ST. PÖLTEN UNIVERSITY
OF APPLIED SCIENCES

Informatik
& Security

Schätzverfahren

Angewandte Statistik

Dipl.-Ing. Hubert Schölnast, BSc

Stand: 3. Juli 2023

Inhaltsverzeichnis

1 Schätzfunktionen 3

1.1 Schätzer 3
1.1.1 Punkt-Schätzer und Intervallschätzer 3
1.2 Güte von Schätzfunktionen 4
1.2.1 Unverzerrtheit (Erwartungstreue) 5
1.2.2 Effizienz 5
1.2.3 Konsistenz 5
1.3 Konstruktion einer Schätzfunktion 6
1.3.1 Methode der kleinsten Quadrate 6
1.3.2 Maximum-Likelihood-Methode 6

2 Punktschätzung 7

2.1 Typische Punktschätzer 7
2.1.1 Schätzer für das arithmetische Mittel 7
2.1.2 Schätzer für die Varianz 8
2.1.3 Schätzer für einen Anteilswert 9
2.2 Wichtige Eigenschaften von Punktschätzungen 9
2.2.1 Unmöglichkeit die Qualität der Schätzung abzuschätzen. 9
2.2.2 Voraussetzung für Intervallschätzung 10

3 Intervallschätzung 10

3.1 Erstellen eines Konfidenzintervall (Grundidee) 11
3.1.1 Inklusionsschluss 11
3.1.2 Wahrscheinlichkeit für ein bestimmtes Intervall 20
3.1.3 Intervall mit einer bestimmten Wahrscheinlichkeit 23
3.2 Konfidenzintervall für das arithmetische Mittel 25
3.2.1 Konfidenzintervall für das arithmetische Mittel, wenn σ bekannt ist 25
3.2.2 Konfidenzintervall für das arithmetische Mittel, wenn σ nicht bekannt ist 26
3.3 Ziehen mit und ohne Zurücklegen 27
3.4 Konfidenzintervall für den Abstand zweier Grundgesamtheiten 28
3.4.1 Wenn die Varianzen beider Grundgesamtheiten bekannt sind 29
3.4.2 Wenn die Varianzen beider Grundgesamtheiten gleich, aber unbekannt sind 29

1 Schätzfunktionen

1.1 Schätzer

Die Begriffe Schätzer, Schätzstatistik und Schätzfunktion bezeichnen alle dieselbe Sache: Eine Regel (eine Folge von Anweisungen, eine Funktion, ...), mit deren Hilfe ein Schätzwert für eine bestimmte statistische Kennzahl (meist ein Lage- oder Streumaß) aus den Daten einer Stichprobe ermittelt werden kann. Dabei soll dieser Schätzwert so gut wie möglich mit der entsprechenden Kennzahl der Grundgesamtheit übereinstimmen. Mit einem Schätzer wird also versucht, anhand einer Stichprobe eine Kennzahl der Grundgesamtheit abzuschätzen.

In der englischsprachigen Literatur findet man diese Begriffe:

- estimator Das ist die Schätzfunktion, so wie sie oben beschrieben ist
- estimand Das ist die Kennzahl, die geschätzt werden soll
- estimate Das ist das Ergebnis der Schätzung (der Wert der Kennzahl)

In der deutschsprachigen Literatur wird leider nicht immer so exakt zwischen diesen Begriffen unterschieden, hier kann mit "Schätzer" je nach Kontext jedes dieser drei Dinge gemeint sein.

Ein einfaches Beispiel:

Wenn man das arithmetische Mittel (AM) einer Grundgesamtheit abschätzen soll, und aus dieser Grundgesamtheit nur eine kleine Stichprobe zur Verfügung hat, berechnet man meist das AM der Stichprobe und verwendet diesen Wert als Schätzwert für das AM der Grundgesamtheit. Dann bezeichnet man das AM der Stichprobe als Schätzer für das AM der Grundgesamtheit.

1.1.1 Punkt-Schätzer und Intervallschätzer

Es gibt Punkt- und Intervallschätzer. Die Punktschätzer liefern einwertige Ergebnisse. Dies steht im Gegensatz zu einem Intervallschätzer, bei dem das Ergebnis ein Bereich von plausiblen Werten ist. »Einwertig« bedeutet nicht unbedingt »eine einzige Zahl«, sondern umfasst auch vektor- oder funktionsbewertete Schätzer, in den meisten Fällen (so auch in diesem Skriptum) ist es aber trotzdem genau eine einzige reelle Zahl.

Alle Funktionen aus dem 1. Semester, die sich ausdrücklich auf Stichproben bezogen haben (z.B. Standardabweichung einer Stichprobe) sind bereits Schätzfunktionen. Darüber hinaus gibt es noch weitere, die hier vorgestellt werden.

1.2 Güte von Schätzfunktionen

Der Wert, der das Ergebnis einer Schätzung ist, wird für gewöhnlich von dem entsprechenden Parameter der Grundgesamtheit abweichen. Im Folgenden wird das Symbol T als Name für eine Kennzahl der Grundgesamtheit verwendet, ohne dass wir uns darauf festlegen wollen, welche Kennzahl das genau ist. Der entsprechende Schätzwert, den wir aus der Stichprobe ermitteln, soll $\hat{T}$ (T-Dach) heißen. Dann ist

$$\mathrm{E}\left(\left(\hat{T} - T\right)^2\right)$$

der Erwartungswert für die quadrierte Abweichung des Schätzwertes $\hat{T}$ vom wahren Wert T. Von diesem Wert möchte man, dass er möglichst klein ist. Mit Hilfe des Steinerschen Verschiebungssatzes[1] kann das zu diesem Ausdruck umgeformt werden:

$$\mathrm{E}\left(\left(\hat{T} - T\right)^2\right) \quad = \quad \mathrm{VAR}\left(\hat{T}\right) + \left(\mathrm{E}\left(\hat{T}\right) - T\right)^2$$

Dabei ist $\mathrm{VAR}\left(\hat{T}\right)$ die Varianz des Schätzwertes und der Ausdruck $\left(\mathrm{E}\left(\hat{T}\right) - T\right)^2$ heißt Verzerrung (engl.: *bias*). Keiner der beiden Ausdrücke kann negativ sein.

Die Werte von $\mathrm{VAR}\left(\hat{T}\right)$ (Varianz des Schätzwertes) und $\mathrm{E}\left(\hat{T}\right)$ (Erwartungswert des Schätzwertes) kann man experimentell erhalten, indem man mehrere gleich große und voneinander unabhängige Stichproben aus der Grundgesamtheit zieht, aus jeder einzelnen Stichprobe den Schätzwert berechnet, und dann aus diesen Schätzwerten das arithmetische Mittel berechnet. Dieser Mittelwert der Schätzwerte nähert sich mit zunehmender Anzahl von Stichproben immer näher dem Erwartungswert des Schätzwertes an (Gesetz der großen Zahlen). Sobald man diesen Erwartungswert kennt, kann man dann auch die Varianz dieser Schätzwerte berechnen.

Tatsächlich wird man aber nicht hundert- oder tausendmal Stichproben aus derselben Grundgesamtheit zeihen, nur um $\mathrm{VAR}\left(\hat{T}\right)$ und $\mathrm{E}\left(\hat{T}\right)$ zu ermitteln. Das ist auch gar nicht notwendig. Es ist aber wichtig zu wissen, dass diese Werte von der Art und Weise abhängen, wie man den jeweiligen Schätzwert berechnet. Unterschiedliche Schätzmethoden werden unterschiedliche Schätzwerte liefern, die dann zu unterschiedlichen Werten für $\mathrm{E}\left(\hat{T}\right)$ und $\mathrm{VAR}\left(\hat{T}\right)$ führen. Wenn man mehrere Schätzmethoden zur Auswahl hat, um einen Schätzwert zu ermitteln, dann kann man von jeder Methode $\mathrm{VAR}\left(\hat{T}\right)$ und $\left(\mathrm{E}\left(\hat{T}\right) - T\right)^2$ berechnen und dann jene Methode wählen, bei der diese Werte am kleinsten sind.

[1] Siehe z.B. https://de.wikipedia.org/wiki/Verschiebungssatz_(Statistik)

Dabei gibt es ein paar Kriterien, die sinnvollerweise erfüllt sein sollten:

1.2.1 Unverzerrtheit (Erwartungstreue)

Die Verzerrung $\left(\mathrm{E}(\hat{T}) - T\right)^2$ kann bestenfalls den Wert 0 annehmen. Dann ist

$$\mathrm{E}(\hat{T}) = T$$

Dann entspricht also der Erwartungswert der Schätzung (das arithmetische Mittel vieler Schätzungen) genau dem zu schätzenden Wert der Grundgesamtheit. Eine erwartungstreue bzw. nichtverzerrende Schätzfunktion hat also die Eigenschaft, dass ihre Schätzergebnisse langfristig genau um den wahren geschätzten Wert streuen.

1.2.2 Effizienz

Wenn zwei erwartungstreue Schätzfunktionen a und b bei derselben Stichprobe die Schätzwerte $\hat{T}_a$ und $\hat{T}_b$ mit den Varianzen $\mathrm{VAR}(\hat{T}_a)$ und $\mathrm{VAR}(\hat{T}_b)$ hervorbringen, ist jene Schätzfunktion zu bevorzugen, bei der die Schätzwerte weniger stark um den Erwartungswert streuen. Es ist also die Funktion zu bevorzugen, die die kleinere Varianz $\mathrm{VAR}(\hat{T})$ aufweist.

Mittlerer quadratischer Fehler

Manchmal wird auch ein möglichst kleiner mittlerer quadratischer Fehler $\left(\hat{T} - T\right)^2$ als Kriterium einer Schätzfunktion genannt. Allerdings stimmt der mittlere quadratische Fehler bei erwartungstreuen Schätzern genau mit der Varianz überein, und bei verzerrenden Schätzern ist die Differenz zwischen dem mittleren quadratischen Fehler und der Varianz genau die Verzerrung. Daher bewirkt eine Verbesserung des mittleren quadratischen Fehlers bei gleichbleibender Verzerrung immer auch eine Verbesserung der Effizienz und umgekehrt.

1.2.3 Konsistenz

Je größer die Stichprobe ist, die man auswertet, desto mehr Informationen über die Grundgesamtheit stehen zur Verfügung. Daher sollte eine Schätzung auch umso genauer sein, je größer der Stichprobenumfang ist. Das heißt, dass die Varianz des Schätzwertes mit zunehmender Stichprobengröße immer kleiner werden sollte. Wenn n_1 und n_2 die Größen zweier Stichproben sind und $\hat{T}_1$ und $\hat{T}_2$ die dazugehörenden Schätzwerte, muss also gelten:

$$n_1 < n_2 \quad \Rightarrow \quad \mathrm{VAR}(\hat{T}_1) > \mathrm{VAR}(\hat{T}_2)$$

Diese Form der Konsistenz wird *schwache Konsistenz* genannt.

Daneben gibt es die *starke Konsistenz*, die sich dadurch auszeichnet, dass zusätzlich auch noch gefordert wird, dass die Varianz des Schätzwertes gegen 0 strebt, wenn der Stichprobenumfang über alle Schranken wächst. Bei der schwachen Konsistenz kann es auch sein, dass die Varianz auch bei (theoretisch) unendlich großen Stichproben größer als 0 ist.

1.3 Konstruktion einer Schätzfunktion

1.3.1 Methode der kleinsten Quadrate

Bei dieser Methode wird versucht, einen mathematischen Ausdruck für die Schätzfunktion $\hat{T} = f(x_i)$ zu finden, bei dem der Erwartungswert $\mathrm{E}\left(\left(\hat{T} - T\right)^2\right)$ ein Minimum hat. Dazu setzt man in die Formel für den Erwartungswert die Formel für die Schätzfunktion ein, differenziert das Ganze einmal, setzt diese erste Ableitung gleich 0 und wandelt diese Gleichung dann in einen brauchbaren Ausdruck um.

Diese Methode liefert immer einen erwartungstreuen und konsistenten Schätzer, allerdings gibt es keine Garantie dafür, dadurch auch einen effizienten Schätzer zu erhalten.

1.3.2 Maximum-Likelihood-Methode

Dabei überlegt man sich, wie eine Grundgesamtheit beschaffen sein müsste, damit sie mit größter Wahrscheinlichkeit genau die Stichprobe hervorbringt, die man erhalten hat. Und dann überlegt man sich, welchen Wert T in dieser Grundgesamtheit haben muss. Das Ergebnis ist natürlich wieder eine Formel, die die Messwerte der Stichprobe verarbeitet um einen Schätzwert für T zu liefern, aber das kann bei manchen Schätzgrößen eine andere Formel sein als die, die man durch die Methode der kleinsten Quadrate erhalten würde.

2 Punktschätzung

Bei der Punktschätzung ordnet man einer Stichprobe einen einzelnen Wert zu und hofft, mit diesem Schätzwert einen plausiblen Wert für die entsprechende Größe der Grundgesamtheit zu erhalten.

2.1 Typische Punktschätzer

Im Folgenden werden ohne explizite Herleitung einige typische und häufig verwendete Punktschätzer vorgestellt. Dabei werden im Kapitel 2 diese Bezeichnungen für die Variablen verwendet:

Beschreibung	Grundgesamtheit (Großbuchstaben)	Stichprobe (Kleinbuchstaben)
Anzahl der Elemente	N	n
Einzelwerte	X_i	x_i
arithmetisches Mittel	$\bar{X}$	$\bar{x}$

2.1.1 Schätzer für das arithmetische Mittel

Zu schätzende Größe: Das arithmetische Mittel der Grundgesamtheit

$$\bar{X} = \frac{1}{N} \cdot \sum_{i=i}^{N} X_i$$

Schätzwert: Das arithmetische Mittel der Stichprobe

$$\bar{x} = \frac{1}{n} \cdot \sum_{i=i}^{n} x_i$$

Dieser Schätzwert ist erwartungstreu und stark konsistent.

2.1.2 Schätzer für die Varianz

Zu schätzende Größe: Die Varianz der Grundgesamtheit. (Der Wert $\bar{X}$ muss zuvor mit der Formel aus 0 berechnet werden.)

$$\sigma^2 = \frac{1}{N} \cdot \sum_{i=i}^{N} (X_i - \bar{X})^2$$

- Schätzwert unter der Voraussetzung, dass $\bar{X}$, also das arithmetische Mittel der Grundgesamtheit, bekannt ist.

$$\sigma^{*2} = \frac{1}{n} \cdot \sum_{i=i}^{n} (x_i - \bar{X})^2$$

- Schätzwerte wenn $\bar{X}$ unbekannt ist. Vor der Berechnung des Schätzwertes für die Varianz muss $\bar{x}$ (Schätzwert für das arithmetische Mittel) berechnet werden. Es gibt zwei unterschiedliche Schätzer für diesem Fall:
  - **unkorrigierte Stichprobenvarianz**

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

Dieser Schätzer ist nicht erwartungstreu. Der Erwartungswert des Schätzwertes ist immer kleiner als die Varianz der Grundgesamtheit. Daher ist dieser Schätzwert auch nicht stark konsistent (aber dennoch schwach konsistent). Der Vorteil dieser Schätzfunktion ist, dass die Schätzwerte gerade bei kleinen Stichprobenumfängen weniger stark von einer Normalverteilung abweichen als bei der korrigierten Stichprobenvarianz. Diese Schätzfunktion erhält man, wenn man mit Hilfe der Maximum-Likelihood-Methode eine Schätzfunktion konstruiert.

  - **korrigierte Stichprobenvarianz**

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

Dieser Schätzer ist erwartungstreu und stark konsistent. Man erhält diese Schätzfunktion durch die Methode der kleinsten Quadrate. Aber die Verteilung der Ergebnisse, die man für mehrere Stichproben aus derselben Grundgesamtheit erhält, weicht von einer Normalverteilung ab. Diese Abweichung ist umso stärker ausgeprägt, je kleiner der Umfang der Stichprobe ist. Weil in der Praxis meist der Erwartungstreue und der starken Konsistenz der Vorzug gegenüber einer optimierten Wiedergabe der Verteilungsfunktion gegeben wird, wird meist die korrigierte Stichprobenvarianz verwendet, wenn das arithmetische Mittel der Grundgesamtheit unbekannt ist (was in praktischen Anwendungen so gut wie immer der Fall ist).

2.1.3 Schätzer für einen Anteilswert

Zu schätzende Größe: Der tatsächliche Anteil (Prozentsatz) jener Elemente der Grundgesamtheit, die eine bestimmte Eigenschaft aufweisen. (Das ist beispielsweise der Anteil grüner Kugeln in einer Urne, die mit grünen und nicht-grünen Kugeln gefüllt ist.)

$$\Pi = \frac{X}{N}$$

Schätzwert:

$$\hat{\pi} = \frac{x}{n}$$

Wenn die Stichprobenziehung mit Zurücklegen erfolgt, sind die Schätzwerte gemäß einer Binomialverteilung verteilt. Wenn ohne Zurücklegen gezogen wird, liegt eine hypergeometrische Verteilung vor. Bei gleicher Stichprobengröße weist die Binomialverteilung aber eine größere Varianz als die hypergeometrische Verteilung auf, daher ist das Ziehen ohne Zurücklegen effizienter als das Ziehen mit Zurücklegen. An diesem Beispiel wird auch deutlich, dass nicht nur die reine mathematische Schätzfunktion die Qualität der Schätzung beeinflusst, sondern auch die Art der Beschaffung der Stichprobe.

2.2 Wichtige Eigenschaften von Punktschätzungen

2.2.1 Unmöglichkeit die Qualität der Schätzung abzuschätzen.

Ein Beispiel: Österreich hat 9 Millionen Einwohner. Davon sind ca. 6,4 Millionen wahlberechtigt. Nehmen wir an, sie befragen 1000 zufällig ausgewählte wahlberechtigte Österreicher, welche Partei sie am nächsten Sonntag wählen würden, wenn eine Nationalratswahl stattfinden würde. Nehmen wir weiters an, davon geben 38 Personen an, sie würden die Partei X wählen. Damit liegt der Schätzwert für diesen Anteilswert bei 3,8%. Eine Partei braucht aber mindestens 4% aller gültigen Stimmen, um einen Sitz im Parlament zu erhalten. Eine Punktschätzung liefert aber keinerlei Anhaltspunkte dafür, wie wahrscheinlich es ist, dass die Partei X dennoch ins Parlament kommt. (Davon könnte abhängen, ob die Partei X aufgibt oder mehr Geld in den Wahlkampf investiert.)

Durch eine Punktschätzung erfahren wir nur einen einzigen Wert. Wenn dieser Wert in der Nähe eines wichtigen Grenzwertes liegt, erfahren wird bei einer Punktschätzung nicht, wie wahrscheinlich es ist, dass der tatsächliche Wert der Grundgesamtheit auf der anderen Seite des Grenzwertes liegt. Wir können bei einer Punktschätzung also keine Aussage darüber machen, wie gut oder schlecht unsere Schätzung ist.

Wenn man solche Angaben über die Verlässlichkeit der Schätzer haben will, muss man Intervallschätzungen machen.

2.2.2 Voraussetzung für Intervallschätzung

Angesehen davon, dass Punktschätzungen schon für sich allein nützliche Informationen bereitstellen, sind ohne sie auch keine Intervallschätzungen möglich. Punktschätzungen liefern die Basiswerte für Intervallschätzungen.

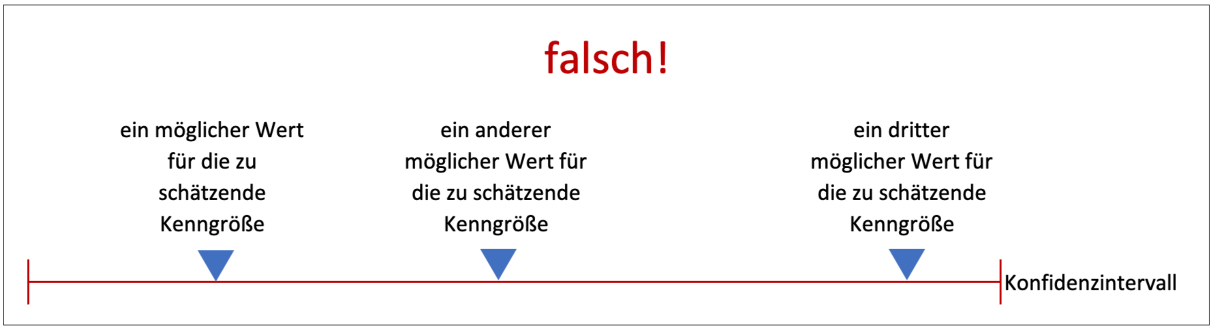
3 Intervallschätzung

Im Gegensatz zur Punktschätzung, wo man für eine bestimmte Kennzahl der Grundgesamtheit nur eine einzige Zahl als Schätzwert bekommt, verfolgt man bei der Intervallschätzung das Ziel, ein Intervall zu erhalten, das so liegt, dass seine Grenzen mit einer gewissen, vorgegebenen Wahrscheinlichkeit die Kennzahl der Grundgesamtheit umschließen. Diese vorgegebene Wahrscheinlichkeit heißt **Konfidenz** oder **Konfidenzniveau**, und das soeben beschriebene Intervall heißt **Konfidenzintervall**.

Beachte, dass die Kennzahl der Grundgesamtheit keine Zufallsgröße ist und nicht von der Stichprobe abhängt. Das gesuchte Intervall hingegen hängt sehr wohl von der Stichprobe ab und ist daher eine Zufallsgröße. Die folgende Formulierung ist daher streng genommen falsch:

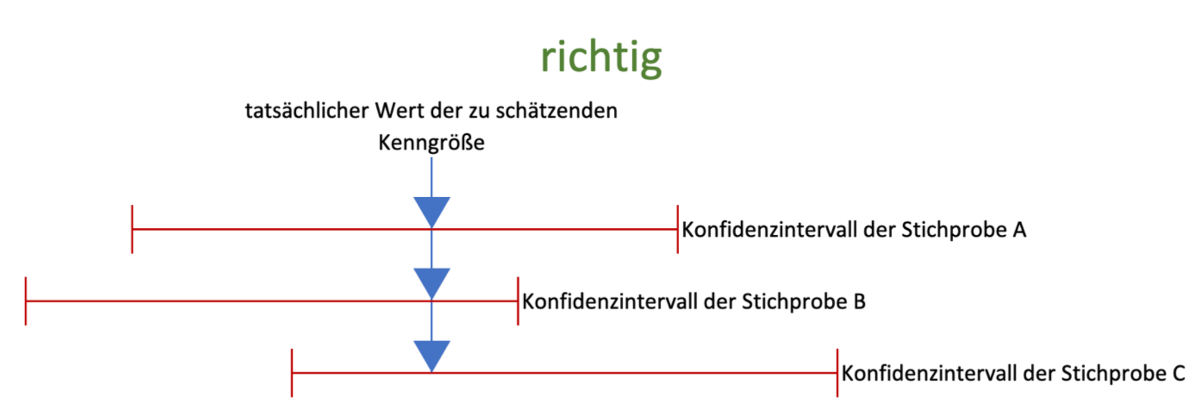
~~Der wahre Wert der zu schätzenden Kenngröße liegt mit einer gewissen Wahrscheinlichkeit im Konfidenzintervall.~~

Diese Formulierung erweckt den Anschein, als wäre das Konfidenzintervall ein festes Intervall, und die Kennzahl der Grundgesamtheit läge das eine Mal da und ein anderes Mal woanders. So ist es aber nicht.



Tatsächlich ist die folgende Formulierung zutreffend:

Das Konfidenzintervall kommt mit einer gewissen Wahrscheinlichkeit so zu liegen, dass seine Grenzen die zu schätzende Kennzahl der Grundgesamtheit umschließen.



Bei dieser Formulierung ist klar, dass das Intervall variabel ist und, abhängig von der gezogenen Stichprobe, mal da und mal dort liegt, während der wahre Wert der Kennzahl der Grundgesamtheit immer gleich ist. Wenn man sich das aber erst einmal bewusst gemacht hat, und auch verstanden hat, spricht nichts dagegen, auch die durchgestrichene Formulierung zu benutzen, weil sie ja auch so wie die fett hervorgehobene Formulierung interpretiert werden kann. Im Zweifel sollte man aber die eindeutigere Formulierung verwenden.

3.1 Erstellen eines Konfidenzintervall (Grundidee)

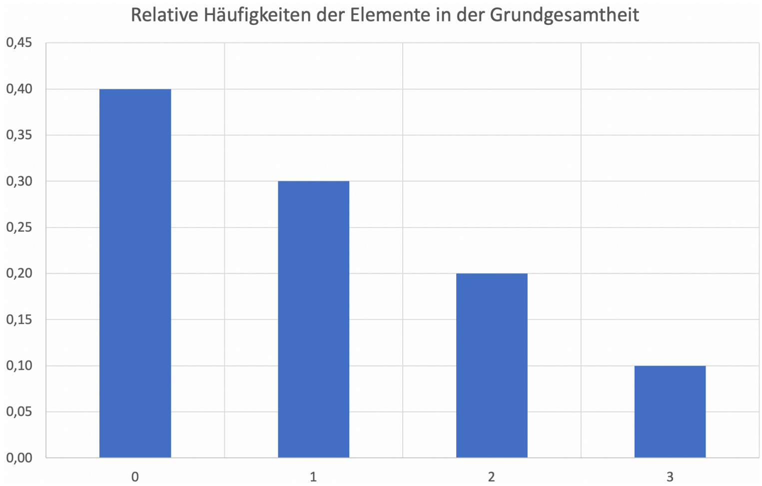
3.1.1 Inklusionsschluss

Eigentlich möchte man in der schließenden Statistik ja von einer bekannten Stichprobe auf bestimmte Eigenschaften einer unbekannten Grundgesamtheit schließen. Bevor wir diese Aufgabe in Abgriff nehmen, wollen wir zuerst den umgekehrten Weg gehen: Wir gehen von einer bekannten Grundgesamtheit aus, und überlegen uns, welche Stichproben man mit welchen Wahrscheinlichkeiten aus dieser Grundgesamtheit ziehen könnte. Daraus kann man nämlich berechnen, wie die Schätzwerte verschiedener Stichproben verteilt sind, und diese Erkenntnis hilft uns dann, das Konfidenzintervall zu verstehen und richtig anzuwenden.

Dazu soll uns dieses fiktive **Beispiel** helfen:

Das statistische Amt eines Landes hat erhoben, wie viele Haushalte wie viele Haustiere haben, und hat dabei folgendes festgestellt:

Anzahl Haustiere	Anteil der Haushalte mit dieser Anzahl
0	40%
1	30%
2	20%
3	10%



Sie werden bemerken, dass diese Zahlen nicht sehr plausibel sind, aber es geht darum, einfache Zahlen zum Rechnen zu haben.

Anhand von Stichproben soll das arithmetische Mittel der Anzahl der Haustiere pro Haushalt abgeschätzt werden.

Sie werden sich sicherlich an diese Formeln erinnern:

- Arithmetisches Mittel der Grundgesamtheit

$$\mu = \sum_{i=1}^{n} x_i \cdot p_i$$

- Varianz der Grundgesamtheit

$$\sigma^2 = \sum_{i=1}^{n} (x_i - \mu)^2 \cdot p_i$$

Die Standardabweichung der Grundgesamtheit σ ist die Quadratwurzel aus der Varianz.

Wir erhalten beim gegebenen Beispiel diese Werte:

$$\mu = 0 \cdot 0{,}4 + 1 \cdot 0{,}3 + 2 \cdot 0{,}2 + 3 \cdot 0{,}1 = 0{,}0 + 0{,}3 + 0{,}4 + 0{,}3 = 1$$

$$\begin{aligned} \sigma^2 &= (0-1)^2 \cdot 0{,}4 + (1-1)^2 \cdot 0{,}3 + (2-1)^2 \cdot 0{,}2 + (3-1)^2 \cdot 0{,}1 \\ &= 0{,}4 + 0{,}0 + 0{,}2 + 0{,}4 = 1 \end{aligned}$$

$$\sigma = \sqrt{1} = 1$$

Um die Größen der Grundgesamtheit besser von den Größen einer Stichprobe unterscheiden zu können, soll folgende Konvention gelten:

Parameter	Grundgesamtheit	Stichprobe
arithmetisches Mittel	μ	$\bar{x}$
Varianz	σ^2	s^2

Stichprobe mit Umfang = 1

Wenn man etwas über die Grundgesamtheit erfahren will, macht es in der Praxis nur wenig Sinn, eine Stichprobe zu ziehen, die nur 1 Element enthält. Da wir uns aber ansehen wollen, welche Stichproben möglich sind, macht es durchaus Sinn, darüber nachzudenken:

Es gibt genau vier mögliche Stichproben für die Anzahl der Haustiere: (0), (1), (2) und (3). Die Wahrscheinlichkeit, die Stichprobe (0) zu ziehen, entspricht genau dem Anteil der Haushalte ohne Haustiere, also 40%. Auch die anderen drei Stichproben haben Wahrscheinlichkeiten, die genau den Anteilen in der Grundgesamtheit entsprechen.

Voraussetzung dafür ist, dass das Ziehen der Stichprobe keinen systematischen Fehler enthält. Wenn man auf der Straße nur Personen befragt, die gerade einen Hund Gassi führt, wird man andere Wahrscheinlichkeiten erhalten.

Das arithmetische Mittel jeder Stichprobe ist natürlich der Wert des einzigen Elements der jeweiligen Stichprobe. Diese arithmetische Mittel treten mit diesen Wahrscheinlichkeiten auf:

$\bar{x}$	Stichproben mit diesem $\bar{x}$	Wahrscheinlichkeit dieses $\bar{x}$ zu finden
0	(0)	40%
1	(1)	30%
2	(2)	20%
3	(3)	10%

Daraus können wir uns den Erwartungswert $E(\bar{x})$ für den Mittelwert $\bar{x}$ ausrechnen. Erinnern wir uns, was der Erwartungswert bedeutet: Wir ziehen hintereinander viele Stichproben. Von jeder Stichprobe berechnen wir das arithmetische Mittel der Anzahl der Haustiere und erhalten auf diese Weise viele Mittelwerte (pro Stichprobe 1 Mittelwert). Der Erwartungswert ist der langfristige Mittelwert dieser Mittelwerte.

$$E(\bar{x}) = \sum_{i=1}^{n} \bar{x}_i \cdot p_i$$

Wir erhalten: $E(\bar{x}) = 1$

Ebenso können wir uns die Varianz dieser arithmetischen Mittelwerte ausrechnen:

$$\sigma_{\bar{x}}^2 = \sum_{i=1}^{n} \left(\bar{x}_i - E(\bar{x})\right)^2 \cdot p_i$$

Ergebnis: $\sigma_{\bar{x}}^2 = 1$

Wir halten fest:

Für $n = 1$:

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

$$\sigma_{\bar{x}}^2 = \sigma^2$$

Wenn man also eine Stichprobe der Größe 1 zieht, sind die Mittelwerte der Stichproben ganz genau gleich verteilt wie die eigentliche Größe in der Grundgesamtheit. Das ist kein großes Wunder, denn schließlich entspricht das arithmetische Mittel jeder Stichprobe genau einem Element der Grundgesamtheit. Daher sind natürlich auch alle statistischen Kennzahlen genau gleich groß wie bei der Grundgesamtheit.

Stichprobe mit Umfang = 2

Jetzt fängt es an interessanter zu werden.

Zunächst vereinbaren wir, dass das Ziehen der Stichprobe mit Zurücklegen erfolgt. Beim Stichprobenumfang 1 hat das noch keine Rolle gespielt, sobald aber mehrere Elemente gezogen werden, hat das Auswirkungen.

Begründung: Die Wahrscheinlichkeit, dass das erste Element der Stichprobe die Ausprägung X hat, entspricht (wenn das Verfahren zum Ziehen der Stichprobe keinen systematischen Fehler hat) genau dem Anteil der Elemente mit dieser Ausprägung in der Grundgesamtheit. Wenn das zweite Element der Stichprobe ohne Zurücklegen gezogen wird, fehlt das gerade zuvor gezogene Element in der Grundgesamtheit.

Damit sinkt beim zweiten Element die Wahrscheinlichkeit, ein Element mit derselben Ausprägung zu erwischen, während die Wahrscheinlichkeit für alle anderen Ausprägungen steigt. Um diesen Effekt zu umgehen, wird jedes Element nach der Ziehung zurückgelegt. Dadurch bleiben die Wahrscheinlichkeiten immer gleich.

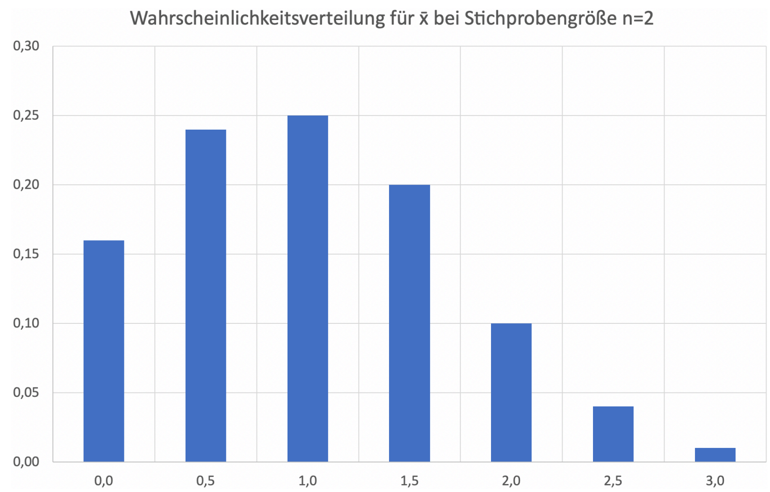
Ein weiterer Vorteil dieser Methode: Die Stichprobe kann dann mehr Elemente enthalten als die Grundgesamtheit. Im (nur in der Theorie möglichen) Extremfall kann die Stichprobe sogar unendlich viele Elemente enthalten. Das wäre beim Ziehen ohne Zurücklegen nicht möglich.

Es sind 16 (4^2) verschiedene Stichproben möglich:

Stichprobe	Wahrscheinlichkeit, diese Stichprobe zu ziehen	arithm. Mittel d. Stichpr. ($\bar{x}$)
$(0,0)$	$0,4 \cdot 0,4 = 0,16$	$0,0$
$(0,1)$	$0,4 \cdot 0,3 = 0,12$	$0,5$
$(0,2)$	$0,4 \cdot 0,2 = 0,08$	$1,0$
$(0,3)$	$0,4 \cdot 0,1 = 0,04$	$1,5$
$(1,0)$	$0,3 \cdot 0,4 = 0,12$	$0,5$
$(1,1)$	$0,3 \cdot 0,3 = 0,09$	$1,0$
$(1,2)$	$0,3 \cdot 0,2 = 0,06$	$1,5$
$(1,3)$	$0,3 \cdot 0,1 = 0,03$	$2,0$
$(2,0)$	$0,2 \cdot 0,4 = 0,08$	$1,0$
$(2,1)$	$0,2 \cdot 0,3 = 0,06$	$1,5$
$(2,2)$	$0,2 \cdot 0,2 = 0,04$	$2,0$
$(2,3)$	$0,2 \cdot 0,1 = 0,02$	$2,5$
$(3,0)$	$0,1 \cdot 0,4 = 0,04$	$1,5$
$(3,1)$	$0,1 \cdot 0,3 = 0,03$	$2,0$
$(3,2)$	$0,1 \cdot 0,2 = 0,02$	$2,5$
$(3,3)$	$0,1 \cdot 0,1 = 0,01$	$3,0$

Wir bewegen die dritte Spalte dieser Tabelle ganz nach links und fassen Stichproben mit demselben arithmetischen Mittel zusammen:

$\bar{x}$	Stichproben	Wahrscheinlichkeit
$0,0$	$(0,0)$	$0,16$
$0,5$	$(0,1), (1,0)$	$0,12 + 0,12 = 0,24$
$1,0$	$(0,2), (1,1), (2,0)$	$0,08 + 0,09 + 0,08 = 0,25$
$1,5$	$(0,3), (1,2), (2,1), (3,0)$	$0,04 + 0,06 + 0,06 + 0,04 = 0,20$
$2,0$	$(1,3), (2,2), (3,1)$	$0,03 + 0,04 + 0,03 = 0,10$
$2,5$	$(2,3), (3,2)$	$0,02 + 0,02 = 0,04$
$3,0$	$(3,3)$	$0,01$



Berechnen wir nun von dieser Verteilung der arithmetischen Mittelwerte den Erwartungswert und die Varianz:

$$E(\bar{x}) = \sum_{i=1}^{n} \bar{x}_i \cdot p_i$$

$$E(\bar{x}) = 0{,}0 \cdot 0{,}16 + 0{,}5 \cdot 0{,}24 + 1{,}0 \cdot 0{,}25 + 1{,}5 \cdot 0{,}20 + 2{,}0 \cdot 0{,}10 + 2{,}5 \cdot 0{,}04 + 3{,}0 \cdot 0{,}01$$

$$E(\bar{x}) = 0{,}00 + 0{,}12 + 0{,}25 + 0{,}30 + 0{,}20 + 0{,}10 + 0{,}03$$

$$E(\bar{x}) = 1{,}00$$

$$\sigma_{\bar{x}}^2 = \sum_{i=1}^{n} \left(\bar{x}_i - E(\bar{x})\right)^2 \cdot p_i$$

$$\sigma_{\bar{x}}^2 = (0{,}0-1)^2 \cdot 0{,}16 + (0{,}5-1)^2 \cdot 0{,}24 + (1{,}0-1)^2 \cdot 0{,}25 + (1{,}5-1)^2 \cdot 0{,}20 + (2{,}0-1)^2 \cdot 0{,}10 + (2{,}5-1)^2 \cdot 0{,}04 + (3{,}0-1)^2 \cdot 0{,}01$$

$$\sigma_{\bar{x}}^2 = 1{,}00 \cdot 0{,}16 + 0{,}25 \cdot 0{,}24 + 0{,}00 \cdot 0{,}25 + 0{,}25 \cdot 0{,}20 + 1{,}00 \cdot 0{,}10 + 2{,}25 \cdot 0{,}04 + 4{,}00 \cdot 0{,}01$$

$$\sigma_{\bar{x}}^2 = 0{,}16 + 0{,}06 + 0{,}00 + 0{,}05 + 0{,}10 + 0{,}09 + 0{,}04$$

$$\sigma_{\bar{x}}^2 = 0{,}50$$

Wir halten fest:

Für $n = 2$:

$$E(\bar{x}) = \mu \qquad\qquad \sigma_{\bar{x}}^2 = \frac{\sigma^2}{2}$$

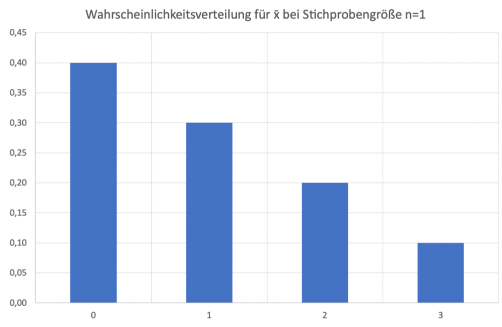
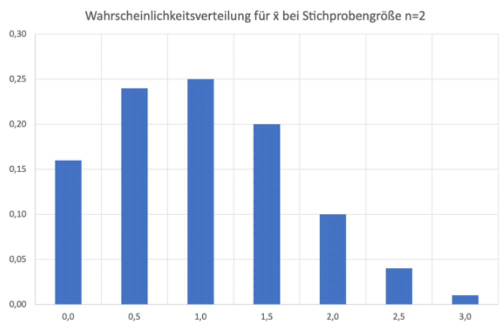
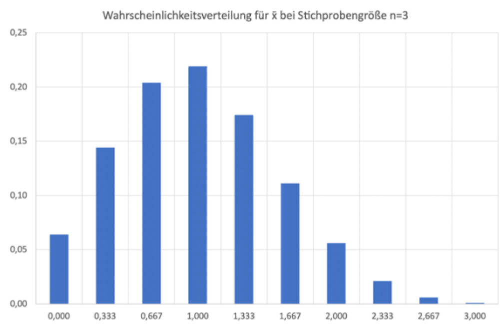
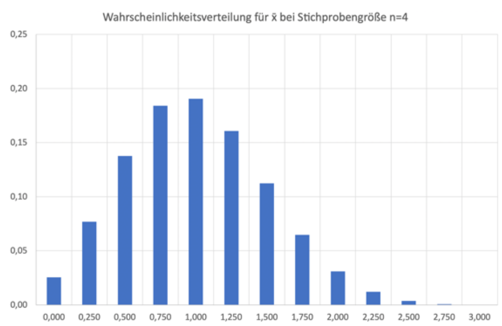
Die Verteilung der Mittelwerte der Stichproben sieht nun anders aus als die Verteilung der Elemente in der Grundgesamtheit. Trotzdem ist der Erwartungswert für das arithmetische Mittel einer Stichprobe der Größe 2 genau gleich groß wie das arithmetische Mittel der Elemente der Grundgesamtheit. Die Varianz des arithmetischen Mittels der Stichproben ist bei einem Stichprobenumfang von 2 Elementen genau die Hälfte der Varianz der Elemente in der Grundgesamtheit.

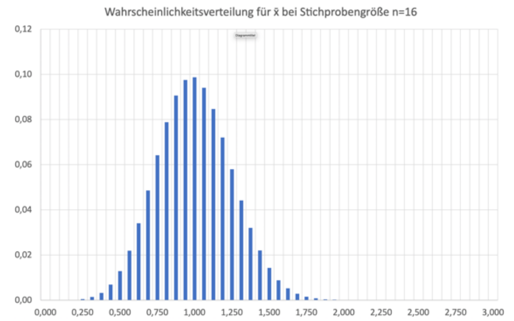
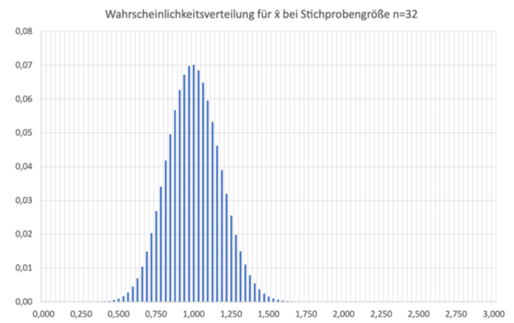
Stichprobe mit Umfang = 3

Es gibt 64 verschiedene Stichproben der Größe 3. Auf die Darstellung der detaillierten Zwischenergebnisse wir hier verzichtet, die Berechnungen laufen nach demselben Schema wie bei $n = 1$ und $n = 2$ ab. Wir erhalten:

$$E(\bar{x}) = 1 = \mu \qquad\qquad \sigma_{\bar{x}}^2 = \frac{1}{3} = \frac{\sigma^2}{3}$$

Stichproben mit größeren Umfängen

Umfang n	Verteilung	Erwartungswert $E(\bar{x})$	Varianz $\sigma_{\bar{x}}^2$
1		$1 = \mu$	$1 = \frac{\sigma^2}{1} = \frac{\sigma^2}{n}$
2		$1 = \mu$	$\frac{1}{2} = \frac{\sigma^2}{2} = \frac{\sigma^2}{n}$
3		$1 = \mu$	$\frac{1}{3} = \frac{\sigma^2}{3} = \frac{\sigma^2}{n}$
4		$1 = \mu$	$\frac{1}{4} = \frac{\sigma^2}{4} = \frac{\sigma^2}{n}$

Umfang n	Verteilung	Erwartungswert $E(\bar{x})$	Varianz $\sigma_{\bar{x}}^2$
8		$1 = \mu$	$\frac{1}{8} = \frac{\sigma^2}{8} = \frac{\sigma^2}{n}$
16		$1 = \mu$	$\frac{1}{16} = \frac{\sigma^2}{16} = \frac{\sigma^2}{n}$
32		$1 = \mu$	$\frac{1}{32} = \frac{\sigma^2}{32} = \frac{\sigma^2}{n}$

Zusammenfassend erkennt man:

- Je größer der Stichprobenumfang ist, desto stärker ähnelt die Verteilung der arithmetischen Mittelwerte der Stichproben einer Normalverteilung.

  Obwohl die Verteilung der Elemente in der Grundgesamtheit eine sehr schiefe Verteilung diskrete Verteilung mit nur 4 möglichen Werten ist, sieht die Verteilung der Stichprobenmittelwerte bei einer Stichprobengröße von 16 einer Normalverteilung bereits zum Verwechseln ähnlich.

  Als Faustregel gilt: Wenn die Stichprobe 30 oder mehr Elemente umfasst, kann man bei jeder beliebigen Verteilung der Elemente der Grundgesamtheit die Verteilung der Stichprobenmittelwerte trotzdem als annähernd normalverteilt angenommen werden. Je stärker die Verteilungsfunktion der Grundgesamtheit eine Normalverteilung ähnelt, desto weniger Elemente benötigt man in den Stichproben, um dort eine Normalverteilung zu erreichen.

- Der Erwartungswert für das arithmetische Mittel einer Stichprobe ist unabhängig von der Größe der Stichprobe und stimmt mit dem arithmetischen Mittel der Ausprägungen der Grundgesamtheit genau überein.

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

- Die Varianz der arithmetischen Mittelwerte vieler Stichproben wird umso kleiner, je größer der Stichprobenumfang ist. Wenn σ^2 die Varianz der Ausprägungen der Grundgesamtheit ist, und wenn n der Stichprobenumfang ist, dann erhält man die Varianz der arithmetischen Mittelwerte vieler Stichproben $\sigma_{\bar{x}}^2$ indem man σ^2 durch die Anzahl der Elemente in der Stichprobe teilt:

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$$

Die Annäherung an die Normalverteilung und die beiden Formeln $E(\bar{x}) = \mu$ und $\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$ wurden hier zwar nur anhand eines Beispiels plausibel gemacht, gelten aber immer (für jede beliebig verteilte Grundgesamtheit). Auf einen ausführlichen Beweis wird an dieser Stelle verzichtet. Mehr darüber kann man in den vorangegangenen Skripten nachlesen:

- Das sich die Verteilung der arithmetischen Mittelwerte der Stichproben bei zunehmendem Stichprobenumfang immer mehr einer Normalverteilung annähert, ist eine direkte Folge des zentralen Grenzwertsatzes (siehe Skriptum »Stetige Verteilungen«, Kapitel 1.2 »Zentraler Grenzwertsatz«)
- Die beiden Formeln bestätigen gemeinsam noch einmal das Gesetz der großen Zahlen (siehe Skriptum »Zufallsvariablen«, Abschnitt 6 »Gesetz der großen Zahlen«)

Definition »Standardfehler des Mittelwerts«

In vielen Statistikbüchern wird im Zusammenhang mit dem Erwartungswert des Mittelwerts vieler Stichproben auch der Ausdruck »Standardfehler« oder »Standardfehler des Mittelwerts« verwendet. Das ist nichts anderes als die Wurzel aus der Varianz der Mittelwerte vieler Stichproben und wird mit dem Symbol $\sigma_{\bar{x}}$ oder auch einfach nur mit s geschrieben:

$$\sigma_{\bar{x}} = \sqrt{\sigma_{\bar{x}}^2} = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}$$

Man erhält also den Standardfehler einer Stichprobe, indem man die Standardabweichung der Grundgesamtheit durch die Wurzel der Stichprobengröße teilt.

3.1.2 Wahrscheinlichkeit für ein bestimmtes Intervall

Wenn man diese Zusammenhänge kennt, kann man sich auch überlegen, wie groß die Wahrscheinlichkeit dafür ist, dass der Mittelwert einer zufälligen Stichprobe innerhalb eines bestimmten Intervalls liegt. Wir können uns bei unserer Haustier-Verteilung beispielsweise fragen, wie groß die Wahrscheinlichkeit dafür ist, dass der Mittelwert einer Stichprobe, die 16 Elemente umfasst, zwischen 0,9 und 1,3 liegt.

Wie wir oben gesehen haben, beträgt in unserem Beispiel bei einer Stichprobengröße von 16 die Varianz $\frac{1}{16}$ und der Erwartungswert ist genau 1. Im Skriptum über stetige Verteilungen wurde gezeigt, wie man diese Verteilung in eine Standardnormalverteilung umrechnet. Dazu muss die gesamte Verteilung mit dieser Formel umgerechnet werden:

$$Z = \frac{X - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$$

mit $\mu_{\bar{x}} = E(\bar{x}) = 1$ und $\sigma_x^2 = \sigma_{\bar{x}}^2 = \frac{1}{16}$ also $\sigma_{\bar{x}} = \sqrt{\frac{1}{16}} = \frac{1}{4}$

Damit müssen wir natürlich auch die beiden Grenzen umformen:

$$g_1 = \frac{0{,}9 - 1}{\frac{1}{4}} = (-0{,}1) \cdot 4 = -0{,}4$$

$$g_2 = \frac{1{,}3 - 1}{\frac{1}{4}} = (0{,}3) \cdot 4 = 1{,}2$$

Die Wahrscheinlichkeit P, dass der Mittelwert im gesuchten Intervall liegt, ist die Fläche unter der Kurve der Wahrscheinlichkeitsdichtefunktion $\varphi(x)$ der Standardnormalverteilung zwischen den Stellen $g_1 = -0{,}4$ und $g_2 = 1{,}2$, und das wiederum ist der Wert des bestimmten Integrals dieser Funktion zwischen diesen beiden Werten:

$$P = \int_{g_1}^{g_2} \varphi(x)\, dx$$

$$g_1 = -0{,}4; \qquad g_2 = 1{,}2; \qquad \varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}x^2}$$

$$P = \frac{1}{\sqrt{2\pi}} \int_{-0{,}4}^{1{,}2} e^{-\frac{1}{2}x^2}\, dx$$

Die Verteilungsfunktion $\Phi(x)$ der Standardnormalverteilung ist aber bereits die Stammfunktion (also das unbestimmte Integral) der Wahrscheinlichkeitsdichtefunktion, daher muss man nur g_1 und g_2 in die Verteilungsfunktion einsetzen, und die Differenz der beiden Funktionswerte ist dann die gesuchte Wahrscheinlichkeit:

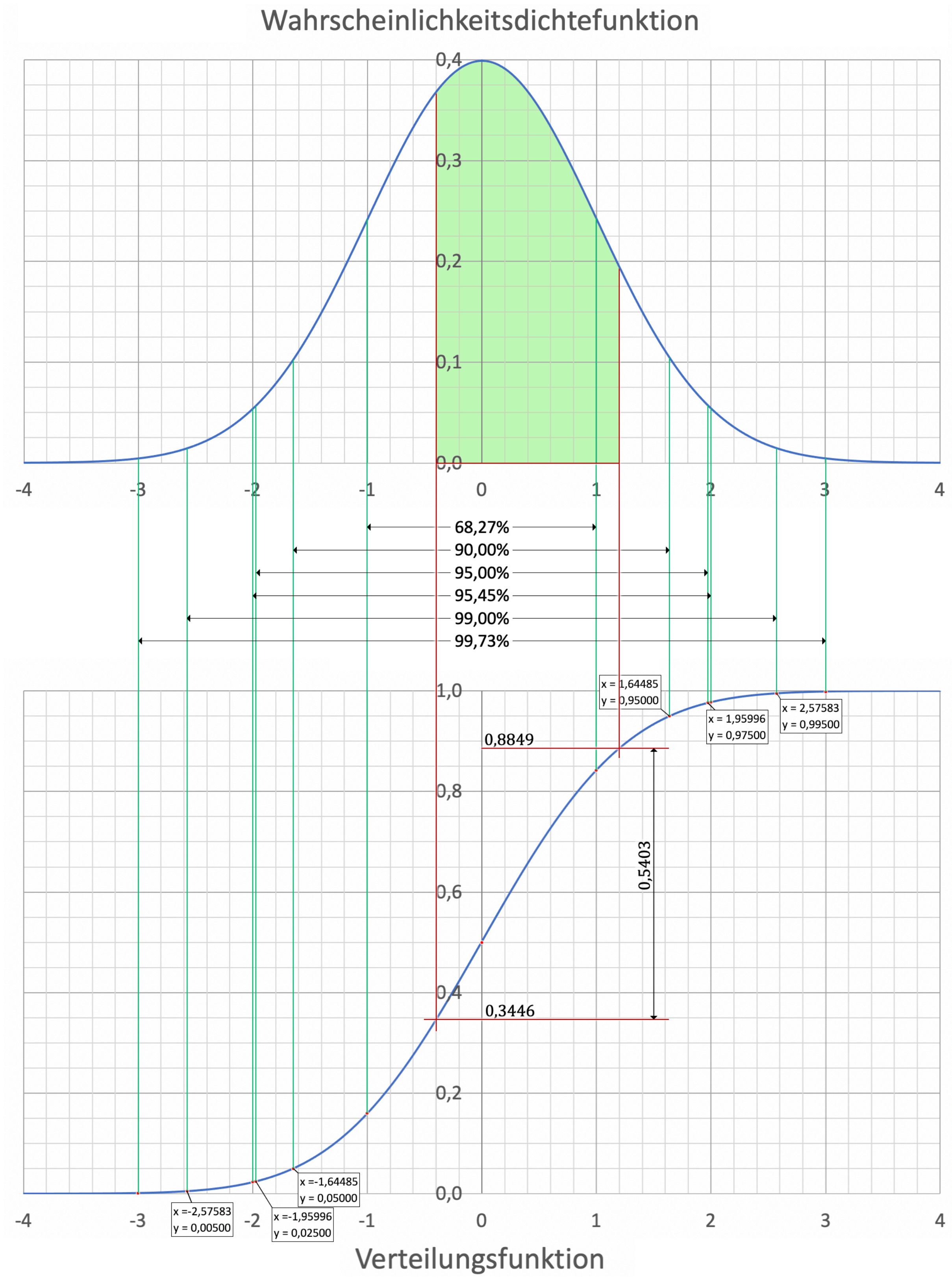
$$P = \Phi(g_2) - \Phi(g_1)$$

$$P = \Phi(1{,}2) - \Phi(-0{,}4)$$

$$P = 0{,}8849 - 0{,}3446$$

$$P = 0{,}5403 = 54{,}03\%$$

Auf der nächsten Seite ist anhand der Standardnormalverteilung und ihrer Verteilungsfunktion dargestellt, wie man von den Grenzen -0,4 und 1,2 zum Ergebnis kommt.



Beispielsanwendung der Standardnormalverteilung

3.1.3 Intervall mit einer bestimmten Wahrscheinlichkeit

Häufiger ist man aber an der Antwort auf die umgekehrte Frage interessiert: Von wo bis wo geht ein Intervall, innerhalb dessen der wahre Mittelwert der Grundgesamtheit mit einer Wahrscheinlichkeit von X Prozent liegt?

Solange man nicht eine der beiden Grenzen des Intervalls oder einen anderen fixen Punkt innerhalb des Intervalls vorgibt, gibt es unendlich viele Antworten auf diese Frage. In der Praxis läuft die Fragestellung daher auf eine dieser Fragen hinaus:

1.a. Wo endet das Intervall, dass bei $-\infty$ beginnt, und das mit einer Wahrscheinlichkeit von X% den wahren Mittelwert der Grundgesamtheit beinhaltet?

1.b. Wo beginnt das Intervall, dass bei $+\infty$ endet, und das mit einer Wahrscheinlichkeit von X% den wahren Mittelwert der Grundgesamtheit beinhaltet?

2. Wo beginnt und wo endet das Intervall, das symmetrisch um den Mittelwert der Stichprobe liegt, und das mit einer Wahrscheinlichkeit von X% den wahren Mittelwert der Grundgesamtheit beinhaltet?

Symmetrische Intervalle

Beginnen wir mit dem symmetrischen Intervall (Fragestellung 2). Man kann natürlich nach beliebigen Wahrscheinlichkeiten fragen, je nach Forschungsfeld haben sich aber diese Standardwerte etabliert:

- 90%: Intervall: von $-g_{90}$ bis $+g_{90}$; Irrtumswahrscheinlichkeit $\alpha = 10\%$

$$1 - \alpha = 0{,}9 = \int_{-g_{90}}^{+g_{90}} \varphi(x)\, dx$$

- 95%: Intervall: von $-g_{95}$ bis $+g_{95}$; Irrtumswahrscheinlichkeit $\alpha = 5\%$

$$1 - \alpha = 0{,}95 = \int_{-g_{95}}^{+g_{95}} \varphi(x)\, dx$$

- 99%: Intervall: von $-g_{99}$ bis $+g_{99}$; Irrtumswahrscheinlichkeit $\alpha = 1\%$

$$1 - \alpha = 0{,}99 = \int_{-g_{99}}^{+g_{99}} \varphi(x)\, dx$$

Die Grenzwerte g_{90}, g_{95} und g_{99} für die Standardnormalverteilung muss man sich nur einmal ausrechnen oder aus einer Tabelle herauslesen, und kann sie dann immer wieder verwenden:

$g_{90} = 1{,}6448536269514$ $\quad \Phi(-g_{90}) = 0{,}050;\ \Phi(g_{90}) = 0{,}950;\ \Phi(g_{90}) - \Phi(-g_{90}) = 0{,}900$
$g_{95} = 1{,}9599639845399$ $\quad \Phi(-g_{95}) = 0{,}025;\ \Phi(g_{95}) = 0{,}975;\ \Phi(g_{95}) - \Phi(-g_{95}) = 0{,}950$
$g_{99} = 2{,}5758293035485$ $\quad \Phi(-g_{99}) = 0{,}005;\ \Phi(g_{99}) = 0{,}995;\ \Phi(g_{99}) - \Phi(-g_{99}) = 0{,}990$

Einseitige Intervalle

Beim einseitigen Intervall ist die Vorgehensweise im Prinzip dieselbe, mit dem Unterschied, dass eine der beiden Intervallgrenzen bei $-\infty$ oder $+\infty$ liegt:
1.a.

$$1-\alpha = \int_{-\infty}^{g} \varphi(x)\,dx$$

1.b.

$$1-\alpha = \int_{g}^{\infty} \varphi(x)\,dx$$

Die Intervallgrenze g hängt also, wie in den beiden obenstehenden Formeln dargestellt, von α ab und kann ebenfalls aus einer Tabelle ausgelesen werden, oder vorberechnet werden. Die weitere Vorgehensweise bei einseitigen Intervallen entspricht – bis auf die Tatsache, dass eine der Intervallgrenzen den fixen Wert $-\infty$ oder $+\infty$ hat – der Vorgehensweise für symmetrische Intervalle.

Fortführung des Beispiels

Wir wollen hier mit einem Intervall fortfahren, das mit einer Wahrscheinlichkeit von 95% den wahren Mittelwert der Grundgesamtheit enthält. Die Irrtumswahrscheinlichkeit α soll also 5% betragen.

Wir wissen nun, dass in der Standardnormalverteilung der X-Wert zwischen $g_1 = -1{,}95996$ und $g_2 = +1{,}95996$ liegen muss. Die Verteilung der Mittelwerte einer Stichprobe aus 16 Elementen, ist aber keine *Standard*normalverteilung, sondern "nur" eine gewöhnliche Normalverteilung (und das auch nur näherungsweise) mit den Parametern $E(\bar{x}) = 1$ und $\sigma_{\bar{x}}^2 = \frac{1}{16}$. Wir können mit der folgenden Formel von der Standardnormalverteilung wieder zurück zur tatsächlichen Verteilung rechnen (siehe Skriptum über stetige Verteilungen, Kapitel 1.1.3):

$$X = \sigma_{\bar{x}} \cdot Z + \mu_{\bar{x}}$$

mit $\mu_{\bar{x}} = E(\bar{x}) = 1$ und $\sigma_{\bar{x}}^2 = \sigma_{\bar{x}}^2 = \frac{1}{16}$ also $\sigma_{\bar{x}} = \sqrt{\frac{1}{16}} = \frac{1}{4}$. Also:

$$X = \frac{1}{4} \cdot Z + 1$$

Wir setzen für Z die Werte für g_1 und g_2 ein und erhalten:

$$x_1 = \frac{1}{4} \cdot g_1 + 1 = \frac{-1{,}95996}{4} + 1 = 0{,}510009$$
$$x_2 = \frac{1}{4} \cdot g_2 + 1 = \frac{+1{,}95996}{4} + 1 = 1{,}489991$$

Damit wissen wir, dass bei einem Stichprobenumfang von 16 Elementen (mindestens) 95% aller Stichproben einen arithmetischen Mittelwert zwischen ca. 0,51 und 1,49 haben werden.

3.2 Konfidenzintervall für das arithmetische Mittel

3.2.1 Konfidenzintervall für das arithmetische Mittel, wenn σ bekannt ist

Wir interessieren uns in der Praxis selten dafür, welche Stichproben man aus einer Grundgesamtheit ziehen kann, und wie weit entfernt vom wahren Mittelwert entfernt der Mittelwert einer theoretischen Stichprobe sein könnte. Vielmehr ist es normalerweise so, dass wir eine Stichprobe mit einer bestimmten Größe haben, daraus das arithmetische Mittel berechnen, und gerne angeben würden, innerhalb welcher Grenzen wir den wahren Mittelwert mit einer bestimmten Wahrscheinlichkeit vermuten. Dazu drehen wir einfach den Spieß um.

Wir kennen μ, den wahren Mittelwert der Grundgesamtheit nicht, haben aber von einem Wahrsager oder einer anderen vertrauenswürdigen Quelle den genauen Wert für die Standardabweichung und damit für die Varianz erfahren, und wir haben aus der Grundgesamtheit eine Stichprobe gezogen, die n Elemente enthält, wobei n groß genug ist, um davon auszugehen, dass die Mittelwerte unterschiedlicher Stichproben normalverteilt sind.

Beispiel: Wir gehen von diesen Daten aus (Körpergrößen von 26 Personen in cm):

160, 161, 163, 165, 169, 169, 170, 170, 172, 173, 175, 175, 175, 175, 175, 175, 177, 178, 180, 180, 181, 181, 185, 187, 189, 190

Das Orakel sagt uns: Die Standardabweichung der Körpergrößen ist genau 8,0.

Das Konfidenzniveau soll 95% sein. Das heißt: Bei 95% aller Stichproben wird das Intervall, das wir nun ausrechnen werden, die wahre mittlere Körpergröße der Gesamtbevölkerung beinhalten. Bei 5% aller Stichproben werden wir ein Intervall erhalten, das vollständig neben dem wahren Mittelwert liegt.

Schritt 1: Punkschätzung

Der erste Schritt ist eine Punktschätzung für das arithmetische Mittel. Das ist genau das arithmetische Mittel der Stichprobe. Wir erhalten damit $\bar{x}$. Der Wert in unserem Beispiel ist genau 175,0.

Schritt 2: Stichprobenvarianz berechnen

Wir kennen ja von einem Orakel die Varianz der Grundgesamtheit, und wir kennen den Umfang der Stichprobe. Daraus berechnen wir, mit welcher Varianz die Mittelwerte vieler Stichproben streuen:

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} = \frac{8^2}{26} = \frac{64}{26} \cong 2{,}4615$$

Der Standardfehler ist die Wurzel daraus:

$$\sigma_{\bar{x}} = \sqrt{2{,}4615} = 1{,}5689$$

Schritt 3: Intervallgrenzen berechnen

Wenn wir die Werte für $\bar{x}$ und $\sigma_{\bar{x}}^2$ in die Formel einsetzen, mit der man eine Standardnormalverteilung in eine beliebige Normalverteilung umrechnet, kann man damit die Grenzwerte für die gegebene Irrtumswahrscheinlichkeit in die gesuchten Intervallgrenzen umrechnen:

$$X = \sigma_{\bar{x}} \cdot Z + \mu_{\bar{x}}$$

$$x_1 = \sigma_{\bar{x}} \cdot (-g_{95}) + \bar{x}$$

$$x_2 = \sigma_{\bar{x}} \cdot (+g_{95}) + \bar{x}$$

Den Wert für $g_{95} = 1{,}9599639845399$ haben wir schon weiter oben ermittelt.

$$x_1 = 1{,}5689 \cdot (-1{,}96) + 175 \cong 171{,}92$$

$$x_2 = 1{,}5689 \cdot (+1{,}96) + 175 \cong 178{,}08$$

Damit sind wir bereits fertig. x_1 und x_2 (ca. 172 und 178 cm) sind die beiden Grenzen des gesuchten Intervalls.

3.2.2 Konfidenzintervall für das arithmetische Mittel, wenn σ nicht bekannt ist

Leider hat man im wirklichen Leben nur selten ein Orakel zur Hand, das einem die Varianz der Grundgesamtheit verrät. Daher muss man die Varianz der Mittelwerte vieler Stichproben aus der einen Stichprobe abschätzen, die man zur Verfügung hat.

Schritt 1: Punkschätzung

Der erste Schritt ist derselbe wie zuvor. Wir erhalten $\bar{x} = 175$.

Schritt 2: Stichprobenvarianz abschätzen

Wir berechnen die erwartungstreue Varianz der Stichprobe (das ist die korrigierte Varianz aus 2.1.2)

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

Das ist der Schätzwert für die Varianz der Grundgesamtheit. Wenn wir die Daten unseres Beispiels in diese Formel einsetzen, erhalten wir $s^2 = 64{,}0$ bzw. $s = 8{,}0$.

Schritt 3: Intervallgrenzen ermitteln

Wie bereits im Skriptum über stetige Verteilungen, im Abschnitt über die t-Verteilung ausführlich dargestellt, ist die Verteilung der Varianzen einer Stichprobe (im Gegensatz zur Verteilung der Mittelwerte vieler verschiedener Stichproben) nicht normalverteilt, sondern unterliegt einer Studentschen t-Verteilung mit $n-1$ Freiheitsgraden. Im Gegensatz zur Standardnormalverteilung, wo die Grenzwerte vom Umfang der Stichprobe unabhängig sind, muss man bei der t-Verteilung diesen Umfang in Form der Anzahl der Freiheitsgrade berücksichtigen.

Wir ermitteln den t-Wert für ein symmetrisches Intervall, wenn das Konfidenzniveau 95% beträgt für $n-1=25$ Freiheitsgrade. Diesen Wert kann man einer vorgefertigten Tabelle entnehme, oder mit Software berechnen. Das Ergebnis ist $t=2{,}0595\cdots$

Wie oben ausführlich ausgeführt, muss man noch durch die Wurzel aus der Stichprobengröße dividieren und erhält dann am Ende diese Formel:

$$x_1=\bar{x}-\frac{s}{\sqrt{n}}\cdot t=175-\frac{8}{\sqrt{26}}\cdot 2{,}0595=175-3{,}23\cdots\cong 171{,}77$$

$$x_2=\bar{x}+\frac{s}{\sqrt{n}}\cdot t=175+\frac{8}{\sqrt{26}}\cdot 2{,}0595=175+3{,}23\cdots\cong 178{,}23$$

Auch hier taucht der Wert $\frac{8}{\sqrt{26}}$ auf. Das ist wieder der Standardfehler. Die Zahl 8 ist jetzt aber nicht die wahre Standardabweichung der Grundgesamtheit, sondern der Schätzwert, den wir aus unserer Stichprobe berechnet haben.

Die (standardisierte) Studentschen t-Verteilung kann bei genügend großem Stichprobenumfang durch eine Standardnormalverteilung approximiert werden. Als Faustregel gilt, dass bei einem Stichprobenumfang von 30 oder mehr Elementen in der Stichprobe die Standardnormalverteilung verwendet werden kann, und man sich das Hantieren mit der Studentschen t-Verteilung und ihren Freiheitsgraden ersparen kann.

3.3 Ziehen mit und ohne Zurücklegen

Wenn Sie sich an den Beginn dieses Abschnitts zurückerinnern, wissen Sie noch, dass wir davon ausgegangen sind, dass wir die Stichproben **mit** Zurücklegen gezogen haben. Es konnte also jedes Element aus der Grundgesamtheit mehrfach in der Stichprobe vorkommen, und die Stichprobe konnte sogar größer als die Grundgesamtheit sein. Der Grund, warum wir das gemacht haben, war einfach, dass beim Ziehen mit Zurücklegen bei jedem neuen Element der Stichprobe die Wahrscheinlichkeiten der Ausprägungen gleich war.

Wenn wir **ohne** Zurücklegen ziehen, ändert sich die Wahrscheinlichkeit, als nächstes ein Element mit einer bestimmten Ausprägung zu ziehen, in Abhängigkeit von den bereits gezogenen Elementen.

Das hat keine Auswirkungen auf den Wert von $\bar{x}$, aber bei der Berechnung von $\sigma_{\bar{x}}^2$ (bei bekannter Varianz σ^2 der Grundgesamtheit) bzw. von $\hat{\sigma}_{\bar{x}}^2$ (wenn diese Varianz unbekannt ist) muss man Korrekturen anbringen:

	Varianz σ^2 der Grundgesamtheit ist bekannt	Varianz σ^2 der Grundgesamtheit ist unbekannt
Stichprobenziehung mit Zurücklegen	$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$	$\hat{\sigma}_{\bar{x}}^2 = \frac{s^2}{n-1}$
Stichprobenziehung ohne Zurücklegen	$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} \cdot \frac{N-n}{N-1}$	$\hat{\sigma}_{\bar{x}}^2 = \frac{s^2}{n-1} \cdot \left(1 - \frac{n}{N}\right)$

Dabei sind:

σ^2 Varianz der Elemente der Grundgesamtheit
s^2 Varianz der Elemente einer Stichprobe
$\sigma_{\bar{x}}^2$ tatsächliche Varianz der arithmetischen Mittelwerte vieler Stichproben
$\hat{\sigma}_{\bar{x}}^2$ aus den Daten einer einzelnen Stichprobe geschätzte Varianz der arithmetischen Mittelwerte vieler Stichproben
N Anzahl der Elemente der Grundgesamtheit
n Anzahl der Elemente der Stichprobe (Stichprobenumfang)

3.4 Konfidenzintervall für den Abstand zweier Grundgesamtheiten

Beispiel:

Man befragt 100 Österreicher und 50 Tschechen nach ihren Bierkonsum (Liter pro Person und Jahr) und möchte nicht nur wissen, um wie viel Bier ein durchschnittlicher Tscheche mehr trinkt als ein durchschnittlicher Österreicher, sondern man möchte auch das symmetrische 95%-Konfidenzintervall für diese Differenz angeben.

Dazu zieht man aus beiden Grundgesamtheiten die Stichproben, die unterschiedlich groß sein dürfen. Aus jeder Stichprobe berechnet man das arithmetische Mittel, und der Erwartungswert für die Differenz der wahren Mittelwerte ist dann die Differenz der beiden Stichprobenmittelwerte.

$$E(\mu_1 - \mu_2) = \bar{x}_1 - \bar{x}_2$$

Wenn $\bar{x}_1$ und $\bar{x}_2$ normalverteilt sind (was immer der Fall ist, wenn die Grundgesamtheiten normalverteilt sind, und was näherungsweise der Fall ist, wenn die Stichproben mehr als ca. 30 Elemente enthalten), dann ist auch die Differenz $\bar{x}_1 - \bar{x}_2$ normalverteilt, und dann kann man das Konfidenzintervall nach derselben Methode ermitteln, wie oben beschrieben. Man muss nur die richtigen Varianzen verwenden:

3.4.1 Wenn die Varianzen beider Grundgesamtheiten bekannt sind

In diesem Fall gilt:

$$\sigma_{\bar{x}}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

Dabei sind σ_1^2 und σ_2^2 die Varianzen der beiden Grundgesamtheiten und n_1 und n_2 sind die Größen der beiden Stichproben. Der so ermittelte Wert für $\sigma_{\bar{x}}^2$ kann wie oben beschrieben verwendet werden.

3.4.2 Wenn die Varianzen beider Grundgesamtheiten gleich, aber unbekannt sind

Man kennt also die Varianzen beider Grundgesamtheiten nicht, kann aber davon ausgehen, dass sie (zumindest annähernd) gleich groß sind.

Dann berechnet man, wie oben beschrieben, die Varianzen s_1^2 und s_2^2 der beiden Stichproben, und daraus die sogenannte gepoolte Varianz

$$s_p^2 = \frac{(n_1 - 1) \cdot s_1^2 + (n_2 - 1) \cdot s_2^2}{n_1 + n_2 - 2}$$

Der Standardfehler der Differenz der Mittelwerte unter der Annahme gleicher Varianzen ist dann:

$$SE = s_p \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

Es ist wieder eine t-Verteilung anzuwenden, und zwar mit $n_1 - n_2 - 2$ Freiheitsgraden.