheinlichkeit α außerhalb des Intervalls) liegt.

Herleitung

Ausgangslage:

$$\bar{x}_n - t_{n-1;1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{s}^2}{n}} \leq \mu \leq \bar{x}_n + t_{n-1;1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{s}^2}{n}},$$

- Da die Varianz der Stichprobe \hat{s}^2 ein erwartungstreuer Schätzer für die (unbekannte) Varianz der Grundgesamtheit σ^2 ist, **kann σ^2 durch \hat{s}^2 ersetzt werden.**
- Da die Varianz nicht bekannt ist, sondern geschätzt werden muss, **erhöht sich die Unsicherheit** der Aussage über den Parameter μ . Um diese Unsicherheit zu kompensieren, muss das Konfidenzintervall etwas breiter angelegt werden. Dies wird erreicht, indem man nicht die Quantile der Standardnormalverteilung $u_{1-\frac{\alpha}{2}}$ verwendet, sondern die Quantile $t_{n-1;1-\frac{\alpha}{2}}$ der so genannten **Student-Verteilung** (auch **t-Verteilung**).

Was ist eine Student- (t-) Verteilung?

Die Quantile der t-Verteilung hängen nicht nur vom Niveau α ab, sondern auch vom Stichprobenumfang n . Bei gleicher Wahrscheinlichkeit sind die Quantile der Student-Verteilung stets größer als die der Standardnormalverteilung. Mit wachsendem Stichprobenumfang nähern sich die Quantile der Student-Verteilung denen der Standardnormalverteilung an. (Ab ca. $n = 30$ kann die Standardnormalverteilung verwendet werden.)

Vergleich zwischen Perzentile der Standardnormal- und der Studentverteilung für $\alpha = 0.05$:

n	$u_{0.975}$	$t_{n-1;0.975}$
10	1.96	2.26
20	1.96	2.09
30	1.96	2.05
40	1.96	2.02
50	1.96	2.01
100	1.96	1.98
1000	1.96	1.96

Schlussfolgerung

Wenn ich eine Zufallsstichprobe vom Umfang n ziehe, liegen der unbekannte Mittelwert der Grundgesamtheit mit einem Anteil von $1 - \alpha$ innerhalb dieses Konfidenzintervalls:

$$\bar{x}_n - t_{n-1;1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{s}^2}{n}} \leq \mu \leq \bar{x}_n + t_{n-1;1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{s}^2}{n}},$$

wobei $\sqrt{\frac{\hat{s}^2}{n}} = \hat{s}e$ der geschätzte Standardfehler ist.

Das Intervall kann auch folgendermaßen angeschrieben werden:

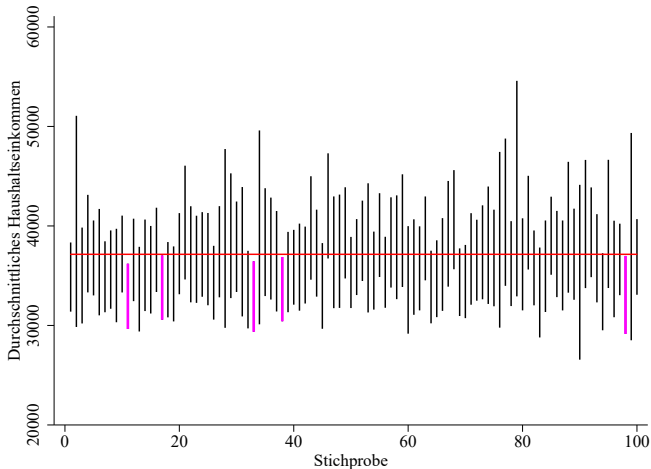
$$[\underline{\mu}, \bar{\mu}] = \bar{x}_n \pm t_{n-1;1-\frac{\alpha}{2}} \cdot \hat{s}e$$

EXCEL-Befehl:

T.INV(0.025; $n - 1$) und T.INV(0.975; $n - 1$) zur Berechnung von Quantilen der t-Verteilung (für $\alpha = 0.05$).

Illustration

95 % der Konfidenzintervalle der Zufallsstichproben enthalten den tatsächlichen Mittelwert. Dass das in diesem Beispiel für genau 95 von 100 Stichproben zutrifft, ist aber ein Zufall.



Statistisches Testen: Intuition

Um Aussagen über die Grundgesamtheit zu treffen, wenn ich nur eine Stichprobe kenne, 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**
- **Alternativhypothese** (H_1): Behauptung über die **Grundgesamtheit**
- Null- und Alternativhypothese **schließen einander aus** und **ergänzen sich** (d.h. eine der beiden Hypothesen wird angenommen)
 - ▶ schließen einander aus: Wenn H_0 richtig ist, dann muss H_1 falsch sein (und umgekehrt)
 - ▶ ergänzen sich: Es muss entweder H_0 oder H_1 richtig sein.
- 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.

Statistische Unsicherheit und Schlussfolgerungen: Gerichtsverhandlung

Es gilt die Unschuldsvermutung!

- **Nullhypothese** (H_0): Der Angeklagte ist unschuldig
- **Alternativhypothese** (H_1): Der Angeklagte ist schuldig

		Entscheidung auf	
		H_0	H_1
wahr ist	H_0	kein Fehler	α -Fehler
	H_1	β -Fehler	kein Fehler

Schlussfolgerung aus diesem Beispiel:

- α -Fehler und β -Fehler hängen zusammen. Durch die Festlegung von α habe ich den α -Fehler unter Kontrolle, den β -Fehler aber nicht.
- Hypothesen sollten so formuliert werden, dass der α -Fehler (den ich unter Kontrolle habe) der schlimmere Fehler ist.
- Wenn die Nullhypothese nicht abgelehnt wird, bedeutet das nicht, dass die Nullhypothese zutrifft.

Man wird freigesprochen, wenn die Schuld nicht bewiesen werden konnte. Das bedeutet nicht, dass dadurch die Unschuld bewiesen wurde. Man kann auch aus *Mangel an Beweisen* freigesprochen werden.

Ablauf eines Statistischen Tests

- Vorbereitung: Allgemeine und spezielle **Voraussetzungen** des Tests **überprüfen** und **Signifikanzniveau** (= Irrtumswahrscheinlichkeit; α -Fehler) **festlegen** ($\alpha = 0.01$ oder $\alpha = 0.05$).
- Durchführung des Tests:
 1. **Hypothesen formulieren:**
 - ★ Als Nullhypothese wird eine möglichst unvoreingenommene Behauptung verwendet.
 - ★ Die Hypothesen sind so zu formulieren, dass der α -Fehler den schwerwiegenden Fehler beinhaltet.
 - ★ Die **nachzuweisende Behauptung** wird als **Alternativhypothese** formuliert. Ich bin (fast) immer an der Alternativhypothese interessiert!
 2. Nach den **vorliegenden Regeln** aufgrund eines Stichprobenergebnisses eine **Entscheidung** für eine der beiden Hypothesen **treffen**.
- **Entscheidung interpretieren**

Zweiseitiger Hypothesentest über Mittelwerte

Es handelt sich um eine Zufallsstichprobe, das Verfahren ist daher zulässig. Das Signifikanzniveau wird mit $\alpha = 0.05$ festgelegt.

1. Hypothesen formulieren:

- ▶ $H_0 : \mu = \mu_0$
- ▶ $H_1 : \mu \neq \mu_0$
- ▶ \rightarrow Zweiseitige Hypothese, da H_0 abgelehnt wird, wenn \bar{x}_n signifikant größer als μ_0 ist **oder** wenn \bar{x}_n signifikant kleiner als μ_0 ist.

2. Entscheidungsregel: mittels geschätztem Konfidenzintervall

- ▶ Der unbekannte Mittelwert der Grundgesamtheit μ liegt mit einer Wahrscheinlichkeit $(1 - \alpha)$ im Intervall $[\underline{\mu}, \bar{\mu}] = \bar{x}_n \pm t_{n-1;1-\alpha/2} \cdot \hat{s}e$
mit $\hat{s}e = \sqrt{\frac{\hat{s}^2}{n}}$
- ▶ Bestimmung des Intervalls $[\underline{\mu}, \bar{\mu}] = \bar{x}_n \pm t_{n-1;1-\alpha/2} \cdot \hat{s}e$
- ▶ Entscheidungsregel:
 - ★ H_0 wird nicht verworfen, wenn $\mu_0 \in [\underline{\mu}, \bar{\mu}]$
 - ★ H_0 zugunsten von H_1 verworfen, wenn $\mu_0 \notin [\underline{\mu}, \bar{\mu}]$

Zweiseitige Hypothesentest über Mittelwerte: Beispiel

Zufallsstichprobe

	n	Mittelwert	Std. Abw.	min	max
x (HH-Einkommen)	100	37,342	21,6387	2,820	102,703

Vorbereitung: Zufallsstichprobe mit $n = 100$, Verfahren daher zulässig; $\alpha = 0.05$

1. Hypothesen formulieren:

- ▶ $\mu_0 = 33,500$
- ▶ $H_0 : \mu = 33,500; H_1 : \mu \neq 33,500$

2. Entscheidungsregel:

- ▶ H_0 wird nicht verworfen, wenn $\mu_0 \in [\underline{\mu}, \bar{\mu}] = \bar{x}_n \pm t_{n-1; 1-\alpha/2} \cdot \hat{s}e$
- ▶ H_0 zugunsten von H_1 verworfen, wenn $\mu_0 \notin [\underline{\mu}, \bar{\mu}]$

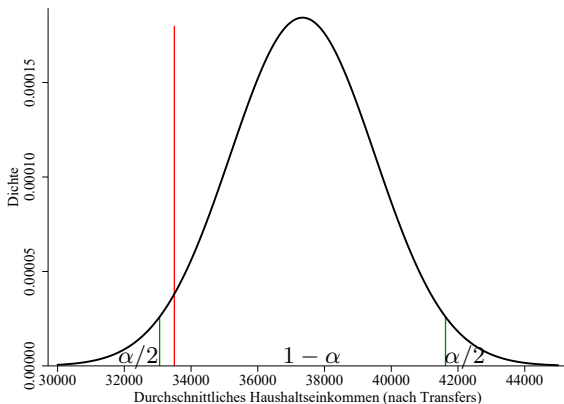
$$\bar{x}_n = 37,342; \hat{s}e = \sqrt{\frac{\hat{s}^2}{n}} = \sqrt{\frac{21,638^2}{100}} = 2,163.8; t_{99; 0.975} = 1.98$$

$$[\underline{\mu}, \bar{\mu}] = 37,342 \pm 1.98 \cdot 2,163.8 = [33,049; 41,636]$$

$$\mu_0 = 33,500 \in [33,049; 41,636] \rightarrow H_0 \text{ wird nicht verworfen.}$$

Zweiseitiger Hypothesentest: Grafische Darstellung

$$n = 100; \alpha = 0.05; \bar{x}_n = 37,342; \hat{s}e = 2,163.8; t_{n-1; 1-\alpha/2} = t_{99; 0.975} = 1.98$$



Anmerkung: t-Verteilung in Abbildung durch Normalverteilung approximiert.

$\mu_0 = 33,500$ liegt innerhalb des 95 %-Konfidenzintervalls
→ H_0 wird nicht verworfen.

Alternative Vorgangsweise: Standardisierte Teststatistik (1)

1. Hypothesen formulieren:

▶ $H_0 : \mu = \mu_0; H_1 : \mu \neq \mu_0$

2. Entscheidungsregel: mittels geschätztem Konfidenzintervall

▶ H_0 wird nicht verworfen, wenn $\mu_0 \in [\underline{\mu}, \bar{\mu}] = \bar{x}_n \pm t_{n-1; 1-\alpha/2} \cdot \hat{s}e$

mit $\hat{s}e = \sqrt{\frac{\hat{s}^2}{n}}$

▶ H_0 zugunsten von H_1 verworfen, wenn $\mu_0 \notin [\underline{\mu}, \bar{\mu}]$

Entscheidungsregel mittels standardisierter Teststatistik:

H_0 wird nicht verworfen, wenn:

$$\bar{x}_n - t_{n-1; 1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{s}^2}{n}} \leq \mu_0 \leq \bar{x}_n + t_{n-1; 1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{s}^2}{n}}$$

$$-t_{n-1; 1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{s}^2}{n}} \leq \mu_0 - \bar{x}_n \leq +t_{n-1; 1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{s}^2}{n}}$$

$$-t_{n-1; 1-\frac{\alpha}{2}} \leq \frac{\mu_0 - \bar{x}_n}{\sqrt{\frac{\hat{s}^2}{n}}} \leq +t_{n-1; 1-\frac{\alpha}{2}}$$

$$\text{t-Wert} = \left| \frac{\mu_0 - \bar{x}_n}{\sqrt{\frac{\hat{s}^2}{n}}} \right| \leq +t_{n-1; 1-\frac{\alpha}{2}}$$

Alternative Vorgangsweise: Standardisierte Teststatistik (2)

1. Hypothesen formulieren:

▶ $H_0 : \mu = \mu_0; H_1 : \mu \neq \mu_0$

2. Entscheidungsregel: mittels standardisierten Teststatistik

▶ H_0 wird nicht verworfen, wenn t-Wert = $\left| \frac{\mu_0 - \bar{x}_n}{\sqrt{\frac{\hat{s}^2}{n}}} \right| \leq t_{n-1; 1 - \frac{\alpha}{2}}$

▶ H_0 zugunsten von H_1 verworfen, wenn t-Wert = $\left| \frac{\mu_0 - \bar{x}_n}{\sqrt{\frac{\hat{s}^2}{n}}} \right| > t_{n-1; 1 - \frac{\alpha}{2}}$

▶ Vorgehensweise:

Berechnung einer standardisierten Teststatistik (t-Wert):

$$\text{t-Wert} = \left| \frac{\mu_0 - \bar{x}_n}{\sqrt{\frac{\hat{s}^2}{n}}} \right| = \left| \frac{33,500 - 37,342}{2,163.8} \right| = 1.78$$

Ermittlung des kritischen Werts c der Verteilung:

$$c = t_{n-1; 1 - \frac{\alpha}{2}} = t_{99; 0.975} = 1.98$$

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

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

Umsetzung in Excel

- $\alpha = 0.05$: wird festgelegt
- $n = 100$; **EXCEL-Befehl** ANZAHL
- $\bar{x}_n = 37,342$: **EXCEL-Befehl** MITTELWERT
- \hat{s} : **EXCEL-Befehl** STABW.S
- $\hat{s}e = \sqrt{\frac{\hat{s}^2}{n}}$
- $t_{n-1;1-\alpha/2} = t_{99;0.975}$: **EXCEL-Befehl** T.INV
 - ▶ T.INV(0.975; 99) gibt das 0.975-Perzentil wieder (für einen Stichprobenumfang $n = 100$ und daher $n - 1 = 99$ Freiheitsgraden).
 - ▶ alternativ dazu kann auch der Befehl T.INV.2S(0.05; 99) verwendet werden

Einseitiger Hypothesentest über Mittelwerte

Zufallsstichprobe

	n	Mittelwert	Std. Abw.	min	max
x (HH-Einkommen)	100	37,342	21,6387	2,820	102,703

Es handelt sich um eine Zufallsstichprobe, das Verfahren ist daher zulässig. Das Signifikanzniveau wird mit $\alpha = 0.05$ festgelegt.

1. Hypothesen formulieren:

▶ $H_0 : \mu \leq 33,500; H_1 : \mu > 33,500$

2. Entscheidungsregel: mittels geschätztem Konfidenzintervall

▶ H_0 wird nicht verworfen, wenn $\mu_0 \in [\underline{\mu}, \infty]$

$$\underline{\mu} = \bar{x}_n - t_{n-1;1-\alpha} \cdot \hat{s}e$$

▶ \bar{H}_0 zugunsten von H_1 verworfen, wenn $\mu_0 \notin [\underline{\mu}, \infty]$

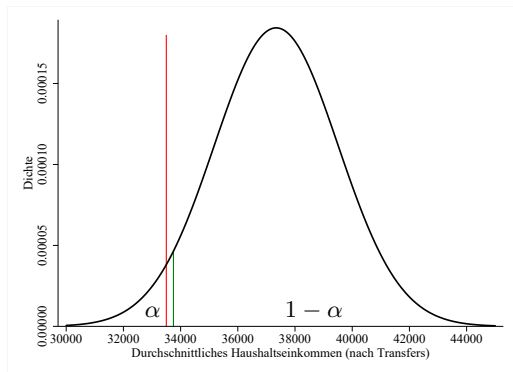
$$\bar{x}_n = 37,342; \hat{s}e = \sqrt{\frac{\hat{s}^2}{n}} = \sqrt{\frac{21,638^2}{100}} = 2,163.8; t_{99;95} = 1.66$$

$$[\underline{\mu}, \infty] = [37,342 - 1.66 \cdot 2,163.8, \infty] = [33,750, \infty]$$

$$\mu_0 = 33,500 \notin [33,750; \infty] \rightarrow H_0 \text{ wird verworfen.}$$

Einseitiger Hypothesentest: Grafische Darstellung

$$n = 100; \alpha = 0.05; \bar{x}_n = 37,342; \hat{s}e = 2,163.8; t_{n-1;1-\alpha} = t_{99;0.95} = 1.66$$



Anmerkung: t-Verteilung in Abbildung durch Normalverteilung approximiert.

$\mu_0 = 33,500$ liegt außerhalb des 95 %-Konfidenzintervalls $\rightarrow H_0$ wird verworfen. In der Grundgesamtheit ist das Durchschnittseinkommen **sehr wahrscheinlich** (mit zumindest 95 %-iger Wahrscheinlichkeit) **höher als 33,500 Euro**. Das Durchschnittseinkommen ist (statistisch) **signifikant** höher 33,550 Euro.

Einseitigen Hypothesentest: Standardisierte Teststatistik (1)

1. Hypothesen formulieren:

Fall 1: $H_0 : \mu \leq \mu_0; H_1 : \mu > \mu_0$

2. Entscheidungsregel: mittels Standardisierter Teststatistik

H_0 wird nicht verworfen wenn $\mu_0 \in [\underline{\mu}, \infty]$

$$\underline{\mu} = \bar{x}_n - t_{n-1;1-\alpha} \cdot \hat{s}e$$

$$\mu_0 \geq \underline{\mu}$$

$$\mu_0 \geq \bar{x}_n - t_{n-1;1-\alpha} \cdot \hat{s}e$$

$$\frac{\mu_0 - \bar{x}_n}{\hat{s}e} \geq -t_{n-1;1-\alpha}$$

Daraus folgt: H_0 wird verworfen, wenn

$$\frac{\mu_0 - \bar{x}_n}{\hat{s}e} < -t_{n-1;1-\alpha}$$

das ist erfüllt, wenn t-Wert = $\left| \frac{\mu_0 - \bar{x}_n}{\hat{s}e} \right| > t_{n-1;1-\alpha}$ **und** $\bar{x}_n > \mu_0$

Einseitigen Hypothesentest: Standardisierte Teststatistik (2)

- **Jedenfalls:**

- ▶ Berechnung einer standardisierten Teststatistik (t-Wert): $t\text{-Wert} = \left| \frac{\mu_0 - \bar{x}_n}{\sqrt{\frac{\hat{s}^2}{n}}} \right|$
- ▶ Ermittlung des kritischen Werts c der Verteilung:
 $c = t_{n-1;1-\alpha} = t_{99;0.95} = 1.66$

- **Fall 1:** $H_0 : \mu \leq \mu_0; H_1 : \mu > \mu_0$

- ▶ H_0 wird (zugunsten von H_1) verworfen, wenn
 $t\text{-Wert} = \left| \frac{\mu_0 - \bar{x}_n}{\sqrt{\frac{\hat{s}^2}{n}}} \right| \geq t_{n-1;1-\alpha}$ **und** $\bar{x}_n > \mu_0$

- **Fall 2:** $H_0 : \mu \geq \mu_0; H_1 : \mu < \mu_0$

- ▶ H_0 wird (zugunsten von H_1) verworfen, wenn
 $t\text{-Wert} = \left| \frac{\mu_0 - \bar{x}_n}{\sqrt{\frac{\hat{s}^2}{n}}} \right| \geq t_{n-1;1-\alpha}$ **und** $\bar{x}_n < \mu_0$