SE Geographie und Ökonomie

Einheit 3:
Univariate Datenanalyse: Schließende Statistik

Bernhard Schmidpeter

bernhard.schmidpeter@jku.at

Institut für Volkswirtschaftslehre

SoSe 2024

Motivation

- In der Praxis ist es oft nicht möglich, alle interessanten Objekte in der Grundgesamtheit zu untersuchen
- Vollerhebungen sind teuer und zeitaufwendig
- Oft steht nur eine Auswahl aus dieser Grundgesamtheit, die Stichprobe, zur Vefügung
- Das Interesse gilt aber normalerweise der Grundgesamtheit (Konsumerhebung, Befragungen zu Schattenwirtschaft, Wahlumfragen)

Motivation

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.

Beispiel: Wahlumfragen

- Interesse an der Grundgesamtheit: Wer gewinnt die nächste Wahl?
- Stichprobe: Bei einer Wahlumfrage unter 1,000 WählerInnen geben 23% der Befragten an, die Partei X wählen zu wollen.
- Frage: Können wir darauf schließ en, daß Partei X tatächlich 23% der Stimmen bekomment?

Lernziele der Einheit 3

- Sie kennen den Unterschied zwischen der **Verteilung einer Zufallsstichprobe** und der **Verteilung der Mittelwerte von Zufallsstichproben**.
- Sie können (ausgehend von einer Stichprobe) ein Intervall bestimmen, sodass der **Mittelwert einer Grundgesamtheit** mit einer gewissen Wahrscheinlichkeit **innerhalb des Intervalls** liegt
- 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: Häfigkeitsverteilung 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

Statistische Unsicherheit und die Rolle des Zufalls: Der Untergang der Titanic

Tabelle: Häfigkeitsverteilung 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 ab (Alter, Körpergröße, Geschlecht, Kabinenkategorie, ...)
- **Aber** auch vom Zufall ab
  - Nicht nur gute Schwimmer haben überlebt
  - Nicht nur junge Personen haben überlebt
- Unterstellen wir kurz, dass die Personeneigenschaften keine Rolle spielen, und jede Person eine Überlebenswahrscheinlichkeit von 33 % hat (fiktiv!)
  - Das Merkmal x (= 1 wenn "überlebt", und 0 sonst) ist eine Zufallsvariable
  - Der Erwartungswert is $\mu = 0.330$ und die Standardabweichung ist $\sigma = 0.470$ (bzw. eine Varianz von $\sigma^2 = 0.221$)
  - Diese zugrunde liegende Verteilung **beobachten wir nicht**

  Hinweis: Für eine binäre Variable x gilt: $E[x] = P(x = 1)$ und $Var(x) = P(x = 1) * [1 - P(x = 1)]$

Warum haben genaue 711 Personen überlebt? Zufall 1

- Wir beobachten lediglich die konkrete Realisationen eines Zufallsexperimentes
- Von 2,201 Passagieren haben 711 ($\overline{x} = 0.323$) überlebt. Die Standardabweichung $s = 0.219$
- Würden wir die Titanic nochmals losschicken und untergehen lassen, dann **würden vermutlich nicht exakt 711 Leute überleben**

$\Rightarrow$ Wie viel mehr oder weniger Personen in unserem Experiment überleben, hängt vom **Zufall** ab.

Warum haben genaue 711 Personen überlebt? Zufall 2

- Wenn wir eine Zufallsstichprobe ziehen, dann wird der Mittelwert $\overline{x}_{100}$ 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

- **Deskriptive Inferenz:** Auf Basis einer Stichprobe können Rückschlüsse auf die Grundgesamtheit gezogen werden.

  Beispiel: Wenn in der Stichprobe der Anteil an Überlebenden bei Gästen der 1. Klasse größer ist als bei Gästen der 3. Klasse ($\overline{x}_{100;1.Klasse} = 0.60 > \overline{x}_{100;3.Klasse} = 0.27$), kann daraus geschlossen werden, dass auch in der Grundgesamtheit der Anteil der Überlebenden bei Gästen der 1. Klasse größer ist?

⇒ Unser Ziel heute: **Rückschlüsse von einer Stichprobe auf die Grundgesamtheit zu ziehen** (deskriptive Inferenz).

- Beobachten wir $\overline{x}_{100;1.Klasse} > \overline{x}_{100;3.Klasse}$ nur aufgrund der Eigenart unserer Stichprobe?
- Oder existiert unsere Beobachtung tatsächlich in der Grundgesamtheit?

Schließende Statistik (2)

- **Kausale Inferenz:** Untersuchung ursächliche Mechanismen hinter beobachteten Zusammenhängen.

  Beispiel: Ist der größere Anteil an Überlebenden bei Gästen der 1. Klasse, der in der Grundgesamtheit beobachtet wird ($\overline{x}_{1.Klasse} = 0.62 > \overline{x}_{3.Klasse} = 0.25$), auf eine höhere Überlebenswahrscheinlichkeit zurückzuführen (d.h. $\mu_{1.Klasse} > \mu_{3.Klasse}$), oder kann der Unterschied lediglich durch Zufall erklärt werden?
- Wir werden uns damit später im Kurs beschäftigen

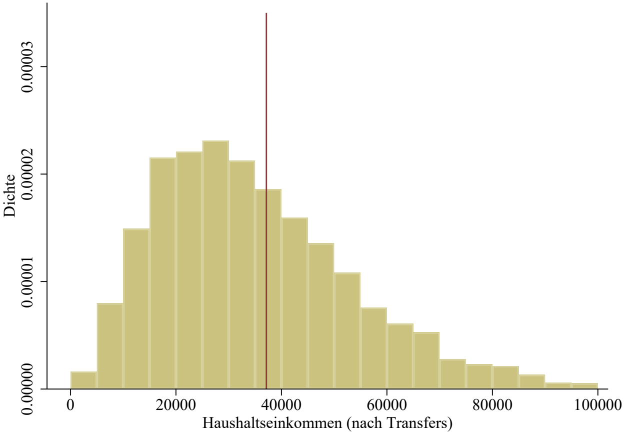
Grundgesamtheit vs. Stichprobe

- Die deskriptive Statistik beschäftigt sich mit der Verteilung einzelner Merkmale
- 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**, praxis-relevant).
- Zwei Aspekte sind von Interesse
  - i) Wenn wir die **Grundgesamtheit** kennen, können wir daraus etwas über die **Stichprobe** lernen (geringe Relevanz)
  - 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.

Wichtig: 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.

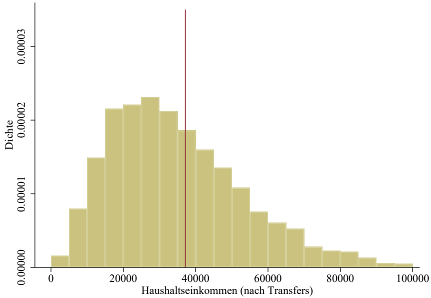
“Grundgesamtheit”

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)



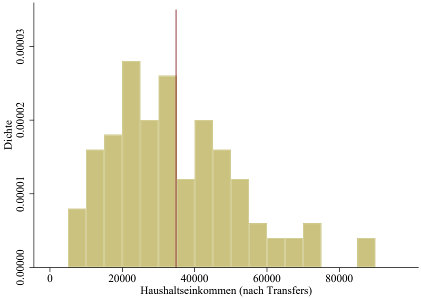
Grundgesamtheit vs. Stichprobe: Beispiel (1)

Histogramm Grundgesamtheit



$(N = 5{,}407;\ \mu = 37{,}150;\ \sigma = 26{,}728)$

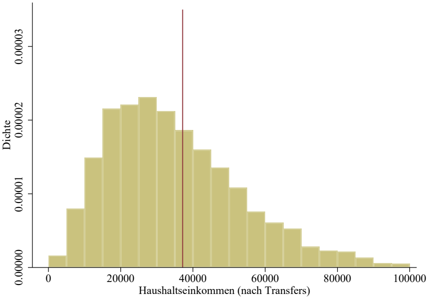
Histogramm Stichprobe



$(n = 100;\ \overline{x}'_{100} = 34{,}864;\ \hat{s}'_{100} = 17{,}637)$

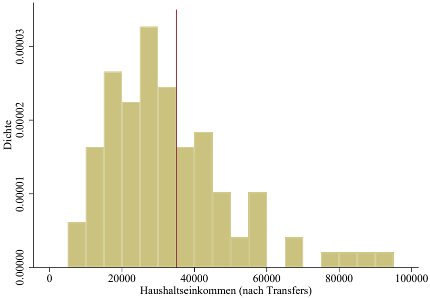
Grundgesamtheit vs. Stichprobe: Beispiel (2)

Histogramm Grundgesamtheit



$(N = 5{,}407;\ \mu = 37{,}150;\ \sigma = 26{,}728)$

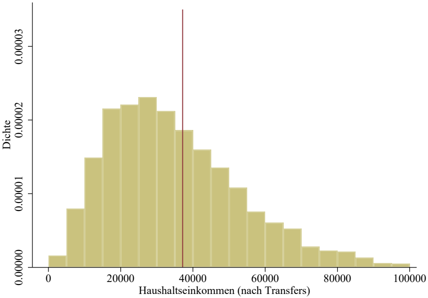
Histogramm Stichprobe



$(n = 100;\ \overline{x}''_{100} = 35{,}019;\ \hat{s}'_{100} = 24{,}460)$

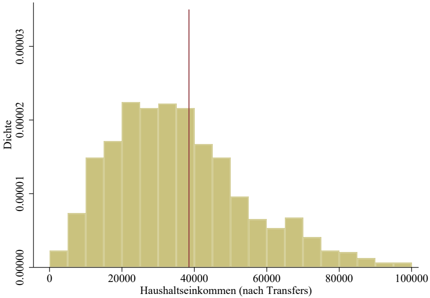
Grundgesamtheit vs. Stichprobe: Beispiel (3)

Histogramm Grundgesamtheit



$(N = 5{,}407;\ \mu = 37{,}150;\ \sigma = 26{,}728)$

Histogramm Stichprobe



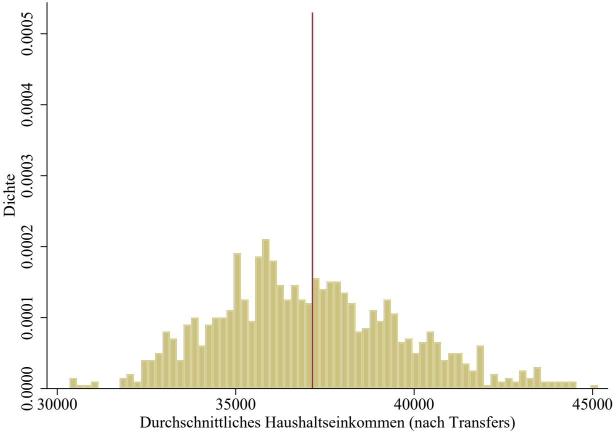
$(n = 1{,}000;\ \overline{x}'_{1000} = 38{,}509;\ \hat{s}'_{1000} = 32{,}111)$

Erkenntnisse: Intuition $1 + 2$

- Erkenntnis 1: Wenn ich Zufallsstichproben ziehe, sieht die Verteilung dieser Stichproben 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)

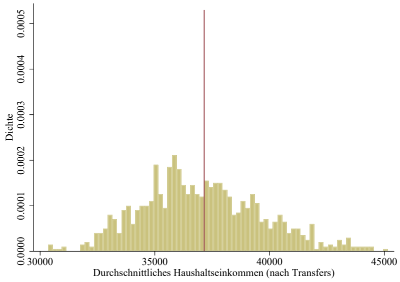
Ich ziehe eine Zufallsstichprobe von $n = 100$ Personen und berechne den Mittelwert der Stichprobe, $\overline{x}_{100}$. Ich wiederhole den Vorgang 1,000 Mal und erhalte daher 1,000 verschiedene Stichprobenmittelwerte (also $\overline{x}'_{100}$, $\overline{x}''_{100}$, $\overline{x}'''_{100}$, usw.).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

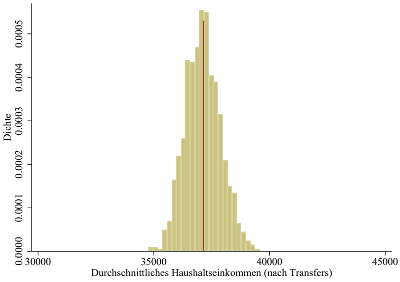


(vertikale rote Gerade markiert den "wahren" Mittelwert $\mu = 37,150$)

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	$507,369$
$\overline{x}$ ($n = 100$, links)	$1,000$	$37,060$	$2,677$	$30,053$	$47,483$
$\overline{x}$ ($n = 1,000$, rechts)	$1,000$	$37,125$	755	$34,573$	$39,920$

Erkenntnisse: Intuition 3 + 4

- Erkenntnis 3: Wenn ich mehrmals eine Stichprobe ziehe und deren Mittelwert berechne, so streuen die Stichprobenmittelwerte um den Mittelwert der Grundgesamtheit.
  - Denken Sie: Ich führe das gleiche Experiment mehrmals durch
- 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:

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

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

  ii) die Varianz der Stichprobe:

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

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

Erkenntnisse: Formal

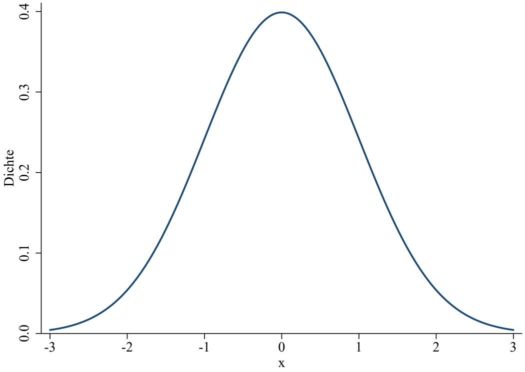
- $\overline{x}_n$ ist selbst eine Zufallsvariable, die einer **Normalverteilung** mit Mittelwert μ und Varianz σ^2/n folgt (**zentraler Grenzwertsatz**): $\overline{x}_n \sim N\left(\mu, \sigma^2/n\right)$
- Die Standardabweichung der Verteilung der Mittelwerte wird auch **Standardfehler** (des Mittelwertes) genannt, und mit *se* (*standard error*) bezeichnet.
- Weil $\overline{x}_n$ den Mittelwert μ in Erwartung abbildet, kann der tatächlich beobachtetet $\overline{x}_n$ stark davon abweichen $\rightarrow$ die Standardabweichung sollte immer mit angegeben werden
  - Frauen verdienen $2,000$ Euro weniger als Männer ($n = 10$, $\sigma = 10,000$)
  - Frauen verdienen $2,000$ Euro weniger als Männer ($n = 50{,}000$, $\sigma = 1,000$)

Die 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

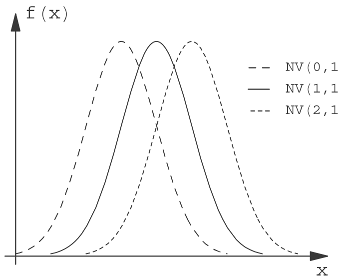
$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{\left(-\frac{(x-\mu)^2}{1\sigma^2}\right)} \quad \text{mit } \sigma \geq 0$$

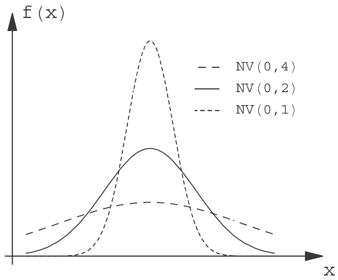
besitzt. (Brauchen Sie nicht wissen.) Eine Normalverteilung $N(\mu = 0, \sigma^2 = 1)$ heißt **Standardnormalverteilung** und sieht so aus:



Die Normalverteilung (2)

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





(siehe Duller, 2019, Abbildung 11.5)

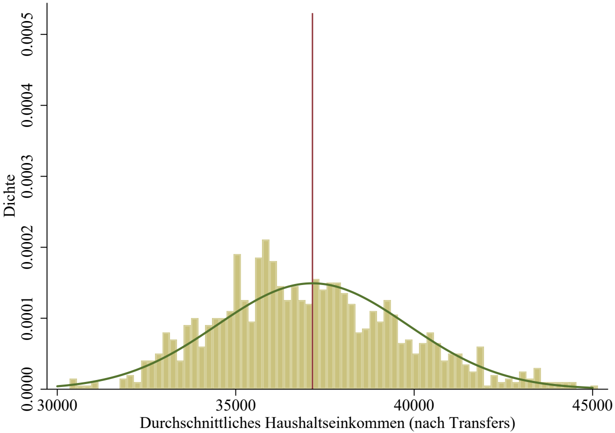
Die Normalverteilung (3)

Warum interessiert uns die Normalverteilung?

- Die (Standard-) Normalverteilung erlaubt, die Verteilung von $\overline{x}_n$ zu normieren
- Die (Standard-) Normalverteilung erlaubt es uns, Wahrscheinlichkeiten der form $P(x \leq X)$ zu bestimmen
- Die (Standard-) Normalverteilung erlaubt es uns, für ein p, einen Wert c zu bestimmen, so dass $P(x \leq c) = p$
- Die (Standard-) Normalverteilung erlaubt es uns, zu entscheiden, ob $\overline{x}_n$ weit von μ entfernt ist

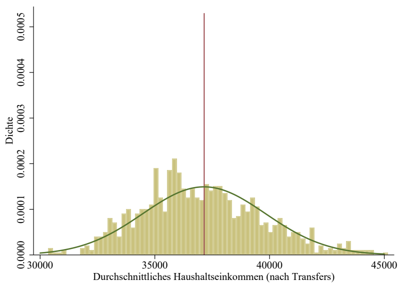
Häufigkeitsverteilung der Mittelwerte (1)

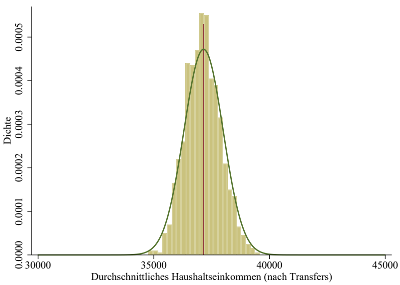
Kann durch Normalverteilung mit $\mu = 37,150$ und Standardabweichung $\sigma/\sqrt{n} = 26,728/\sqrt{100} = 2672.8$ approximiert werden (Grundgesamtheit SOEP):



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:





Häufigkeitsverteilung der Mittelwerte (3)

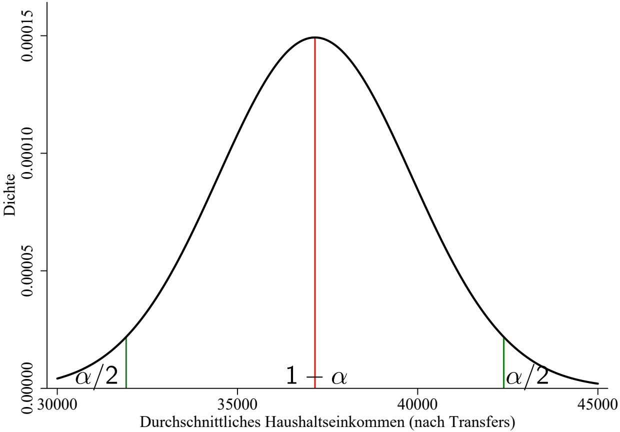
	N	Mittelwert	Std. Abw.	min	max
x (HH-Einkommen)	$5,407$	$37,150$	$26,728$	583	$507,369$
$\overline{x}$ $(n = 100)$	$1,000$	$37,060$	$2,677$	$30,053$	$47,483$
$\overline{x}$ $(n = 1,000)$	$1,000$	$37,125$	755	$34,573$	$39,920$

- Normalerweise beobachten wir x nicht
- Wie wissen wir dann, ob $\overline{x}$ die Wirklichkeit μ gut abbildet?
- Können wir hierzu die Tatsache verwenden, dass $\overline{x}$ normalverteilt ist?

$\Rightarrow$ Bestimmung eines Intervall, sodass $\overline{x}_n$ mit einer bestimmten Wahrscheinlichkeit $1 - \alpha$ innerhalb des Intervalls (bzw. mit einer Wahrscheinlichkeit α außerhalb) liegt

Intervall: Intuition

Verteilung der Stichprobenmittelwerte eines Merkmals x mit dem Stichprobenumfang $n = 100$, $\bar{x}_{100} = 37,060$ und Standardfehler $se = \sigma/\sqrt{n} = 2,677/\sqrt{100} \approx 2,673$ für $\alpha = 0.05$ (5 %):



Intervall: Intuition

- Von der Graphik können wir sehen, dass

$$P(-\tilde{u}_{\frac{\alpha}{2}} \leq \overline{x} \leq \tilde{u}_{\frac{\alpha}{2}}) = 1 - \alpha$$

- Weil intuitive, das Intervall ist schwer zu bestimmen
  - Intervallgrenzen hängen von der prżisen Fragestellung ab
  - Vergleichen Sie die Verteilung von Einkommen und Vermögen

- Um uns das Leben zu erleichtern, verwenden wir einen "Trick":

$$\text{Wenn} \quad \bar{x} \sim N(\mu, \sigma^2/n) \quad \text{dann} \quad \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

- Die Standardnormalverteilung hat einfach zu bestimmende Intervallgrenzen
  - Für $\alpha = 10\%$, $u_{.05} = 1.644854$
  - Für $\alpha = 5\%$, $u_{.025} = 1.959964$
  - Für $\alpha = 1\%$, $u_{.025} = 2.575829$

Intervall: Intuition (2)

- Unter Verwendung unseres "Tricks"haben wir

$$P(-u_{\frac{\alpha}{2}} \leq \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \leq u_{\frac{\alpha}{2}}) = 1-\alpha$$

- Und deshalb gilt:

$$\begin{aligned}
&P(-u_{\frac{\alpha}{2}} \leq \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \leq u_{\frac{\alpha}{2}}) \\
&= P\left(-u_{\frac{\alpha}{2}}\left(\sigma/\sqrt{n}\right) \leq \bar{x}-\mu \leq u_{\frac{\alpha}{2}}\left(\sigma/\sqrt{n}\right)\right) \\
&= P\left(-u_{\frac{\alpha}{2}}\left(\sigma/\sqrt{n}\right) - \bar{x} \leq -\mu \leq u_{\frac{\alpha}{2}}\left(\sigma/\sqrt{n}\right) - \bar{x}\right) \\
&= P\left(\bar{x} - u_{\frac{\alpha}{2}}\left(\sigma/\sqrt{n}\right) \leq -\mu \leq \bar{x} + u_{\frac{\alpha}{2}}\left(\sigma/\sqrt{n}\right)\right)
\end{aligned}$$

- Ersetzen σ durch $\hat{s}$ gibt uns unser Konfidenzintervall

$$\left[\bar{x} - u_{\frac{\alpha}{2}}\left(\hat{s}/\sqrt{n}\right), \bar{x} + u_{\frac{\alpha}{2}}\left(\hat{s}/\sqrt{n}\right)\right)$$

Hinweise

- 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**)

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

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 Konfedenzintervalls:

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

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

Das Intervall kann auch folgendermaßen angeschrieben werden:

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

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

Statistisches Testen: Intuition

- Wir können unsere Intuition über Konfidenzintervalle verwenden um sogenannte **statistische Tests** durchzuführen.
- Ein statistischer Test ist eine Regel zur Entscheidung unter Unsicherheit, da wir die Grundgesamtheit nicht kennen
- Die Entscheidung ist zwischen zwei Behauptungen zu treffen, die als Hypothesen bezeichnet werden
  - Frauen verdienen unterschiedlich viel als Männer (zweiseitiger Test)
  - 3. Klasse Passagiere hatten ein geringere Überlebenswahrscheinlichkeit als 1. Klasse Passagiere (einseitiger Test)

Statistisches Testen: Intuition

- **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

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

		Entscheidung auf H_0	Entscheidung auf 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 ($H1$ ist immer Hypothese von Intresse)
- Wenn die Nullhypothese nicht abgelehnt wird, bedeutet das nicht, dass die Nullhypothese zutrifft.(z.B.Mangel an Beweisen)

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 schwerwiegenderen 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 $\overline{x}_n$ signifikant größer als μ_0 ist **oder** wenn $\overline{x}_n$ 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 $[\underline{\mu}, \overline{\mu}] = \overline{x}_n \pm t_{n-1;1-\alpha/2} \cdot \hat{se}$ mit $\hat{se} = \sqrt{\frac{\hat{s}^2}{n}}$
   - Bestimmung des Intervalls $[\underline{\mu}, \overline{\mu}] = \overline{x}_n \pm t_{n-1;1-\alpha/2} \cdot \hat{se}$
   - Entscheidungsregel:
     - H_0 wird nicht verworfen, wenn $\mu_0 \in [\underline{\mu}, \overline{\mu}]$
     - H_0 zugunsten von H_1 verworfen, wenn $\mu_0 \notin [\underline{\mu}, \overline{\mu}]$

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

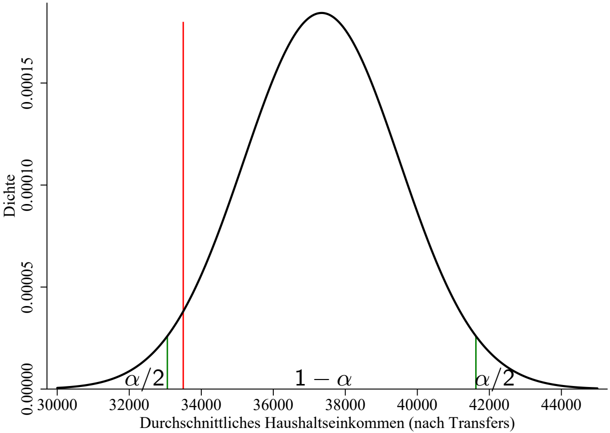
$\overline{x}_n = 37,342$; $\hat{se} = \sqrt{\frac{\hat{s}^2}{n}} = \sqrt{\frac{21,638^2}{100}} = 2,163.8$; $t_{99;0.975} = 1.98$

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

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

Zweiseitiger Hypothesentest: Grafische Darstellung

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



Alternative Vorgangsweise: Standardisierte Teststatistik (1)

Entscheidungsregel mittels standardisierten Teststatistik:

H_0 wird nicht verworfen, wenn:

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

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

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

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

Alternative Vorgangsweise: Standardisierte Teststatistik (2)

Entscheidungsregel mittels standardisierten Teststatistik:

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

  Berechnung einer standardisierten Teststatistik (t-Wert):

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

  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
- $\overline{x}_n = 37,342$: **EXCEL-Befehl** MITTELWERT
- $\hat{s}$: **EXCEL-Befehl** STABW.S
- $\hat{se} = \sqrt{\frac{\hat{s}^2}{n}}$
- $t_{n-1;1-\alpha/2} = t_{99;0.975}$: **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 $\mu_0 \in [\underline{\mu}, \infty]$
     $\underline{\mu} = \overline{x}_n - t_{n-1;1-\alpha} \cdot \hat{se}$
   - H_0 zugunsten von H_1 verworfen, wenn $\mu_0 \notin [\underline{\mu}, \infty]$

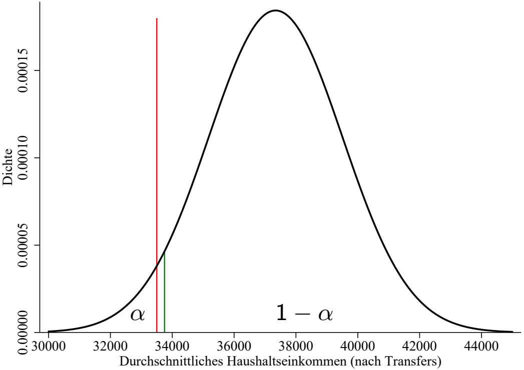
$$\overline{x}_n = 37,342; \ \hat{se} = \sqrt{\frac{\hat{s}^2}{n}} = \sqrt{\frac{21,638^2}{100}} = 2,163.8; \ t_{99;95} = 1.66$$

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

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

Einseitiger Hypothesentest: Grafische Darstellung

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



Anmerkung: t-Verteilung in Abbildung durch Normalverteilung approximiert.

$\mu_0 = 33,500$ liegt außerhalb des 95 %-Konfidenzintervalls $\longrightarrow$ 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 $\mu_0 \in [\underline{\mu}, \infty]$
   $\underline{\mu} = \overline{x}_n - t_{n-1;1-\alpha} \cdot \hat{se}$

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

$$\mu_0 \geq \overline{x}_n - t_{n-1;1-\alpha} \cdot \hat{se}$$

$$\frac{\mu_0 - \overline{x}_n}{\hat{se}} \geq -t_{n-1;1-\alpha}$$

Daraus folgt: H_0 wird verworfen, wenn

$$\frac{\mu_0 - \overline{x}_n}{\hat{se}} < -t_{n-1;1-\alpha}$$

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

Einseitigen Hypothesentest: Standardisierte Teststatistik (2)

- **Jedenfalls:**
  - Berechnung einer standardisierten Teststatistik (t-Wert):
    t-Wert $= \left| \frac{\mu_0 - \overline{x}_n}{\sqrt{\frac{\hat{s}^2}{n}}} \right|$
  - 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-Wert $= \left| \frac{\mu_0 - \overline{x}_n}{\sqrt{\frac{\hat{s}^2}{n}}} \right| \geq t_{n-1;1-\alpha}$ **und** $\overline{x}_n > \mu_0$
- Fall 2: $H_0: \ \mu \geq \mu_0$; $H_1: \ \mu < \mu_0$
  - H_0 wird (zugunsten von H_1) verworfen, wenn
    t-Wert $= \left| \frac{\mu_0 - \overline{x}_n}{\sqrt{\frac{\hat{s}^2}{n}}} \right| \geq t_{n-1;1-\alpha}$ **und** $\overline{x}_n < \mu_0$