8. Stichproben

8.1 Grundgesamtheit und Stichprobe

In der **induktiven Statistik** (**beurteilende Statistik**) gehen wir von **Stichprobendaten** aus. Speziell stammen die Daten aus **Zufallsstichproben**. Die Stichprobenergebnisse sollen auf die Grundgesamtheit übertragen werden. Im Folgenden beschränken wir den Begriff der Stichprobe stets auf das Konzept der Zufallsstichprobe.

Eine **Grundgesamtheit** enthält alle interessierenden statistischen Einheiten (Käufer, Regionen etc.), während eine **Stichprobe** nur zufällig ausgewählte Einheiten aus der Grundgesamtheit umfasst.

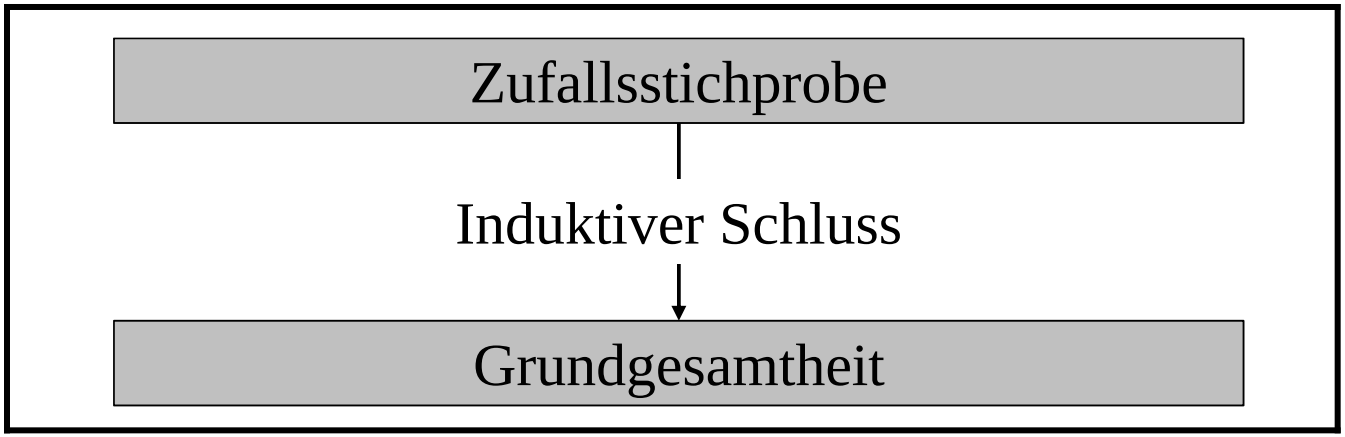
Aus einer **Vollerhebung** lassen sich **Kenngrößen (Parameter)** wie z.B. Mittelwerte, Streuungsmaße, Anteilswerte einer **Grundgesamtheit** prinzipiell **exakt** bestimmen. Eine **Stichprobe** enthält dagegen **Informationen**, die **mit Zufallsschwankungen behaftet** sind. Die **Verallgemeinerung von Stichprobenergebnissen** auf die Grundgesamtheit ist aufgrund unvollständiger Informationen daher stets mit Ungenauigkeiten verbunden. Dieser Rückschluss von den Beobachtungen in der Stichprobe auf bestimmte Charakteristika der Grundgesamtheit wird als **induktiver Schluss** bezeichnet.

Gründe für Stichprobenerhebungen:

- Kostenersparnis
- Zeitgewinn (Aktualität)
- Gründlichere Durchführung
- Praktische Unmöglichkeit von Vollerhebungen (zerstörende Tests von Produkten)

Unter einem **induktiven Schluss** versteht man die Verallgemeinerung von Stichprobenergebnissen auf die Grundgesamtheit.

Abbildung: Induktiver Schluss



Stichprobenfehler: Zufällige Abweichung eines Stichprobenergebnisses von einem Parameter der Grundgesamtheit

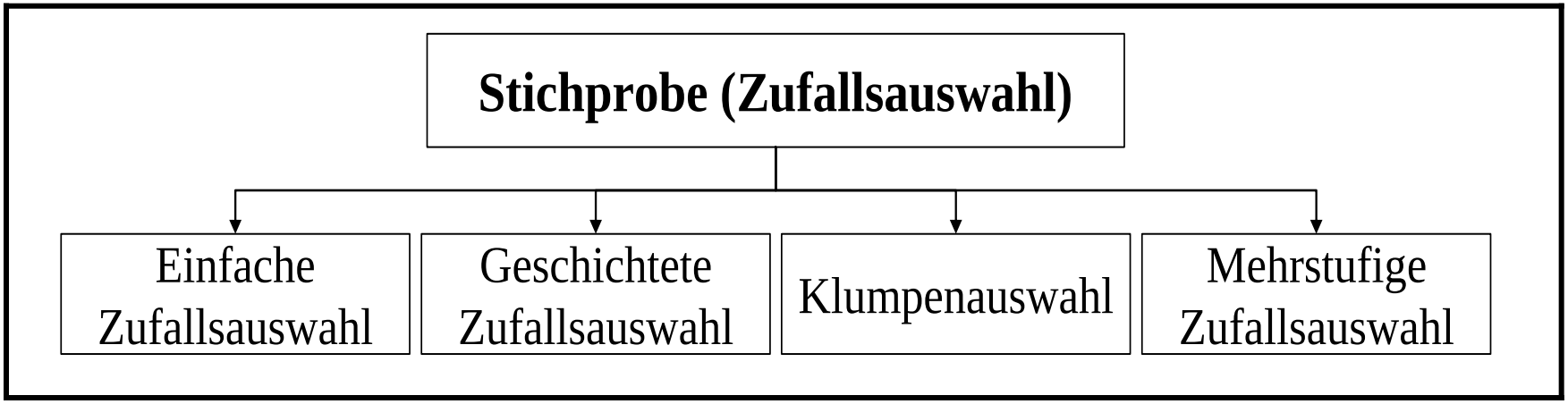
8.2 Zufallsauswahl

Bei einer **Zufallsauswahl** gelangt jede statistische Einheit der Grundgesamtheit mit einer **berechenbaren Wahrscheinlichkeit** in die **Stichprobe**. Über den **Zufallsmechanismus** wird die **Repräsentativität** sichergestellt.

Unter Anwendung des Wahrscheinlichkeitskalküls bei der Auswahl der statistischen Einheiten können bei einer Zufallsauswahl Aussagen über die Genauigkeit der Ergebnisse gemacht werden. Der induktive Schluss ist daher stets mit einem bestimmten Grad an Sicherheit verbunden.

Stichprobenfehler und **Sicherheitsgrad** lassen sich bei Zufallsstichproben mit Hilfe der Wahrscheinlichkeitsrechnung **quantifizieren**.

Abbildung: Verfahren der Zufallsauswahl



● Einfache Zufallsauswahl

Bei der **einfachen Zufallsauswahl** gelangt jede statistische Einheit der Grundgesamtheit mit **gleicher Wahrscheinlichkeit** in die Stichprobe.

Die Durchführung der einfachen Zufallsauswahl setzt im Allgemeinen voraus, dass alle **N Elemente der Grundgesamtheit** nummeriert in Form einer Liste oder Datei vorliegen. Zu unterscheiden sind verschiedene **Techniken zur Auswahl der n Stichprobenelemente**:

1. Auswahl mit (Pseudo-)Zufallszahlen

Die Elemente der Grundgesamtheit werden durchnummeriert. Die zu ziehenden Untersuchungseinheiten entnimmt man einer **Zufallszahlentabelle** oder einem **Zufallszahlengenerator**, der beispielsweise in ein Computerprogramm integriert ist.

Die **Zufallszahlen** werden in beiden Fällen über einen **Algorithmus mit** einer **langen Periode** erzeugt. Aus diesem Grund spricht man auch von **Pseudo-Zufallszahlen**. Obwohl Pseudo-Zufallszahlen durch eine Rechenvorschrift ermittelt werden, lassen sie sich nicht von **"echten" Zufallszahlen** (z.B. Urnenauswahl) unterscheiden.

2. Systematische Zufallsauswahl

a) Periodische Auswahl:

Per Zufall wird aus den ersten N/n Untersuchungseinheiten ein Startpunkt ausgewählt und anschließend jedes N/n-te Element gezogen.

b) Schlussziffernverfahren:

Entnommen werden alle Elemente, deren Schlussziffer in einer numerischen Liste einen durch einen Zufallsmechanismus bestimmten Wert aufweist.

Beispiel 8.1:

Bei der Vorgabe der Schlussziffer 6 erreicht man z. B., dass exakt 10 %, bei der Vorgabe der Schlussziffer 6 und zusätzlich den Schlussziffernkombinationen 58, 11, 35 und 99, dass exakt 14 % der Elemente aus der Grundgesamtheit in die Stichprobe gelangen. Mit dem Schlussziffernverfahren lässt sich also jeder angestrebte Auswahlsatz genau erreichen. ♦

Die **Prinzipien des induktiven Schlusses** lassen sich besonders klar im Falle einer einfachen Zufallsauswahl verdeutlichen, auf die wir uns im Folgenden konzentrieren werden. Eine **einfache Zufallsstichprobe** liegt unter folgenden Voraussetzungen vor:

- Alle Elemente der Grundgesamtheit haben die **gleiche Chance (Wahrscheinlichkeit)**, in die Stichprobe zu gelangen (=**uneingeschränkte Zufallstichprobe**).

- Die **Ziehungen** erfolgen **unabhängig** voneinander. Die Wahrscheinlichkeit, ausgewählt zu werden, ist dann bei jeder Ziehung gleich (**Ziehung mit Zurücklegen** <u>oder</u> **Ziehung aus einer unendlichen Grundgesamtheit**).

Approximativ liegen **unabhängige Ziehungen** vor, wenn aus einer **endlichen Grundgesamtheit ohne Zurücklegen** gezogen wird, der **Auswahlsatz n/N** jedoch **klein** ist (Faustregel: $n/N \leq 0{,}05$).

Beispiel 8.2:

Ein Meinungsforschungsinstitut besitzt eine Datei mit allen Telefonnummern Deutschlands. Der Computer wählt zufällig eine Telefonnummer aus, der Haushalt mit dieser Telefonnummer wird anschließend befragt. Um zu gewährleisten, dass Haushalte nicht mehrmals befragt werden, wird jede Telefonnummer höchstens einmal gezogen. Da der Auswahlsatz geringer als 5% ist (aus 20 Millionen Telefonnummern werden 1000 ausgewählt; somit gilt $1000/20.000.000 < 0{,}05$) sind die Bedingungen einer einfachen Zufallsstichprobe approximativ erfüllt. ♦

8.3 Stichprobenvariablen und -funktionen

● Stichprobenvariablen

Vor der Erhebung ist nicht bekannt, welche n Merkmalsträger in die Stichprobe gelangen. Damit liegen keine Informationen über die Ausprägungen beim **Merkmal X** vor. Somit sind **n Zufallsvariablen** vorhanden, die mit $\mathbf{X_i, i=1,2,\ldots,n}$ symbolisiert werden und als **Stichprobenvariablen** bezeichnet werden. Die Stichprobenvariable X_i bezeichnet also die potenzielle Beobachtung der Zufallsvariablen X, die bei der i-ten Durchführung des Zufallsvorgangs gemacht wird.

Nach der Erhebung liegen konkrete Realisationen der Stichprobenvariable X_i vor, die wir mit x_i bezeichen.

Zufallsstichprobe: n-dimensionale Zufallsvariable $(X_1, X_2, \ldots, X_n)$

Konkrete Stichprobe: Realisation der n-dimensionalen ZV $(x_1, x_2, \ldots, x_n)$

Beispiel 8.3:

Es sollen zehn Studierende nach ihrem Einkommen (Merkmal X) befragt werden (n = 10). Damit liegen vor der Erhebung 10 Zufallsvariablen (=Stichprobenvariablen) vor. Es ist nämlich nicht bekannt, was für ein Einkommen beispielsweise der erste Befragte haben wird:

$$X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9 \text{ und } X_{10}.$$

Nach der Erhebung sind konkrete Realisationen vorhanden. Gibt der erste Student ein Einkommen von 449€ an, so ist x_1=449. Insgesamt mögen sich bei der Befragung von 10 Studierenden folgende Stichprobenrealisationen ergeben haben:

$$x_1 = 449,\ x_2 = 662,\ x_3 = 389,\ x_4 = 802,\ x_5 = 523,$$

$$x_6 = 339,\ x_7 = 540,\ x_8 = 603,\ x_9 = 744,\ x_{10} = 459.$$ ♦

Die bei einer **einfachen Zufallsstichprobe** geltende **Unabhängigkeit** (bzw. Approximative Unabhängigkeit) **der Ziehungen** bedeutet stets, dass die **Stichprobenvariablen X_i unabhängig** (bzw. approximativ unabhängig) voneinander sind.

Eine **identische Verteilung** bedeutet insbesondere, dass die **Stichprobenvariablen X_i** den **gleichen Erwartungswert** und die **gleiche Varianz** haben:

(8.1) $E(X_i) = \mu$ für alle i

und

(8.2) $V(X_i) = \sigma^2$ für alle i.

Beispiel 8.4:

Das Durchschnittseinkommen aller Kasseler Studierenden (=Grundgesamtheit) soll 604€ bei einer Varianz von 6400€^{2} betragen. Wird eine Stichprobe von n=100 Kasseler Studenten gezogen, dann beträgt der Erwartungswert des Einkommens bei jedem der befragten Studenten 604€. Gleichzeitig ist bei jedem der befragten Studenten im Mittel mit einer Abweichung von 80€ (=$\sigma=\sqrt{6400}$) von diesem Erwartungswert zu rechnen.

Die Göttinger oder Marburger Studenten bilden dagegen andere Grundgesamtheiten, die durch völlig unterschiedliche Parameter μ und σ^2 gekennzeichnet sein können. ♦

- **Stichprobenfunktionen**

Aus einer realisierten Stichprobe (x_1, x_2, ... ,x_n) lassen sich nun **empirische Kennzahlen** berechnen, die als **Schätzwerte** für die unbekannten Parameter der Grundgesamtheit **oder Teststatistiken** verwendet werden können. Sie werden allgemein als **Stichprobenfunktonen** bezeichnet, die die in der Stichprobe enthaltenen Informationen mit Hilfe geeigneter Funktionen g verdichten.

Vor dem Ziehen der Stichprobe sind die **Stichprobenfunktionen** selbst **Zufallsvariablen**,

$$g(X_1,X_2,\ldots,X_n),$$

da sie je nachdem, welche Einheiten in die Zufallsauswahl gelangen, unterschiedliche Werte annehmen können. **Nach dem Ziehen der Stichprobe** liegt eine Realisation, d.h. ein **konkreter Wert einer Stichprobenfunktion** vor:

$$g(x_1,x_2,\ldots,x_n).$$

Beispiel 8.5:

Stichprobenfunktionen sind z.B. das arithmetische Mittel und unterschiedliche Streuungsmaße

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i\,, \qquad \rightarrow \qquad \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 \qquad \rightarrow \qquad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2$$

$$S^{*2} = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 \qquad \rightarrow \qquad s^{*2} = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2,$$

$$D^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \mu\right)^2 \qquad \rightarrow \qquad d^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \mu\right)^2$$

der Anteilswert

$$\overline{P}_n = Y_n / n = \sum_{i=1}^{n} \underbrace{X_i}_{0,1-\text{Var.}} / n \qquad \rightarrow \qquad \overline{p}_n = y_n / n = \sum_{i=1}^{n} \underbrace{x_i}_{0,1-\text{Werte}} / n$$

oder die Teststatistik .

$$Z = (\overline{X} - \mu)/(\sigma/\sqrt{n}) \qquad \rightarrow \qquad z = (\overline{x} - \mu)/(\sigma/\sqrt{n})$$

♦

8.4 Stichprobenverteilung

Unter einer **Stichprobenverteilung** versteht man die Wahrscheinlichkeitsverteilung einer Stichprobenfunktion $g(X_l, X_2, \ldots, X_n)$ aus allen möglichen Stichproben vom Umfang n. Die Stichprobenverteilung lässt sich bei kleinem Stichprobenumfang kombinatorisch (tabellarisch) durch Auswerten aller möglichen Stichproben bestimmen.

Erwartungswert und Varianz der Stichprobenverteilung von $\overline{X}$

Einfache Zufallsstichprobe (Auswahl mit Zurücklegen)	Uneingeschränkte Zufallsstichprobe (Auswahl ohne Zurücklegen)
$E(\overline{X}) = \mu$	$E(\overline{X}) = \mu$
$V(\overline{X}) = \frac{\sigma^2}{n}$	$V(\overline{X}) = \frac{\sigma^2}{n} \cdot \frac{N-n}{N-1}$

N: Umfang der Grundgesamtheit, n: Stichprobenumfang

Bei großem Stichprobenumfang sind für gebräuchliche Stichprobenfunktionen die Stichprobenverteilungen approximativ bekannt. Insbesondere sind die beiden Stichprobenfunktionen $\overline{X}$ und $\overline{P}$ aufgrund des zentralen Grenzwertsatzes approximativ normalverteilt.

Beispiel 8.6:

Aus der Patentstatistik geht die Anzahl der Patentmeldungen der 5 Unternehmen einer High-Tech-Branche hervor:

Unternehmen	A	B	C	D	E
Anzahl der Patentmeldungen	2	4	2	6	1

a) Geben Sie den Erwartungswert und die Varianz der Patentmeldungen für die Grundgesamtheit der Unternehmen der High-Tech-Branche an!

$$E(X) = \mu = \sum_{j=1}^{4} x_j \cdot p_j = 1 \cdot \frac{1}{5} + 2 \cdot \frac{2}{5} + 4 \cdot \frac{1}{5} + 6 \cdot \frac{1}{5}$$

$$= \frac{1}{5} + \frac{4}{5} + \frac{4}{5} + \frac{6}{5} = \frac{15}{5} = 3$$

$$V(X) = \sigma^2 = \sum_{j=1}^{4} x_j^2 \cdot p_j - [E(X)]^2$$

$$= 1^2 \cdot \frac{1}{5} + 2^2 \cdot \frac{2}{5} + 4^2 \cdot \frac{1}{5} + 6^2 \cdot \frac{1}{5} - 3^2$$

$$= \frac{1}{5} + \frac{8}{5} + \frac{16}{5} + \frac{36}{5} - 9 = \frac{61}{5} - 9 = 12{,}2 - 9 = 3{,}2$$

b) Bestimmen Sie die Stichprobenverteilung des arithmetischen Mittels $\overline{X}$ der Anzahl der Patentmeldungen bei uneingeschränkten Zufallsstichproben (Ziehen ohne Zurücklegen) des Umfangs 2!

Stichprobenverteilung von $\overline{X}$ für n=2

$\overline{X}$	Stichproben $\binom{N}{n} = \binom{5}{2} = 10$	$P(\overline{X} = \overline{x})$
1,5	$\{1, 2_1\}, \{1, 2_2\}$	2/10 = 0,20
2	$\{2_1, 2_2\}$	1/10 = 0,10
2,5	$\{1, 4\}$	1/10 = 0,10
3	$\{2_1, 4\}, \{2_2, 4\}$	2/10 = 0,20
3,5	$\{1, 6\}$	1/10 = 0,10
4	$\{2_1, 6\}, \{2_2, 6\}$	2/10 = 0,20
5	$\{4, 6\}$	1/10 = 0,10

c) Berechnen Sie den Erwartungswert der Stichprobenverteilung von $\overline{X}$ für n=2!

$$E(\overline{X}) = \sum_{\ell=1}^{7} \overline{x}_\ell \cdot P(\overline{X}_\ell = \overline{x}_\ell)$$

$$= 1{,}5 \cdot 0{,}2 + 2 \cdot 0{,}1 + 2{,}5 \cdot 0{,}1 + 3 \cdot 0{,}2 + 3{,}5 \cdot 0{,}1 + 4 \cdot 0{,}2 + 5 \cdot 0{,}1$$

$$= 0{,}3 + 0{,}2 + 0{,}25 + 0{,}6 + 0{,}35 + 0{,}8 + 0{,}5 = 3$$

d) Berechnen Sie die Varianz der Stichprobenverteilung von $\overline{X}$ für n=2

- unter Verwendung der in Teil b) ermittelten Wahrscheinlichkeitsverteilung,
- unter Verwendung der in Teil a) ermittelten Varianz der Grundgesamtheit!

Varianz der Stichprobenverteilung	
mit der Wahrscheinlichkeitsvertei-lung aus Teil b)	mit der Varianz σ^2 der Grund-gesamtheit aus Teil a)
$V(\overline{X}) = \sum_{\ell=1}^{7} \overline{x}_\ell^2 \cdot P(\overline{X}_\ell = \overline{x}_\ell) - [E(\overline{X})]^2$ $= 1{,}5^2 \cdot 0{,}2 + 2^2 \cdot 0{,}1 + 2{,}5^2 \cdot 0{,}1 + 3^2 \cdot 0{,}2$ $+ 3{,}5^2 \cdot 0{,}1 + 4^2 \cdot 0{,}2 + 5^2 \cdot 0{,}1 - 3^2$ $= 0{,}45 + 0{,}4 + 0{,}625 + 1{,}8 + 1{,}225 + 3{,}2 + 2{,}5$ $- 9 = 10{,}2 - 9 = 1{,}2$	$V(\overline{X}) = \frac{\sigma^2}{n} \cdot \frac{N-n}{N-1}$ $= \frac{3{,}2}{2} \cdot \frac{5-2}{5-1}$ $= 1{,}6 \cdot \frac{3}{4} = 1{,}2$

9. Punktschätzung

9.1 Schätzfunktionen und ihre Eigenschaften

Eine Stichprobenfunktion, die zur Schätzung eines unbekannten Parameters in der Grundgesamtheit eingesetzt wird, heißt **Schätzfunktion**. Sie liefert einen **Punktschätzer** für den unbekannten Parameter der Grundgesamtheit.

Unbekannter Parameter der Grundgesamtheit: $\boldsymbol{\theta}$

Schätzfunktion für $\boldsymbol{\theta}$**:** $\hat{\theta} = g(X_1, X_2, ..., X_n)$

Punktschätzung:

$$\hat{\theta} = g(X_1, X_2, ..., X_n) \rightarrow \boldsymbol{\theta}$$

Punktschätzer (Schätzwert): $\hat{\theta} = g(x_1, x_2, ..., x_n)$

Die Stichprobenfunktion $\hat{\theta}$ wird als Schätzfunktion bezeichnet, weil ihre Realisationen konkrete Schätzwerte $\hat{\theta}$ für den unbekannten Parameter $\boldsymbol{\theta}$ der Grundgesamtheit liefern. Man beachte, dass mit $\hat{\theta}$ zugleich eine Zufallsvariable als auch deren Realisation bezeichnet wird. Aus dem jeweiligen Kontext geht hervor, welche Interpretation für $\hat{\theta}$ gerade vorliegt.
In Abhängigkeit davon, welche Untersuchungseinheiten der Grundgesamtheit in die Stichprobe gelangen, werden wir unterschiedliche Punktschätzer erhalten. Ein Punktschätzer ist also immer mit einer Unsicherheit behaftet. In der Regel wird er nicht exakt mit dem unbekannten Parameter der Grundgesamtheit übereinstimmen.

Häufig stehen mehrere Schätzfunktionen zur Punktschätzung eines unbekannten Parameters zur Verfügung. Welche Schätzfunktion soll dann verwendet werden?

Beispiel 9.1:

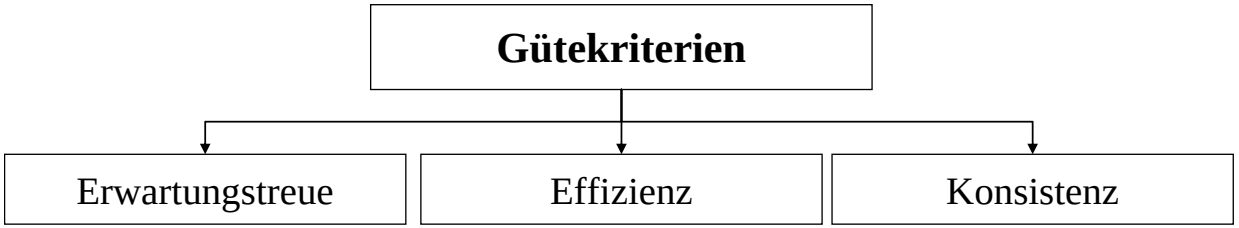
- Bei symmetrischen Verteilungen kann ein unbekannter Erwartungswert μ durch den Stichprobenmittelwert $\bar{x}$ und durch den Stichprobenmedian $\tilde{x}$ geschätzt werden. Soll $\bar{X}$ oder $\tilde{X}$ als Schätzfunktion zur Schätzung von μ verwendet werden?

- Bei einer Poisson-Verteilung stimmen Erwartungswert E(X) und Varianz V(X) überein. Sie werden durch den Parameter λ wiedergegeben. Soll der Parameter λ durch das Stichprobenmittel $\bar{x}$ oder die Stichprobenvarianz s^2 geschätzt werden?

- Warum sollte die unbekannte Varianz σ^2 einer Verteilung durch die Stichprobenvarianz mit dem Faktor 1/(n-1) statt mit dem Faktor 1/n geschätzt werden? ♦

Um die Frage zu beantworten, welcher Punktschätzer bei Schätzungen eines unbekannten Parameters einer Grundgesamtheit vorzuziehen ist, benötigt man Kriterien, nach denen die Güte von Schätzfunktionen beurteilt werden kann. Drei wichtige Gütekriterien werden im Folgenden

Abbildung: Gütekriterien für eine Schätzfunktion



- **Erwartungstreue**

Aufgrund des Stichprobenfehlers wird ein Punktschätzer $\hat{\theta}$ den unbekannten Parameter der Grundgesamtheit nicht genau treffen. Der Punktschätzer sollte aber im Durchschnitt mit dem unbekannten Parameter θ übereinstimmen. Diese Eigenschaft bezeichnet man als Erwartungstreue.

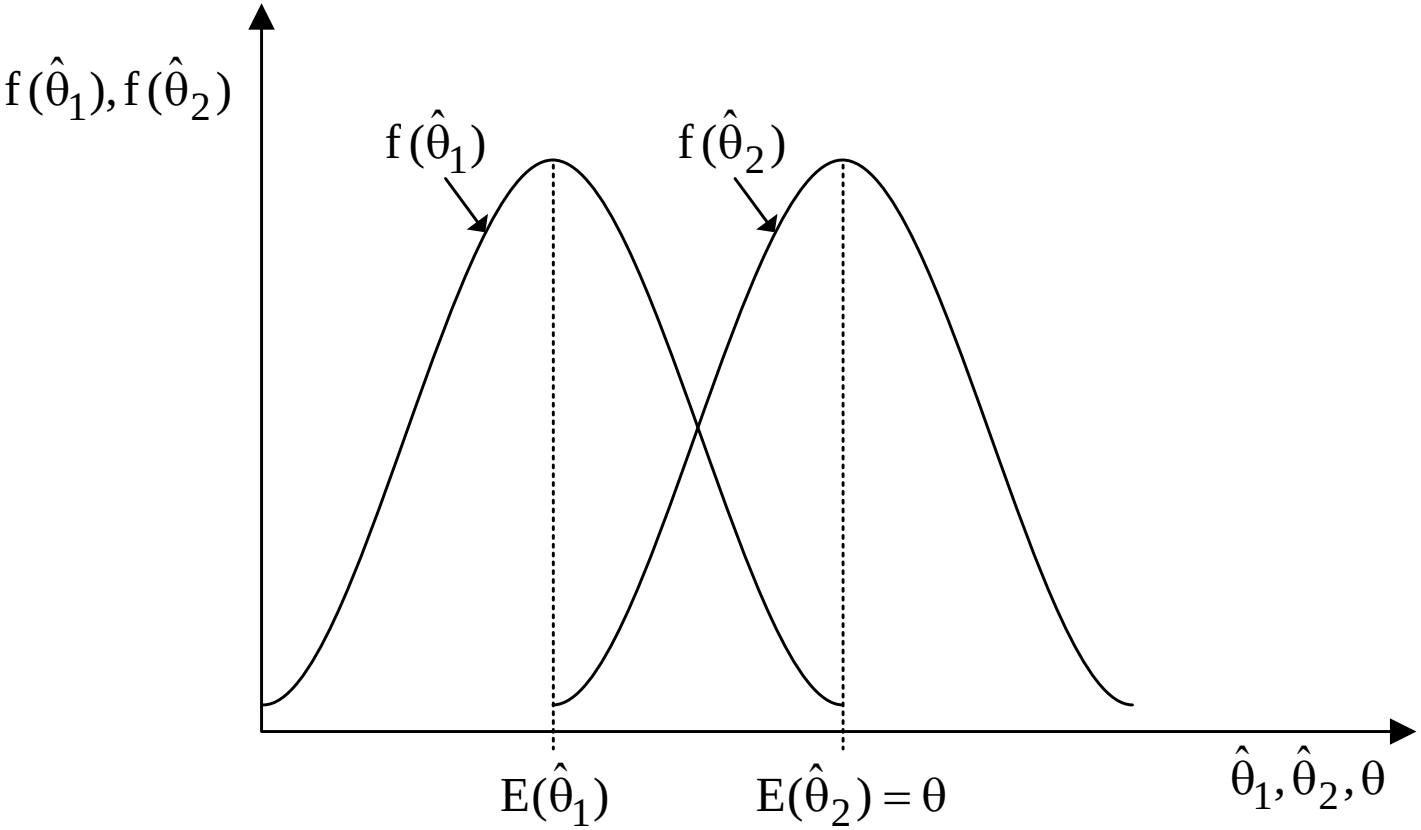
Eine Schätzfunktion $\hat{\theta} = g(X_1, \ldots, X_n)$ für den unbekannten Parameter der Grundgesamtheit ist **erwartungstreu (unverzerrt)**, falls

(9.1) $$E(\hat{\theta}) = \theta$$

gilt. Für $E(\hat{\theta}) - \theta \neq 0$ ist die Schätzfunktion **verzerrt**. Das Ausmaß der Verzerrung wird auch als **Bias** bezeichnet.

Bei einem erwartungstreuen Schätzer stimmt der Erwartungswert der Schätzfunktion mit dem Parameter der Grundgesamtheit überein. Der Bias ist dann gleich 0.

Abbildung: Erwartungstreuer und verzerrter Schätzer



$E(\hat{\theta}_1) \neq \theta \quad \rightarrow \quad \hat{\theta}_1$ ist kein erwartungstreuer Schätzer für θ

$E(\hat{\theta}_2) = \theta \quad \rightarrow \quad \hat{\theta}_2$ ist ein erwartungstreuer Schätzer für θ

Beispiel 9.1:

Gegeben ist eine beliebige Grundgesamtheit eines quantitativen Merkmals, das durch die Zufallsvariable X beschrieben wird. Der Erwartungswert μ der Verteilung von X ist unbekannt.

Wir wollen das arithmetische Mittel

$$\overline{X} = \frac{1}{n} \cdot \sum_{i=1}^{n} X_i$$

als Schätzer für den unbekannten Parameter μ verwenden und fragen, ob diese Schätzfunktion erwartungstreu ist.

Hierzu bilden wir den Erwartungswert von $\overline{X}$:

$$E(\overline{X}) = E\left(\frac{1}{n} \cdot \sum_{i=1}^{n} X_i\right) = \frac{1}{n} \cdot E\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n} \cdot \sum_{i=1}^{n} E(X_i) \quad .$$

Mit $E(X_i) = \mu$ für alle i=1,2,…,n erhalten wir

$$E(\overline{X}) = \frac{1}{n} \cdot \left(\underbrace{\mu + \mu + \ldots + \mu}_{n-\text{mal}}\right) = \frac{1}{n} \cdot n \cdot \mu = \mu \quad .$$

Damit ist gezeigt, dass $\overline{X}$ ein erwartungstreuer Schätzer für μ ist. ♦

Analog kann gezeigt werden, dass der **Stichprobenanteilswert**

$$\overline{P} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

mit X_i gleich 1, falls das Ereignis A vorliegt und 0 sonst, eine erwartungstreue Schätzfunktion für die **unbekannte Wahrscheinlichkeit** oder den **unbekannten Anteilswert p der Grundgesamtheit** ist:

$$E(\overline{P}) = p.$$

Aufwändiger ist der Nachweis, dass die Schätzfunktion

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$$

erwartungstreu ist:

$$E(S^2) = \sigma^2.$$

Die **Stichprobenvarianz** ist also dann ein **erwartungstreuer Schätzer** für die **unbekannte Varianz σ^2 der Grundgesamtheit**, wenn die Summe der Abweichungsquadrate mit dem **Faktor 1/(n-1)** multipliziert wird.

Wird dagegen die Stichprobenvarianz mit dem Faktor 1/n gebildet,

$$S^{*2} = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$$

ist sie nicht erwartungstreu. Der Bias beträgt hier

$$E\left(S^{*2}\right) - \sigma^2 = \frac{n-1}{n}\sigma^2 - \sigma^2 = -\frac{\sigma^2}{n} \quad .$$

● Effizienz

Die Erwartungstreue bezieht sich auf den Durchschnitt aller Stichproben eines bestimmten Umfangs. Werden mehrere Stichproben entnommen, dann werden sich die einzelnen Schätzwerte mehr oder weniger voneinander unterscheiden. Die **Effizienz** einer Schätzfunktion bezieht sich auf die **Streuung eines Schätzers**. Je geringer die Varianz der Schätzfunktion ist, desto präziser ist die Schätzung.

Die erwartungstreue Schätzfunktion mit der geringsten Varianz heißt **effizient**.

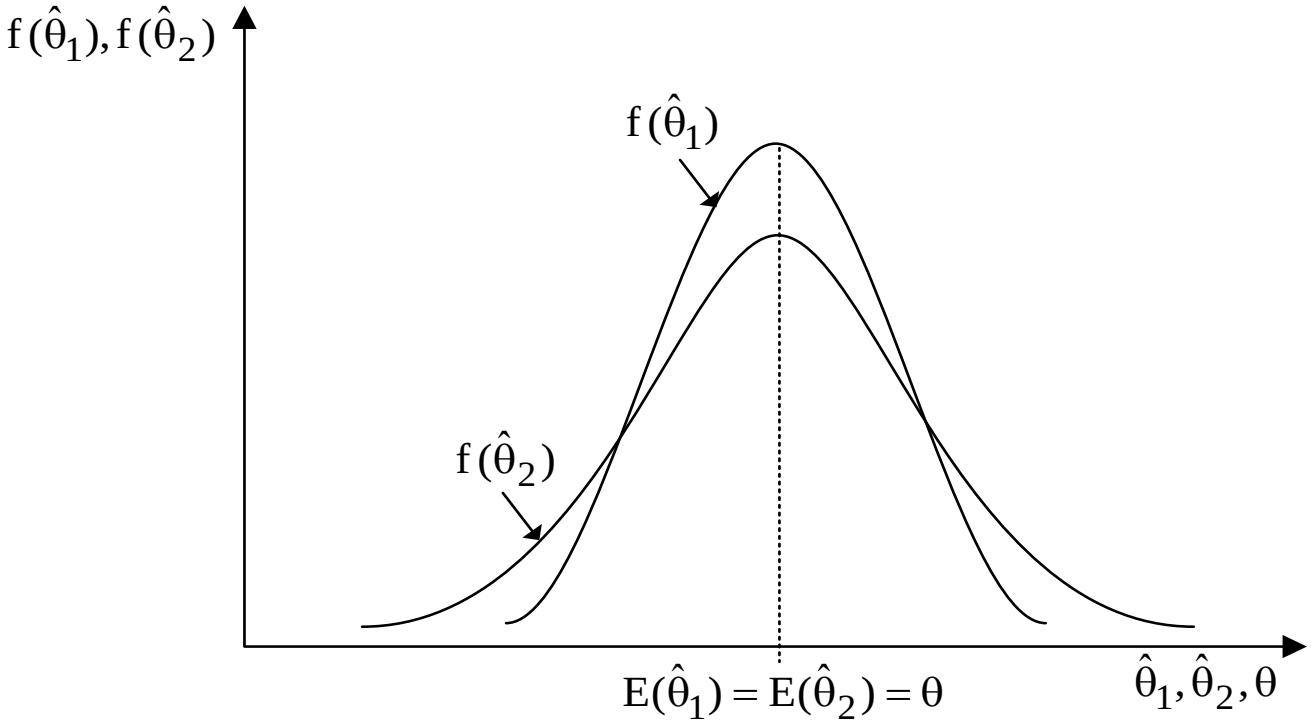
Aufschluss über die **relative Effizienz** zweier Schätzfunktionen erhält man durch einen Vergleich ihrer Varianzen. Es seien $\hat{\theta}_1 = g_1(X_1,X_2,...,X_n)$ und $\hat{\theta}_2 = g_2(X_1,X_2,...X_n)$ zwei erwartungstreue Schätzfunktionen für den unbekannten Parameter der Grundgesamtheit. Die Schätzfunktion $\hat{\theta}_1$ ist **relativ effizient** zu $\hat{\theta}_2$, wenn für jeden gegebenen Stichprobenumfang n

(9.2) $$V(\hat{\theta}_1) \leq V(\hat{\theta}_2) \quad .$$

gilt.

Das Gütekriterium der Effizienz besagt also, dass von zwei alternativen erwartungstreuen Schätzfunktionen diejenige zu wählen ist, die bei festem Stichprobenumfang n die kleinere Varianz aufweist.

Abbildung: Relative Effizienz eines Schätzers



$E(\hat{\theta}_1) = E(\hat{\theta}_2) = \theta$ und $V(\hat{\theta}_1) \leq V(\hat{\theta}_2)$

$\rightarrow$ $\hat{\theta}_1$ ist relativ effizient zu $\hat{\theta}_2$

Beispiel 9.2:

Zur Schätzung des unbekannten Erwartungswertes μ einer Grundgesamtheit werden die Schätzfunktionen

$$\hat{\mu}_1 = \frac{1}{n}\left(X_1 + X_2 + \ldots + X_n\right)$$

und

$$\hat{\mu}_2 = \frac{1}{n}\left(a_1 X_1 + a_2 X_2 + \ldots + a_n X_n\right)$$

mit $a_1 + a_2 + \ldots + a_n = n$

in Betracht gezogen. Beide Schätzfunktionen sind erwartungstreu. Die Erwartungstreue von $\hat{\mu}_1$ ist bereits in Beispiel 9.1 gezeigt worden, da es sich hierbei um das arithmetische Mittel $\overline{X}$ handelt. Die Erwartungstreue von $\hat{\mu}_2$ ergibt sich aus

$$E\left(\hat{\mu}_2\right) = \frac{1}{n}\left[a_1 E\left(X_1\right) + \ldots + a_n E\left(X_n\right)\right]$$

$$= \frac{1}{n}\left(a_1\mu + \ldots + a_n\mu\right) = \frac{1}{n}\mu\left(a_1 + \ldots + a_n\right) = \frac{1}{n}\mu \cdot n = \mu \quad .$$

Welche der beiden Schätzfunktionen hat aber die kleinere Varianz? Die Schätzfunktion mit der kleineren Varianz ist die relativ effizientere Schätzfunktion.

Als Varianz von $\hat{\mu}_1$ erhalten wir

$$V(\hat{\mu}_1) = V\left[\frac{1}{n}(X_1 + X_2 + \ldots + X_n)\right]$$

$$\underset{\substack{\uparrow \\ \text{Unabhäng.}}}{=} \frac{1}{n^2}[V(X_1) + V(X_2) + \ldots + V(X_n)]$$

$$= \frac{1}{n^2}\underbrace{(\sigma^2 + \sigma^2 + \ldots + \sigma^2)}_{n-\text{mal}} = \frac{1}{n^2} \cdot n \cdot \sigma^2 = \frac{\sigma^2}{n} \ .$$

Für die Varianz von $\hat{\mu}_2$ ergibt sich

$$V(\hat{\mu}_2) = \frac{1}{n^2}\left[a_1^2 V(X_1) + a_2^2 V(X_2) + \ldots + a_n^2 V(X_n)\right]$$

$$= \frac{1}{n^2}\left(a_1^2\sigma^2 + a_2^2\sigma^2 + \ldots + a_n^2\sigma^2\right) = \frac{1}{n^2}\sigma^2\left(a_1^2 + a_2^2 + \ldots + a_n^2\right).$$

Falls alle a_i gleich 1 sind, ist $\left(a_1^2 + \ldots + a_n^2\right) = n$, so dass die Varianzen von $\hat{\mu}_1$ und $\hat{\mu}_2$ identisch sind. Sofern nur ein a_i ungleich 1 ist, nimmt die Summe der Größen a_i^2 einen Wert größer als n an. Die Varianz von $\hat{\mu}_2$ ist dann größer als die Varianz von $\hat{\mu}_1$. Allgemein gilt damit

$$V(\hat{\mu}_1) \leq V(\hat{\mu}_2),$$

was bedeutet, dass der Schätzer $\hat{\mu}_1$ relativ effizient zu $\hat{\mu}_2$ ist. ♦

● Konsistenz

Eine Reihe von Schätzfunktionen, die in der statistischen Praxis auftreten, haben die Eigenschaft, dass ihr Erwartungswert bei steigendem Stichprobenumfang n gegen den unbekannten Parameter θ strebt (**asymptotische Erwartungstreue**),

(9.3) $$\lim_{n\to\infty} E\left(\hat{\theta}\right) = \theta \ ,$$

und auch ihre Varianz verschwindet, d.h. gegen 0 geht:

(9.4) $$\lim_{n\to\infty} V\left(\hat{\theta}\right) = 0 \ .$$

Man bezeichnet diese asymptotische Güteeigenschaft einer Schätzfunktion als **Konsistenz**.

Ein **konsistenter Schätzer**, der die Bedingungen (9.3) und (9.4) erfüllt, fällt bei einem über alle Grenzen steigenden Stichprobenumfang mit dem unbekannten Parameter θ zusammen. Die Konsistenz einer Schätzfunktion ist also die Garantie, dass sich ein steigender Aufwand lohnt, um im Durchschnitt bessere Schätzergebnisse zu erhalten. **Konsistenz** wird in der Praxis häufig als **minimale Güteanforderung** angesehen, die an eine Schätzfunktion zu stellen ist.

Beispiel 9.3:

Wir wollen prüfen, ob die Schätzfunktion

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

eine konsistente Schätzung für den unbekannten Erwartungswert μ einer Grundgesamtheit ist.

1. Asymptotische Erwartungstreue

   In Beispiel 9.1 haben wir gezeigt, dass $\overline{X}$ erwartungstreu ist. Da diese Eigenschaft für jeden beliebigen Stichprobenumfang gilt, ist sie auch für $n \to \infty$ erfüllt:

$$\lim_{n\to\infty} E\left(\overline{X}_n\right) = \mu$$

2. Verschwindende Varianz

   Mit $\mathrm{Var}\left(\overline{X}_n\right) = \sigma^2 / n$ erhalten wir

$$\lim_{n\to\infty} V\left(\overline{X}\right) = \lim_{n\to\infty} \frac{1}{n}\sigma^2 = \lim_{n\to\infty} \frac{1}{n} \cdot \lim_{n\to\infty} \sigma^2 = 0 \cdot \sigma^2 = 0 \ .$$

Das Stichprobenmittel $\overline{X}$ ist somit ein konsistenter Schätzer für den unbekannten Mittelwert (Erwartungswert) μ der Grundgesamtheit. Bei steigendem Stichprobenumfang n konzentrieren sich also die Realisationen von $\overline{X}_n$ immer enger um μ, da die Streuung geringer wird. ♦

Beispiel 9.4:

Zehn Kasseler Studierende sind nach ihrem Einkommen (Merkmal X) befragt worden. Aufgrund einer einfachen Zufallsstichprobe haben sich folgende Stichprobenrealisationen ergeben:

$$x_1 = 449,\ x_2 = 662,\ x_3 = 389,\ x_4 = 802,\ x_5 = 523,$$
$$x_6 = 339,\ x_7 = 540,\ x_8 = 603,\ x_9 = 744 \text{ und } x_{10} = 459.$$

Aus den Stichprobendaten erhält man das arithmetische Mittel

$$\begin{aligned}\bar{x} &= \frac{1}{10} \cdot \sum_{i=1}^{10} x_i \\ &= \frac{1}{10} \cdot (449 + 662 + 389 + 802 + 523 + 339 + 540 + 603 + 744 + 459) \\ &= \frac{1}{10} \cdot 5.510 = 551\ [€].\end{aligned}$$

Der Schätzwert für das Durchschnittseinkommen aller Kasseler Studenten beträgt also 551 €. Da wir nur 10 Merkmalswerte verwendet haben, wird das arithmetische Mittel aller Studenten, μ, von dem Punktschätzer abweichen. Der Stichprobenmittelwert $\overline{X}$ in Höhe von 551 € ist jedoch eine **erwartungstreue (unverzerrte) und konsistente Schätzung** des Durchschnittseinkommens aller Kasseler Studierenden.

Man weiß, dass es bei einer einfachen Zufallsstichprobe keine erwartungstreue Schätzfunktion mit einer geringeren Varianz als die Varianz von $\overline{X}$ gibt. Aus diesem Grund ist $\overline{X}$ ein **effizienter Schätzer**.

Die Varianz von $\overline{X}$, $\sigma^2_{\overline{x}} = \sigma^2 / n$, kann durch $\hat{\sigma}^2_{\overline{x}} = s^2 / n$ geschätzt werden. Mit der erwartungstreuen Stichprobenvarianz

$$s^2 = \frac{1}{10-1} \cdot \left[(449-551)^2 + (662-551)^2 + \ldots + (459-551)^2 \right]$$

$$= \frac{1}{9} \cdot 206.236 = 22.915{,}11\ [€^2],$$

für die unbekannte Varianz σ^2 der Einkommen aller Kasseler Studenten erhält man

$$\hat{\sigma}^2_{\overline{x}} = \frac{s^2}{n} = \frac{22.915{,}11}{10} = 2291{,}51 [€^2].$$

Der Standardfehler der Schätzung (=Standardabweichung von $\overline{X}$) beträgt

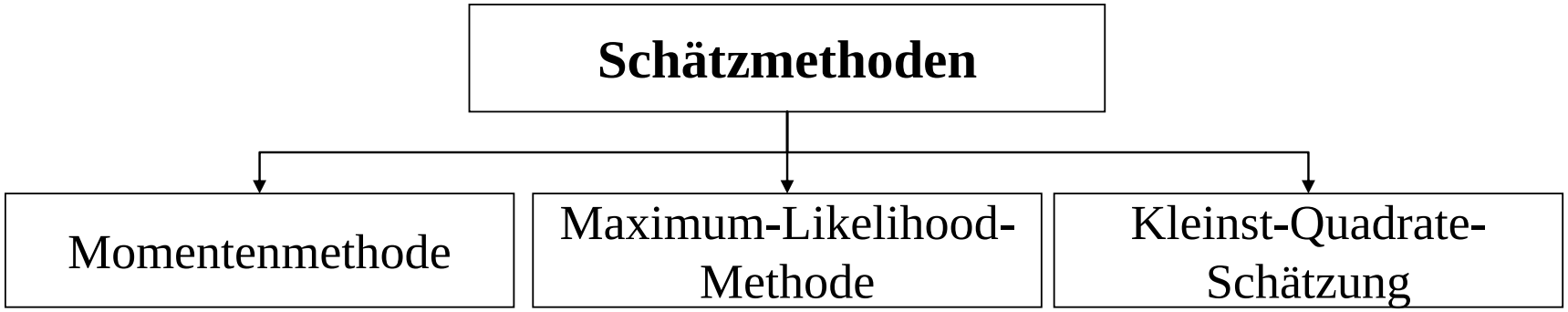
$$\hat{\sigma}_{\overline{x}} = \sqrt{\frac{s^2}{n}} = \sqrt{2291{,}51} = 47{,}87\ [€].$$

Die **Effizienz** von $\overline{X}$ zeigt sich darin, dass kein anderer erwartungstreuer Schätzer für μ eine größere Präzision (Genauigkeit), d.h. einen kleineren Standardfehler besitzt. ♦

9.2 Methoden der Punktschätzung

Methoden der Punktschätzung sind Verfahren zur Konstruktion von Schätzfunktionen für die unbekannten Parameter der Grundgesamtheit. Bisher wurde aufgezeigt, welche Eigenschaften Punktschätzer aufweisen sollen. Wie kann man für einen Parameter der Grundgesamtheit aber eine Schätzfunktion entwickeln?

Abbildung: Wichtige Schätzmethoden



Im Folgenden stellen wir die Momentenmethode und die Methode der kleinsten Quadrate als Punktschätzmethoden vor.

• Momentenmethode

Die Momente stellen eine Verallgemeinerung des arithmetischen Mittels dar. Da die gewöhnlichen Momente,

(9.5) $\mu_r = E(X^r), \quad r=1,2,\ldots,$

in der Grundgesamtheit unbekannt sind, werden sie durch die vergleichbaren empirischen Momente der Stichprobe,

(9.6) $$\overline{X}_r = \frac{1}{n}\sum_{i=1}^{n} X_i^r, r = 1,2,\ldots,$$

ersetzt.

Die **Momentenmethode** verwendet die empirischen Momente einer Stichprobe, um unbekannte Parameter der Grundgesamtheit zu schätzen.

Beispiel 9.6:

Der unbekannte Erwartungswert μ ist gleich dem ersten gewöhnlichen Moment μ_1 der Verteilung von X in der Grundgesamtheit. Nach der Momentenmethode ist die Schätzfunktion durch das erste empirische Moment der Stichprobe, also durch

$$\overline{X}_1 = \frac{1}{n}\sum_{i=1}^{n} X_i^1 = \frac{1}{n}\sum_{i=1}^{n} X_i$$

d.h. durch das arithmetische Mittel $\overline{X}$ gegeben. Wie wir gesehen haben, ist $\overline{X}$ ein erwartungstreuer, effizienter und konsistenter Schätzer für μ. ♦

Beispiel 9.7:

Auch die unbekannte Varianz der Grundgesamtheit σ^2 kann nach dem Varianzverschiebungssatz unter Verwendung der ersten beiden Momente dargestellt werden:

$$\sigma^2 = \mu_2 - \mu_1^2 = E\left(X^2\right) - [E(X)]^2 \cdot$$

Einen Momentenschätzer erhält man, indem man die unbekannten Momente der Grundgesamtheit durch die empirischen Momente der Stichprobe ersetzt:

$$\overline{X}_2 - \overline{X}_1^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} X_i^1\right)^2$$

Nach einigen Umformungen ergibt sich daraus

$$\overline{X}_2 - \overline{X}_1^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2,$$

d.h. die nicht-erwartungstreue Stichprobenvarianz S^{*2}. Die Momentenmethode liefert also nicht grundsätzlich erwartungstreue Schätzfunktionen. ♦

• Methode der kleinsten Quadrate

Die Methode der kleinsten Quadrate wurde bereits in der Deskriptiven Statistik zur Schätzung der Regressionskoeffizienten herangezogen. Der unbekannte Parameter der Grundgesamtheit wird hierbei so bestimmt, dass die quadrierten Abweichungen der Beobachtungswerte von ihrem Schätzwert minimiert werden.

Mit der **Kleinst-Quadrate-Schätzung** ermittelt man den Schätzer so, dass die Summe der quadrierten Abweichungen von dem Parameterschätzer minimiert wird. Aufgrund der Quadrierung werden absolut große Abweichungen überproportional berücksichtigt.

Beispiel 9.8:

Ausgegangen wird von den n Zufallsvariablen einer Stichprobe,

$$X_1, X_2, \ldots, X_n,$$

die zur Schätzung des arithmetischen Mittels der Grundgesamtheit μ herangezogen werden. Der Kleinst-Quadrate-Schätzer $\hat{\mu}$ für den unbekannten Mittelwert der Grundgesamtheit ergibt sich durch Minimierung der quadrierten Abweichungsquadratsumme:

$$Q(\mu) = \sum_{i=1}^{n} (X_i - \mu)^2 \rightarrow \underset{\mu}{\mathrm{Min}} \; .$$

Leitet man $Q(\mu)$ nach μ ab,

$$\frac{dQ(\mu)}{d\mu} = \frac{d\left[\sum_{i=1}^{n}(X_i-\mu)^2\right]}{d\mu} = \sum_{i=1}^{n}\frac{d(X_i-\mu)^2}{d\mu}$$

erhält man unter Anwendung der Kettenregel den Ausdruck

$$\frac{dQ(\mu)}{d\mu} = \sum_{i=1}^{n}\underbrace{2\cdot(X_i-\mu)}_{\text{äußere Ableitung}}\cdot\underbrace{(-1)}_{\text{innere Ableitung}}.$$

Zur Bestimmung des Kleinst-Quadrate-Schätzers für μ muss die erste Ableitung gleich null gesetzt werden. Diese Beziehung muss der Kleinst-Quadrate-Schätzer $\hat{\mu}$ erfüllen:

$$\sum_{i=1}^{n}2\cdot(X_i-\hat{\mu})\cdot(-1)=0.$$

Aus

$$\sum_{i=1}^{n}(X_i-\hat{\mu})=0, \qquad \sum_{i=1}^{n}X_i-\sum_{i=1}^{n}\hat{\mu}=0 \quad \text{und} \quad n\cdot\hat{\mu}=\sum_{i=1}^{n}X_i$$

erhält man den Kleinst-Quadrate-Schätzer

$$\hat{\mu}=\frac{1}{n}\cdot\sum_{i=1}^{n}X_i=\overline{X}$$

für den unbekannten Mittelwert μ der Grundgesamtheit, der dem arithmetischen Mittel der Stichprobe entspricht. ♦