26 Statistische Schätztheorie

Die *Schätztheorie* gehört neben der *Testtheorie* (siehe Kap.27–29) zu wichtigen Gebieten der *mathematischen Statistik*. Ihre Aufgabe besteht darin, aufgrund von *Stichproben* Methoden zur Ermittlung von *Schätzungen* für unbekannte *Parameter* und *Verteilungsfunktionen* einer betrachteten Grundgesamtheit (*Zufallsgröße* X) anzugeben.

☞

Im Rahmen des vorliegenden Buches beschränken wir uns auf *Schätzungen unbekannter Parameter* für eine Zufallsgröße X. Schätzungen für unbekannte Verteilungsfunktionen lassen sich unter Verwendung empirischer Verteilungsfunktionen durchführen (siehe Abschn.23.3).

♦

26.1 Einführung

Bei statistischen Untersuchungen können folgende zwei Fälle auftreten:
Man kennt für eine vorliegende *Grundgesamtheit* (siehe Kap.23), deren betrachtetes Merkmal durch eine *Zufallsgröße* X beschrieben wird,

* weder die *Verteilungsfunktion* (Wahrscheinlichkeitsverteilung) noch deren *Parameter* (Erwartungswert, Varianz/Streuung,...).

* die *Verteilungsfunktion* (Wahrscheinlichkeitsverteilung), aber nicht deren *Parameter* (Erwartungswert, Varianz/Streuung,...).

Der letzte Fall tritt bei einer Reihe praktischer Untersuchungen auf, in denen man die *Verteilungsfunktion* aufgrund des zentralen Grenzwertsatzes (siehe Abschn.20.3) bzw. der Eigenschaften bekannter Verteilungsfunktionen (siehe Kap.18) näherungsweise kennt. Illustrieren wir die Problematik an einfachen Beispielen.

Beispiel 26.1:

a) Betrachten wir die Aufgabe aus Beispiel 18.2b. Dort haben wir angenommen, daß bei der Produktion von Bolzen deren *Länge* (*Zufallsgröße* X) *normalverteilt* ist. Die *Normalverteilung* kann man hier aufgrund des zentralen Grenzwertsatzes näherungsweise voraussetzen, da sich bei der Produktion der Bolzen eine Reihe unabhängiger zufälliger Effekte überlagert. Des weiteren haben wir im Beispiel 18.2b den *Erwartungswert*

μ=50 mm (Sollwert) und die *Standardabweichung* σ=0.2 mm angenommen, d.h. eine N(50,0.2)-*Normalverteilung*. Ein Bolzen kann in diesem Beispiel nicht mehr verwendet werden (d.h. ist defekt), wenn seine Länge um mehr als 0.25 mm vom Sollwert (Erwartungswert) 50 mm abweicht.

Hier liegt der *Idealfall* vor, daß

* *Verteilungsfunktion* (*Normalverteilung*)
* *Erwartungswert*
* *Standardabweichung*

der betrachteten Zufallsgröße X *bekannt* sind.

a1)Es kann jedoch der Fall auftreten daß nur der *Erwartungswert* μ=50 mm als *Sollwert* bekannt ist, während man die *Standardabweichung* σ nicht kennt. Für diesen unbekannten Parameter σ lassen sich *Schätzwerte* anhand entnommener *Stichproben* berechnen.
Wenn man in der betrachteten *Grundgesamtheit* der *Bolzen* das *Merkmal* der *Länge* (in mm) als *Zufallsgröße* X verwendet, kann man mit einem aus einer Stichprobe berechneten *Schätzwert* für die *Standardabweichung* σ mittels der *Verteilungsfunktion* der *Normalverteilung* N(50,σ) bzw. der *standardisierten Normalverteilung* Φ(x) z.B. die *Wahrscheinlichkeit* dafür berechnen, daß ein der Produktion entnommener *Bolzen defekt* ist:

$$P(|X-50|>0.25) \;=\; 1 \;-\; P(|X-50|\le 0.25)$$

$$= \; 1 - P(50-0.25 \le X \le 50+0.25) = 1 - P(50.25) + P(49.75)$$

$$= 1 \;-\; P\left(\frac{50-0.25-50}{\sigma} \le \frac{X-50}{\sigma} \le \frac{50+0.25-50}{\sigma}\right)$$

$$= 1-(\Phi(0.25/\sigma) \;-\; \Phi(-0.25/\sigma)) = 2\cdot(\,1-\Phi(0.25/\sigma)\,)$$

Einen *Schätzwert* für die unbekannte *Standardabweichung* σ der *Zufallsgröße* X erhält man beispielsweise, wenn man für eine aus der Grundgesamtheit entnommene *Stichprobe* die *Stichprobenstandardabweichung* s_X berechnet (siehe Abschn.24.5.1).

a2)Häufiger tritt bei dem betrachteten Beispiel der Fall auf, daß der *Erwartungswert* μ *unbekannt* ist, während man die *Standardabweichung* σ kennt, da die zur Produktion der Bolzen verwandten Maschinen Toleranzgrenzen haben. In diesem Fall ist man an Schätzwerten für den unbekannten Erwartungswert interessiert, die z.B. aus entnommenen Stichproben durch Berechnung des *Stichprobenmittels* gewonnen werden.

b) Die Anzahl der beim Zerfall einer radioaktiven Substanz pro Zeiteinheit zerfallenen Atome läßt sich durch eine *Zufallsgröße* X beschreiben, die der *Poisson-Verteilung* genügt. Diese diskrete Verteilung ist durch die *Wahrscheinlichkeiten*

$$P(X=k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda} \qquad (k = 0, 1, 2, \ldots)$$

vollständig charakterisiert, wenn man den Parameter λ (*Erwartungswert*) kennt (siehe Abschn.18.1). Anhand von *Stichproben* kann man *Schätzwerte* für den unbekannten Erwartungswert λ berechnen, wenn man z.B. die aus den Stichproben berechneten *Stichprobenmittel* heranzieht (siehe Abschn.24.5.1).

c) Bei der *Produktion* von *Leuchtraketen* ist die *Wahrscheinlichkeit* p gesucht, daß eine beliebig entnommene Rakete funktioniert, d.h. brauchbar ist.
Der Ereignisraum Ω für dieses Zufallsexperiment enthält nur die beiden Elementarereignisse ω_0 (unbrauchbar) und ω_1 (brauchbar), d.h., er hat die Form

$$\Omega = \{\omega_0, \omega_1\}$$

Um die Wahrscheinlichkeit p für einen bestimmten Produktionszeitraum exakt zu bestimmen, müßte man alle produzierten Raketen ausprobieren. Dieses Verfahren ist aber praktisch (ökonomisch) nicht vertretbar. Deshalb wird nur eine *zufällige Stichprobe* entnommen, um einen *Schätzwert* für die unbekannte Wahrscheinlichkeit p zu erhalten. Da das zu untersuchende *Merkmal* einen *qualitativen Charakter* (unbrauchbar, brauchbar) besitzt, definieren wir eine zugehörige *Zufallsgröße* X beispielsweise in der Form

$$X(\omega_0) = 0 \quad , \quad X(\omega_1) = 1$$

Diese *Zufallsgröße* X genügt einer *Null-Eins-Verteilung* (siehe Abschn. 18.1) mit dem *Erwartungswert*

$$E(X) = p$$

Für eine aus der Produktion entnommene (eindimensionale) *Stichprobe*

$$x_1 , x_2 , \ldots , x_n$$

vom *Umfang n* kann man deshalb als *Schätzwert* für die unbekannte Wahrscheinlichkeit p den *empirischen Mittelwert* (*Stichprobenmittel*)

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

verwenden, wenn für die Werte x_i der Stichprobe 0 bzw. 1 gesetzt wird, falls die entsprechende Rakete unbrauchbar bzw. brauchbar war.

♦

☞

Die gegebenen Beispiele lassen erkennen, daß zur *Bestimmung* unbekannter *Parameter* Θ der Wahrscheinlichkeitsverteilung einer Grundgesamtheit nur die Möglichkeit bleibt, durch entnommene *Zufallsstichproben* Informationen zu erhalten. Dabei stellt eine *zufällig entnommene Stichprobe* eine *Realisierung* der die betrachteten *Merkmale* der *Grundgesamtheit* beschreibenden *Zufallsgrößen* X, Y, ... dar, wie aus Kap.23 zu ersehen ist.
Die *Schätzung* unbekannter *Parameter* Θ einer Grundgesamtheit (*Zufallsgröße* X) bildet einen Schwerpunkt der *Schätztheorie.* In der Schätztheorie unterscheidet man zwischen Punkt- und Intervallschätzungen für die unbekannten Parameter:

* Eine *Punktschätzung* liefert einen *Schätzwert* $\hat{\Theta}$ für den unbekannten Parameter Θ.

* Eine *Intervallschätzung* liefert ein *Intervall,* in dem der unbekannte Parameter Θ mit einer vorgegebenen Wahrscheinlichkeit liegt.

Beide Arten von Schätzungen lernen wir in den Abschn.26.3 und 26.4 kennen. Da Schätzungen auf der Basis von *Schätzfunktionen* durchgeführt werden, besprechen wir diese vorher im Abschn.26.2.

♦

☞

Bei einer Reihe praktischer Aufgaben kennt man die *Verteilungsfunktionen näherungsweise* aufgrund des zentralen Grenzwertsatzes (siehe Abschn. 20.3) bzw. der Eigenschaften bekannter Verteilungsfunktionen (siehe Kap. 18), so daß nur *unbekannte Parameter* zu *schätzen* sind.
Falls jedoch die *Verteilungsfunktion* F(x) einer Grundgesamtheit (*Zufallsgröße* X) unbekannt ist, kann man durch eine entnommene Stichprobe mittels der *empirischen Verteilungsfunktion*

$$\hat{F}_n(x)$$

eine *Schätzung* für F an der Stelle x geben (siehe Abschn.23.3).

♦

26.2 Schätzfunktionen

Die *Schätzung* $\hat{\Theta}$ eines unbekannten *Parameters* Θ einer Grundgesamtheit (*Zufallsgröße* X) wird für eine *mathematische Stichprobe*

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

vom Umfang n mittels *Stichprobenfunktionen*

$$\hat{\Theta} = \hat{\Theta}(X_1, X_2, \ldots, X_n)$$

durchgeführt. Diese Stichprobenfunktionen sind ebenfalls wieder *Zufallsgrößen*, deren Wahrscheinlichkeitsverteilung von der Verteilungsfunktion der Grundgesamtheit (*Zufallsgröße* X) abhängen kann (siehe Abschn.23.5).

☞

Im Rahmen der *Schätztheorie* heißen Stichprobenfunktionen *Punktschätzungen, Punktschätzfunktionen, Schätzfunktionen, Schätzungen* oder *Schätzer.* Wir verwenden im vorliegenden Buch die Bezeichnung *Punktschätzfunktion.* Eine *Punktschätzfunktion* ist somit eine *Zufallsgröße* und besitzt eine *Wahrscheinlichkeitsverteilung.*
Eine *Realisierung*

$$\hat{\Theta}_r = \hat{\Theta}(x_1, x_2, \ldots, x_n)$$

der *Punktschätzfunktion*

$$\hat{\Theta} = \hat{\Theta}(X_1, X_2, \ldots, X_n)$$

für eine *konkrete eindimensionale Stichprobe*

$x_1, x_2, \ldots, x_n$

vom *Umfang n* bezeichnet man als *Schätzwert*

$\hat{\Theta}_r$

für den *unbekannten Parameter*

Θ

d.h., man erhält diesen Schätzwert, indem man in die Punktschätzfunktion die Zahlenwerte einer konkreten Stichprobe einsetzt.

♦

Beispiel 26.2:

Im Beispiel 23.4 haben wir die *Punktschätzfunktion*

$$\overline{X} = \overline{X}(X_1, X_2, \ldots, X_n) = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

für die *mathematische Stichprobe*

$X_1, X_2, \ldots, X_n$

kennengelernt, die man zur *Schätzung* des unbekannten *Erwartungswertes* E(X) einer vorliegenden Grundgesamtheit (*Zufallsgröße* X) heranziehen kann. Die so berechnete *Funktion* $\overline{X}$ (*Stichprobenmittel*) ist ebenfalls eine *Zufallsgröße*, die für großes n *näherungsweise normalverteilt* ist, selbst wenn die Zufallsgröße X der Grundgesamtheit eine andere Wahrscheinlichkeitsverteilung besitzt. Dies ist eine Folgerung aus dem *zentralen Grenz-*

wertsatz, da sich $\overline{X}$ als Summe unabhängiger Zufallsgrößen darstellt (siehe Abschn. 20.3). Für eine *konkrete Stichprobe*

$$x_1 , x_2 , \dots , x_n$$

vom *Umfang n* ergibt sich durch Einsetzen als *Schätzwert* der *empirische Mittelwert* $\overline{x}$ (*Stichprobenmittel*)

$$\overline{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

für den unbekannten *Erwartungswert* E(X). Einen weiteren Schätzwert kann man unter Verwendung des *empirischen Medians* gewinnen (siehe Abschn. 24.5.1 und Beispiel 26.3).

♦

26.3 Punktschätzungen

Eine *Punktschätzfunktion*

$$\hat{\Theta} = \hat{\Theta}(X_1 , X_2 , \dots , X_n)$$

ist eine *Stichprobenfunktion,* die dadurch charakterisiert ist, daß man durch Einsetzen einer *konkreten eindimensionalen Stichprobe*

$$x_1 , x_2 , \dots , x_n$$

vom *Umfang n* einen einzigen Zahlenwert gewinnt, den man als *Schätzwert* für den *unbekannten Parameter* Θ einer vorliegenden Grundgesamtheit (*Zufallsgröße* X) verwendet. Hiermit erhält man aber keine Aussagen über die *Genauigkeit* des berechneten *Schätzwertes.*

☞

Bei *Stichproben* mit *kleinem Umfang* können bei *Punktschätzungen* erhebliche *Fehler* auftreten. Deshalb werden in der statistischen Schätztheorie zusätzlich *Intervallschätzungen* entwickelt, die bzgl. der Genauigkeit Aussagen liefern (siehe Abschn.26.4).

♦

☞

In einer Reihe von Fällen lassen sich zur *Schätzung* eines *unbekannten Parameters*

Θ

mehrere Punktschätzfunktionen

$\hat{\Theta}$

angeben. So kann man z.B. zur Schätzung des unbekannten *Erwartungswertes* E(X) einer *Zufallsgröße* X den

* *empirischen Mittelwert*

* *empirischen Median*

heranziehen (siehe Beispiel 26.3).

♦

Von dem Statistiker R.A. Fisher wurden folgende *Eigenschaften* für eine *Punktschätzfunktion*

$$\hat{\Theta} = \hat{\Theta}(X_1, X_2, \ldots, X_n)$$

gefordert, um eine möglichst *"gute" Schätzung* für einen *unbekannten Parameter*

$$\Theta$$

zu erhalten:

- *Erwartungstreue* (*Unverzerrtheit*)

  Eine *Punktschätzfunktion*

  $$\hat{\Theta} = \hat{\Theta}(X_1, X_2, \ldots, X_n)$$

  eines *unbekannten Parameters* Θ heißt *erwartungstreu* (*unverzerrt*), wenn gilt

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

  d.h., der Erwartungswert der gelieferten Schätzung $\hat{\Theta}$ ist gleich dem gesuchten Parameter Θ.
  Gilt diese Eigenschaft nur in der schwächeren Form für $n \to \infty$, so heißt die Schätzung $\hat{\Theta}$ *asymptotisch erwartungstreu* bzw. *asymptotisch unverzerrt.*

- *Konsistenz*

  Eine *Punktschätzfunktion*

  $$\hat{\Theta} = \hat{\Theta}(X_1, X_2, \ldots, X_n)$$

  eines *unbekannten Parameters* Θ heißt *konsistent*, wenn sie für $n \to \infty$ gegen den gesuchten Parameter Θ in Wahrscheinlichkeit konvergiert, d.h. für beliebiges $\varepsilon > 0$ folgendes

  $$\lim_{n \to \infty} P(|\hat{\Theta}(X_1, X_2, \ldots, X_n) - \Theta| < \varepsilon) = 1$$

  gilt.

- *Effizienz*

  Betrachten wir zwei *erwartungstreue* und *konsistente Punktschätzfunktionen*

  $$\hat{\Theta}_1 = \hat{\Theta}_1(X_1, X_2, \ldots, X_n) \quad \text{und} \quad \hat{\Theta}_2 = \hat{\Theta}_2(X_1, X_2, \ldots, X_n)$$

für einen *unbekannten Parameter* Θ. Dann heißt die *Punktschätzfunktion*

$$\hat{\Theta}_1 = \hat{\Theta}_1(X_1, X_2, \ldots, X_n)$$

effizienter (*wirksamer*) als

$$\hat{\Theta}_2 = \hat{\Theta}_2(X_1, X_2, \ldots, X_n)$$

wenn für ihre *Varianzen/Streuungen* gilt

$$\sigma^2(\hat{\Theta}_1) < \sigma^2(\hat{\Theta}_2)$$

Als *effiziente Punktschätzfunktion* wird diejenige mit der kleinsten Varianz/Streuung bezeichnet. Eine nichteffiziente Punktschätzfunktion wird als *ineffizient* bezeichnet.

☞

Gibt es für eine praktische Aufgabe mehrere Punktschätzfunktionen, die alle gegebenen Eigenschaften besitzen, so wird man eine von diesen verwenden. Wenn nur Schätzfunktionen vorliegen, die einige dieser Eigenschaften haben, so muß der Anwender über die Auswahl entscheiden.

♦

☞

Das Nachprüfen der gegebenen Eigenschaften ist für eine vorliegende Punktschätzfunktion nicht immer einfach. Deshalb geben wir im folgenden die Eigenschaften für die zwei häufig benötigten Punktschätzfunktionen *Stichprobenmittel* und *Stichprobenvarianz.*

♦

☞

Betrachten wir *Punktschätzfunktionen* für die beiden Parameter *Erwartungswert* und *Varianz/Streuung* einer *Zufallsgröße* X. Man kann beweisen, daß die *Punktschätzfunktion*

* $$\overline{X} = \frac{X_1 + X_2 + \ldots + X_n}{n} = \frac{1}{n} \cdot \sum_{i=1}^{n} X_i$$ (*Stichprobenmittel*)

für die *Zufallsgröße* X mit einer beliebigen Wahrscheinlichkeitsverteilung *erwartungstreu, konsistent* und *effizient* ist.
Eine *Realisierung* des *Stichprobenmittels* für eine *konkrete eindimensionale Stichprobe*

$$x_1, x_2, \ldots, x_n$$

vom *Umfang n* liefert als *Schätzwert* für den *Erwartungswert*

$$E(X) = \mu$$

den *empirischen Mittelwert* (siehe Abschn.24.5.1)

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

Der *Median* (siehe Abschn.24.5.1) liefert ebenfalls eine *erwartungstreue Schätzung* für den Erwartungswert, die allerdings nicht effizient ist.

* $$S_X^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (X_i - \overline{X})^2 \qquad \text{(\textit{Stichprobenvarianz})}$$

für die *Zufallsgröße* X mit einer beliebigen Verteilung *erwartungstreu* und *konsistent* ist.
Eine *Realisierung* der *Stichprobenvarianz* für eine *konkrete eindimensionale Stichprobe*

$x_1 \, , \, x_2 \, , \, \ldots \, , \, x_n$

vom *Umfang n* liefert als *Schätzwert* für die *Varianz/Streuung*

$V(X) = \sigma^2$

die *empirischen Varianz/Streuung* (siehe Abschn.24.5.1)

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

Wenn man in der Formel der *Stichprobenvarianz* statt durch n–1 durch n dividiert, d.h., die Formel in der Form

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

verwendet, so ist die hierdurch erhaltene Punktschätzfunktion *nicht erwartungstreu*, sondern nur *asymptotisch erwartungstreu*.

* $$S_X = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^{n} (X_i - \overline{X})^2} \qquad \text{(\textit{Stichprobenstandardabweichung})}$$

für die *Zufallsgröße* X mit einer beliebigen Verteilung als *Schätzung* für die *Standardabweichung* σ *nicht erwartungstreu* ist. Obwohl die Stichprobenvarianz

S_X^2

erwartungstreu ist, kann für S_X nur *asymptotisch erwartungstreu* nachgewiesen werden.

♦

☞

Die Systeme MATHCAD und MATLAB stellen zur Berechnung von empirischen *Mittelwert*, *Median* und *Varianz/Streuung* vordefinierte Funktionen zur Verfügung, die wir bereits im Abschn.24.5.1 kennenlernten. Damit können beide Systeme erfolgreich zur *Schätzung* von *Erwartungswert* und *Varianz/ Streuung* einer Grundgesamtheit (*Zufallsgröße* X) herangezogen werden, wie wir im Beispiel 26.3 illustrieren. Da man in beiden Systemen beliebige Funktionen einfach definieren kann (siehe Abschn.13.1.3) lassen sie sich auch zur Berechnung beliebiger *Punktschätzfunktionen* einsetzen.

♦

Beispiel 26.3:

Betrachten wir ein *Zahlenbeispiel* für die *Schätzung* (*Punktschätzung*) der unbekannten Parameter

* *Erwartungswert* μ
* *Varianz/Streuung* σ^2

für eine gegebene *Zufallsgröße* X:

Wir untersuchen die Produktion eines Zubehörteils, d.h., die betrachtete Grundgesamtheit besteht aus den in einem bestimmten Zeitraum produzierten Teilen. Das zu untersuchende Merkmal dieser Grundgesamtheit sei die Länge des Zubehörteils (in cm) und werde durch die *Zufallsgröße* X beschrieben. Für die Zufallsgröße X kann näherungsweise eine *Normalverteilung* aufgrund des zentralen Grenzwertsatzes vorausgesetzt werden, da sich bei der Produktion des Zubehörteils eine Reihe unabhängiger zufälliger Effekte überlagert. Für die folgenden Rechnungen benötigen wir keine Aussagen über die Wahrscheinlichkeitsverteilung von X.

Um Aussagen über die *Zufallsgröße* X zu erhalten, wird eine Zufallsstichprobe von sieben Teilen aus der Produktion entnommen und gemessen, wobei folgende Zahlenwerte (in cm) für die Länge erhalten werden:

2.33 2.34 2.37 2.32 2.35 2.37 2.34

Diese Meßwerte bilden eine *Stichprobe* vom *Umfang* *7* aus der betrachteten Grundgesamtheit der Zubehörteile, für die wir im folgenden mittels MATHCAD und MATLAB als *Schätzwert* für

- den unkannten *Erwartungswert*
  * 2.3457 mittels des *empirischen Mittelwertes* (erwartungstreue und effiziente Schätzung)
  * 2.34 mittels des *Medians* (erwartungstreue aber ineffiziente Schätzung)
- die unbekannte *Varianz/Streuung*

0.0003619 mittels der *empirischen Varianz/Streuung* (erwartungstreue und effiziente Schätzung)

erhalten:

MATHCAD

$$x := \begin{pmatrix} 2.33 \\ 2.34 \\ 2.37 \\ 2.32 \\ 2.35 \\ 2.37 \\ 2.34 \end{pmatrix}$$

mean (x) = 2.3457 **median** (x) = 2.34 **var** (x) = 3.619×10^{-4}

MATHCAD

MATLAB

```
>> x = [ 2.33  2.34  2.37  2.32  2.35  2.37  2.34 ] ;
>> mean ( x )
ans =
   2.3457
>> median ( x )
ans =
   2.3400
>> var ( x )
ans =
  3.6190e-004
```

MATLAB

♦

☞

Bisher haben wir vorgegebene *Punktschätzfunktionen* betrachtet und diese bzgl. spezieller Eigenschaften (Erwartungstreue, Konsistenz, Effizienz...) untersucht. Es stellt sich nun die Frage, wie man überhaupt zu einer *Punktschätzfunktion* gelangt, die zusätzlich noch gewisse Eigenschaften besitzt. Hierzu wurden Methoden entwickelt, von denen wir die

* *Maximum-Likelihood-Methode*
* *Methode der kleinsten Quadrate*
* *Momentenmethode*

in den folgenden Abschn.26.3.1–26.3.3 kennenlernen.

♦

26.3.1 Maximum-Likelihood-Methode

Die von R.A. Fisher entwickelte *Maximum-Likelihood-Methode* zur Konstruktion einer *Punktschätzfunktion* setzt voraus, daß die Wahrscheinlichkeitsverteilung der Grundgesamtheit (*Zufallsgröße* X) bekannt ist, während gewisse Parameter unbekannt sind. Wir beschränken uns im folgenden auf einen unbekannten Parameter Θ, da die Problematik bei mehreren Parametern analog ist, wie wir im Beispiel 26.4c für zwei Parameter illustrieren.
Ausgangspunkt der *Maximum-Likelihood-Methode* ist eine *konkrete Stichprobe* vom *Umfang n*, die aus der zu untersuchenden Grundgesamtheit (*Zufallsgröße* X) entnommen wurde. Das *Prinzip* dieser Methode besteht darin, bei vorliegenden Daten einer Stichprobe dasjenige Modell zu verwenden, unter welchem diese Daten die *größte Wahrscheinlichzkeit* des Auftretens besitzen.
Um das *Prinzip* einer *Maximum-Likelihood-Methode* bei einem unbekannten Parameter Θ zu realisieren, bildet man die *Likelihood-Funktion*

$$L(x_1, x_2, \ldots, x_n;\Theta)$$

der *konkreten Stichprobe*

$$x_1, x_2, \ldots, x_n$$

vom *Umfang n*, die von den Zahlenwerten dieser Stichprobe und dem *unbekannten Parameter* Θ abhängt und folgendermaßen *definiert* ist:

- $L(x_1, x_2, \ldots, x_n;\Theta) =$

  $$P(X = x_1; \Theta)\cdot P(X = x_2; \Theta)\cdot \ldots \cdot P(X = x_n; \Theta)$$

  für eine *diskrete Zufallsgröße* X mit den *Einzelwahrscheinlichkeiten*

  $$P(X = x_i; \Theta) \qquad (i = 1, 2, \ldots, n)$$

  Wenn man die *Wahrscheinlichkeitsfunktion* f verwendet (siehe Abschn. 18.1), so schreibt sich die *Likelihood-Funktion* für diskrete Zufallsgrößen in der Form

  $$L(x_1, x_2, \ldots, x_n;\Theta) = f(x_1; \Theta)\cdot f(x_2; \Theta)\cdot \ldots \cdot f(x_n; \Theta)$$

- $L(x_1, x_2, \dots, x_n; \Theta) = f(x_1; \Theta) \cdot f(x_2; \Theta) \cdot \dots \cdot f(x_n; \Theta)$

  für eine *stetige Zufallsgröße* X mit der *Wahrscheinlichkeitsdichte* $f(t; \Theta)$

☞

Wir haben gesehen, daß sich die *Likelihood-Funktion* allgemein in der Form

$$L(x_1, x_2, \dots, x_n; \Theta) = f(x_1; \Theta) \cdot f(x_2; \Theta) \cdot \dots \cdot f(x_n; \Theta) = \prod_{k=1}^{n} f(x_k; \Theta)$$

schreibt, worin f für

* *diskrete Zufallsgrößen*

  die Wahrscheinlichkeitsfunktion

* *stetige Zufallsgrößen*

  die Wahrscheinlichkeitsdichte

bezeichnet. Wenn man in die *Likelihood-Funktion* die Zahlenwerte

$x_1, x_2, \dots, x_n$

einer konkreten eindimensionalen Stichprobe vom Umfang n einsetzt, so ist sie nur noch eine *Funktion* des *unbekannten Parameters* Θ.

♦

☞

Die *Vorgehensweise* bei der Anwendung der *Maximum-Likelihood-Methode* besteht im folgenden:

* Es wird ein *Schätzwert* (*Maximum-Likelihood-Schätzwert*) für den unbekannten *Parameter* Θ in Abhängigkeit von einer vorliegenden *Stichprobe* so bestimmt, daß die *Likelihood-Funktion* bzgl. Θ ein *Maximum* annimmt.

* Unter der Voraussetzung, daß die *Likelihood-Funktion differenzierbar* ist, kann man die notwendige *Bedingung*

  $$\frac{\partial}{\partial \Theta} L(x_1, x_2, \dots, x_n; \Theta) = 0$$

  für ein *relatives Maximum* heranziehen und eventuell noch zusätzlich die *hinreichende Bedingung* verwenden, um das Maximum von L zu bestimmen. Damit erhält man als Lösung der gegebenen Gleichung einen *Schätzwert* (*Maximum-Likelihood-Schätzwert*) für den unbekannten *Parameter* Θ.

* Häufig ist es für die Berechnung günstiger, den *natürlichen Logarithmus* der *Likelihood-Funktion* zu benutzen, d.h.

$$\ln L(x_1, x_2, \ldots, x_n; \Theta) = \sum_{k=1}^{n} \ln f(x_k; \Theta)$$

Diese Funktion besitzt offensichtlich die gleichen Maxima wie die ursprüngliche Funktion. Man erhält hierfür die folgende notwendige Bedingung für ein Maximum:

$$\frac{\partial}{\partial \Theta} \sum_{k=1}^{n} \ln f(x_k; \Theta) = \sum_{k=1}^{n} \frac{1}{f(x_k; \Theta)} \cdot \frac{\partial}{\partial \Theta} f(x_k; \Theta) = 0$$

♦

Für die mittels der *Maximum-Likelihood-Methode* bestimmte *Punktschätzfunktion* (*Maximum-Likelihood-Schätzfunktion*) kann man unter gewissen Bedingungen beweisen, daß sie *asymptotisch erwartungstreu*, *konsistent* und *asymptotisch normalverteilt* sind.

♦

Wenn man die *Likelihood-Funktion* für eine Aufgabe per Hand aufgestellt hat, kann man zur Bestimmung ihres Maximums die Systeme MATHCAD und MATLAB heranziehen, indem man die vordefinierten Funktionen zur Differentiation und Gleichungsauflösung beider Systeme verwendet, wie im folgenden Beispiel 26.4a–c illustriert wird.
MATLAB besitzt zusätzlich *vordefinierte Funktionen* zur *Maximum-Likelihood-Methode:*

MATLAB

MATLAB stellt mit der allgemeinen *vordefinierten Funktion*

mle ('*Verteilung*' , x)

eine *Maximum-Likelihood-Schätzfunktion* für die unbekannten Parameter einer Zufallsgröße X mit der Wahrscheinlichkeitsverteilung '*Verteilung*' zur Verfügung, wobei sich die Zahlenwerte der konkreten Stichprobe im Vektor **x** befinden müssen. Dabei ist für '*Verteilung*' der MATLAB-Name für die vorliegende Wahrscheinlichkeitsverteilung einzusetzen. Die Anwendung dieser Funktion illustrieren wir im Beispiel 26.4a.
Des weiteren sind in MATLAB noch *Maximum-Likelihood-Schätzfunktionen* für spezielle Wahrscheinlichkeitsverteilungen vordefiniert, wie z.B.

* **binofit** (x , n)

  für die *Binomialverteilung*

* **normfit** (x)

für die *Normalverteilung*

* **poissfit** (x)

für die *Poisson-Verteilung*

Die Anwendung dieser Funktionen gestaltet sich problemlos, wobei **x** den Vektor der Zahlenwerte der vorliegenden Stichprobe darstellt. Der Anwender erhält zusätzliche Informationen zu diesen Funktionen aus der Hilfe von MATLAB. Deshalb sind im folgenden alle Funktionen zur Parameterschätzung aus dem MATLAB-Hilfefenster aufgelistet:

Parameter estimation.

betafit	*- Beta parameter estimation.*
binofit	*- Binomial parameter estimation.*
expfit	*- Exponential parameter estimation.*
gamfit	*- Gamma parameter estimation.*
mle	*- Maximum likelihood estimation.*
normfit	*- Normal parameter estimation.*
poissfit	*- Poisson parameter estimation.*
raylfit	*- Rayleigh parameter estimation.*
unifit	*- Uniform parameter estimation.*
weibfit	*- Weibull parameter estimation.*

Wir illustrieren die Anwendung dieser in MATLAB vordefinierten Funktionen im Beispiel 26.4a und c für die Funktionen **normfit**, **poissfit** und **mle**.

MATLAB

Beispiel 26.4:

a) Bestimmen wir einen *Maximum-Likelihood-Schätzwert* für den *unbekannten Parameter* $\Theta = \lambda$ (Erwartungswert) einer *Poisson-verteilten* Grundgesamtheit (*diskrete Zufallsgröße* X) anhand einer konkreten *Stichprobe*

$$x_1 \, , \, x_2 \, , \, \ldots \, , \, x_n$$

vom *Umfang n.* Mit der *Wahrscheinlichkeit*

$$P(X=k) \; = \; \frac{\lambda^k}{k!} \cdot e^{-\lambda} \qquad\qquad (\, k = 0 \, , 1 \, , 2 \, , \ldots)$$

für die *Poisson-Verteilung* (siehe Abschn.18.1) ergibt sich damit die folgende *Likelihood-Funktion*

$$L\,(x_1 \, , \, x_2 \, , \, \ldots \, , \, x_n \, ; \lambda\,) \; = \; e^{-n \cdot \lambda} \cdot \prod_{k=1}^{n} \frac{\lambda^{x_k}}{x_k!}$$

Da sich die Maximierung dieser Funktion schwierig gestaltet, empfiehlt sich die *Logarithmierung* der *Likelihood-Funktion:*

$$\ln L(x_1 , x_2 , \dots , x_n ; \lambda) = -n\cdot\lambda + \ln\lambda\cdot\sum_{k=1}^{n} x_k - \sum_{k=1}^{n} \ln(x_k !)$$

für die die notwendige Bedingung

$$\frac{\partial}{\partial\lambda} \ln L(x_1 , x_2 , \dots , x_n ; \lambda) = -n + \frac{1}{\lambda}\cdot\sum_{k=1}^{n} x_k = 0$$

die Lösung

$$\lambda = \frac{1}{n}\cdot\sum_{k=1}^{n} x_k = \overline{x}$$

liefert, d.h., der berechnete *Maximum-Likelihood-Schätzwert* für den *Erwartungswert* λ ist gleich dem *empirischen Mittelwert* $\overline{x}$.

Die per Hand durchgeführten Rechnungen können von MATHCAD erledigt werden:

MATHCAD

MATHCAD berechnet mit seinem vordefinierten Operator zur Differentiation und dem Schlüsselwort **solve** zur Gleichungslösung (siehe Abschn.12.5) das per Hand berechnete Ergebnis nur bei Verwendung der logarithmierten Likelihood-Funktion:

$$\frac{d}{d\lambda}\left(\ln(\lambda)\cdot\sum_{k=1}^{n} xk - \sum_{k=1}^{n} \ln(xk!) - n\cdot\lambda\right) = 0 \text{ solve}, \lambda \rightarrow \frac{\displaystyle\sum_{k=1}^{n} x}{n}$$

MATHCAD

MATLAB

Mit MATLAB ist es dem Autor nicht gelungen, die erforderlichen symbolischen Rechnungen zur Berechnung der *Maximum-Likelihood-Schätzwertes* für den *Erwartungswert* λ der *Poisson-Verteilung* mittels der vordefinierten Funktionen **diff** und **solve** zur Differentiation bzw. Gleichungslösung durchzuführen.

Deshalb wenden wir die vordefinierte Funktion **poissfit** von MATLAB an, um für eine konkrete Stichprobe den *Maximum-Likelihood-Schätzwert* für den *Erwartungswert* λ zu berechnen:

Dazu betrachten wir die Anzahl X der in einem kurzen Zeitintervall zerfallenden Atome eines radioaktiven Stoffes, die einer Poisson-Verteilung genügen soll, und verwenden folgende konkrete Stichprobe:

10 11 9 12 13 8

MATLAB berechnet mittels

```
>> x = [ 10 11 9 12 13 8 ] ;
>> poissfit ( x )
ans =
  10.5000
```

als *Maximum-Likelihood-Schätzwert* des *Erwartungswertes* λ den zu Beginn theoretisch berechneten empirischen Mittelwert $\bar{x}$, wie die Berechnung des empirischen Mittelwerts mittels der in MATLAB vordefinierten Funktion **mean** zeigt:

```
>> mean ( x )
ans =
  10.5000
```

Das gleiche Ergebnis ergibt sich, wenn man die allgemeine Funktion **mle** auf die *Poisson-Verteilung* anwendet:

```
>> mle ( 'poiss' , x )
ans =

  10.5000
```

b) Verwenden wir die *Maximum-Likelihood-Methode* um die unbekannte *Wahrscheinlichkeit* p=P(A) für das *Eintreten* eines *Ereignisses* A zu *schätzen*, d.h., es ist der *Parameter* $\Theta = p$ zu *schätzen*. Dazu verwendet man eine Null-Eins-verteilte *Zufallsgröße* X der folgenden Form

$$X = \begin{cases} 0 & \text{Ereignis A nicht eingetreten} \\ 1 & \text{Ereignis A eingetreten} \end{cases}$$

mit den *Wahrscheinlichkeiten*

$P(X=0) = 1 - p$ und $P(X=1) = p$

Zur Bestimmung eines Schätzwertes für die unbekannte Wahrscheinlichkeit p führt man n Zufallsexperimente (zufällige Versuche) durch (z.B.

n=100) und beobachtet, wie oft hierbei das Ereignis A eintritt. Damit haben wir eine *konkrete eindimensionale Stichprobe*

$x_1, x_2, \dots, x_{100}$

vom *Umfang 100* vorliegen, wobei die *Stichprobenwerte* nur die Zahlen 1 oder 0 annehmen können, je nachdem ob das Ereignis A eingetreten ist oder nicht.
Wenn das Ereignis A in der Stichprobe z.B. 70 mal eintritt, so hat die *Likelihood-Funktion* folgende Gestalt

$$L(x_1, x_2, \dots, x_n; p) = P(X=0;p)^{30} \cdot P(X=1;p)^{70}$$
$$= (1-p)^{30} \cdot p^{70}$$

Wenn wir die *logarithmierte Likelihood-Funktion*

$$\ln L(x_1, x_2, \dots, x_n; p) = 30 \cdot \ln(1-p) + 70 \cdot \ln p$$

nach p differenzieren und Null setzen, so ergibt sich die folgende *Gleichung* als notwendige Bedingung für ihr Maximum

$$-\frac{30}{1-p} + \frac{70}{p} = 0$$

Aus dieser Gleichung erhält man durch einfache Auflösung nach p die *Lösung*

$$p = \frac{70}{100} = 0.7$$

Die durchgeführten Rechnungen lassen sich in MATHCAD und MATLAB mit den Möglichkeiten zur Differentiation und Gleichungslösung erledigen:

MATHCAD

Wir verwenden den vordefinierten *Operator* zur *Differentiation* und das *Schlüsselwort* **solve** zur *Gleichungslösung* (siehe Abschn. 13.2 und 12.5), um die notwendige Bedingung für ein Maximum aufzustellen und nach p aufzulösen:

$$\frac{d}{dp}(30 \cdot \ln(1-p) + 70 \cdot \ln(p)) = 0 \;\mathbf{solve}\;, p \rightarrow \frac{7}{10}$$

Damit berechnet MATHCAD als *Lösung* 7/10.

MATHCAD

Wir verwenden die beiden *vordefinierten Funktionen* **diff** und **solve** zur *symbolischen Differentiation* bzw. *Gleichungslösung* (siehe Abschn. 13.2 und 12.5), um die notwendige Bedingung für ein Maximum aufzustellen und nach p aufzulösen:

```
>> syms p ; solve ( ' diff ( 30 * log ( 1 - p ) + 70 * log ( p ) = 0 , p ) '
   , ' p ' )
ans =
7/10
```

Damit berechnet MATLAB als *Lösung* 7/10.

Die berechnete Lösung 0.7 ist ein *Maximum-Likelihood-Schätzwert* für die unbekannte *Wahrscheinlichkeit* p. Man sieht, daß dieser *Maximum-Likelihood-Schätzwert* die *relative Häufigkeit* liefert (siehe Abschn.16.2.2).

c) Verwenden wir die *Maximum-Likelihood-Methode* um den *unbekannten Erwartungswert* μ und die *unbekannte Varianz/Streuung* σ^2 für eine $N(\mu,\sigma)$-*normalverteilte Zufallsgröße* X zu schätzen. In diesem Beispiel sind zwei Parameter zu schätzen, d.h. die *Likelihood-Funktion* hat für eine *konkrete eindimensionale Stichprobe*

$x_1 \, , \, x_2 \, , \, \ldots \, , \, x_n$

vom *Umfang n* für die *Normalverteilung* die folgende Gestalt

$$L(x_1 , x_2 , \ldots , x_n ; \mu , \sigma^2) = \frac{1}{\left(\sqrt{2 \cdot \pi \cdot \sigma^2}\right)^n} \cdot e^{-\frac{1}{2 \cdot \sigma^2} \cdot \sum_{i=1}^{n} (x_i - \mu)^2}$$

d.h., sie ist eine Funktion der zwei Variablen μ und σ^2. Zur einfacheren Berechnung eines lokalen Maximums wird die *Likelihood-Funktion* logarithmiert. Danach ergeben sich als notwendige Bedingungen für ein Maximum die beiden Gleichungen

$$\frac{\partial \ln L}{\partial \mu} = \frac{1}{\sigma^2} \cdot \sum_{i=1}^{n} (x_i - \mu) = 0$$

$$\frac{\partial \ln L}{\partial \sigma^2} = -\frac{n}{2 \cdot \sigma^2} + \frac{1}{2 \cdot \sigma^4} \cdot \sum_{i=1}^{n} (x_i - \mu)^2 = 0$$

deren *Lösungen*

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

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

man z.B. durch Elimination erhält.

Die erhaltenen Lösungen liefern die *Maximum-Likelihood-Schätzfunktionen* für die beiden unbekannten Parameter *Erwartungswert* μ und *Varianz/Streuung* σ^2 der *Normalverteilung*. Man sieht, daß sich als *Maximum-Likelihood-Schätzfunktion* für

* den *Erwartungswert*

  das *Stichprobenmittel* $\quad \bar{X} = \frac{1}{n} \cdot \sum_{i=1}^{n} X_i$

* die *Varianz/Streuung*

  die *Stichprobenvarianz* $\quad S_X^2 = \frac{1}{n} \cdot \sum_{i=1}^{n} (X_i - \bar{X})^2$

ergeben, wobei allerdings für die *Varianz/Streuung* nur der asymptotisch erwartungstreue Wert geliefert wird, da in der berechneten Formel durch n dividiert wird (siehe Abschn.24.5.1).

Berechnen wir die eben per Hand erhaltenen *Maximum-Likelihood-Schätzfunktionen* mittels MATHCAD und MATLAB:

MATHCAD

Mittels MATHCAD läßt sich das Ergebnis ohne Logarithmierung der Likelihood-Funktion berechnen. Wir setzen die Variable

σ^2 gleich s.

Damit ermittelt MATHCAD mit seinem vordefinierten Operator zur Differentiation und dem Schlüsselwort **solve** zur Gleichungslösung (siehe Abschn.12.5) das gegebene Ergebnis, wobei das Ergebnis für s in einer anderen Form geschrieben wird. Die Überführung in die oben gegebene Form überlassen wir dem Leser:

$$\begin{bmatrix} \dfrac{d}{d\mu}\dfrac{1}{\left(\sqrt{2\cdot\pi\cdot s}\right)^n}\cdot e^{\frac{-1}{2\cdot s}\cdot\sum_{i=1}^{n}(x_i-\mu)^2} = 0 \\[2ex] \dfrac{d}{ds}\dfrac{1}{\left(\sqrt{2\cdot\pi\cdot s}\right)^n}\cdot e^{\frac{-1}{2\cdot s}\cdot\sum_{i=1}^{n}(x_i-\mu)^2} = 0 \end{bmatrix} \text{solve}, \begin{pmatrix}\mu \\ s\end{pmatrix} \rightarrow \begin{bmatrix} \dfrac{\sum_{i=1}^{n} x_i}{n} & \dfrac{\left[-\left(\sum_{i=1}^{n} x_i\right)^2 + \sum_{i=1}^{n}(x_i)^2\cdot n\right]}{n^2} \end{bmatrix}$$

MATHCAD

MATLAB

Bei MATLAB ist es dem Autor nicht gelungen, die erforderlichen symbolischen Rechnungen zur Berechnung der *Maximum-Likelihood-Schätzfunktion* für *Erwartungswert* und *Varianz/Streuung* der *Normalverteilung* mittels der vordefinierten Funktionen **diff** und **solve** zur Differentiation bzw. Gleichungslösung durchzuführen.
Deshalb wenden wir die vordefinierten Funktionen **normfit** und **mle** von MATLAB an, um für eine konkrete Stichprobe die *Maximum-Likelihood-Schätzwert* für *Erwartungswert* und *Varianz/Streuung* zu berechnen:
Wir betrachten dazu die Aufgabe aus Beispiel 26.3, in der wir die Produktion eines Zubehörteils untersuchen d.h., die betrachtete Grundgesamtheit besteht aus den in einem bestimmten Zeitraum produzierten Teilen. Das zu untersuchende Merkmal dieser Grundgesamtheit sei die Länge des Zubehörteils (in cm) und werde durch die *Zufallsgröße* X beschrieben.
Um Aussagen über die *Zufallsgröße* X zu erhalten, werden zufällig sieben Teile aus der Produktion entnommen und gemessen, wobei wir folgende Zahlenwerte (in cm) annehmen:

2.33 2.34 2.37 2.32 2.35 2.37 2.34

Diese Meßwerte bilden eine *Stichprobe* vom *Umfang* 7 aus der betrachteten Grundgesamtheit der Zubehörteile.

MATLAB berechnet mittels

```
>> x = [ 2.33 2.34 2.37 2.32 2.35 2.37 2.34 ] ;

>> [ m , s ] = normfit ( x )

m =

   2.3457
```

```
s =

   0.0190
```

für die gegebene Stichprobe als *Maximum-Likelihood-Schätzwert* den *empirischen Mittelwert* m =2.3457 und die *empirische Standardabweichung* s=0.0190

Bei Anwendung der allgemeine Funktion **mle** auf die *Normalverteilung* berechnet MATLAB:

```
>> x = [ 2.33 2.34 2.37 2.32 2.35 2.37 2.34 ] ;

>> mle ( 'norm' , x )

ans =

   2.3457    0.0176
```

d.h., für die berechnete empirische Standardabweichung ergibt sich eine kleine Abweichung.

♦

26.3.2 Methode der kleinsten Quadrate

Die bekannte bereits von Gauß angewandte *Methode der kleinsten Quadrate* haben wir schon bei der empirischen Regression kennengelernt (siehe Abschn.24.6), so daß wir das Prinzip nur anhand eines Beispiels illustrieren. Des weiteren werden wir diese Methode im Kap.30 im Rahmen der Korrelation und Regression erfolgreich einsetzen.
Die *Methode der kleinsten Quadrate* ist eng mit der *Maximum-Likelihood-Methode* verwandt und kann noch benutzt werden, wenn die Wahrscheinlichkeitsverteilung der betrachteten Grundgesamtheit (*Zufallsgröße* X) nicht bekannt ist.
Betrachten wir als Anwendungsbeispiel die Schätzung des unbekannten Erwartungswertes $E(X)=\mu$ einer Grundgesamtheit (*Zufallsgröße* X) mittels der Methode der kleinsten Quadrate.

Beispiel 26.5:

Für eine *konkrete eindimensionale Stichprobe*

$x_1 , x_2 , \ldots , x_n$

vom *Umfang n* soll mittels der *Methode der kleinsten Quadrate* ein *Schätzwert* (*Kleinst-Quadrate-Schätzwert*) für den unbekannten *Erwartungswert*

$E(X) = \mu$

einer gegebenen Grundgesamtheit (*Zufallsgröße* X) bestimmt werden.
Das Prinzip der *Methode der kleinsten Quadrate* besteht darin, einen *Schätzwert* $\bar{x}$ für μ so zu bestimmen, daß die Summe der Quadrate der Abweichungen zwischen den Stichprobenwerten und dem gesuchten Parameter μ minimal werden, d.h., es ist die folgende Minimierungsaufgabe zu lösen:

$$\sum_{i=1}^{n} (x_i - \bar{x})^2 = \underset{\mu}{\text{Minimum}} \sum_{i=1}^{n} (x_i - \mu)^2$$

Die Lösung dieser Aufgabe ergibt sich einfach aus der notwendigen Bedingung für ein relatives Minimum durch Nullsetzen der ersten Ableitung bzgl. μ. Man erhält das *Ergebnis*

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

d.h., der ermittelte *Kleinst-Quadrate-Schätzwert* ist gleich dem *empirischen Mittelwert* $\bar{x}$. Als *Kleinst-Quadrate-Schätzfunktion* ergibt sich damit das *Stichprobenmittel* $\overline{X}$

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

♦

☞

Die Systeme MATHCAD und MATLAB können erfolgreich für die Methode der kleinsten Quadrate herangezogen werden, da sie die hier auftretenden Minimierungsaufgaben mittels ihrer vordefinierten Funktionen zur Differentiation und Gleichungslösung lösen können.

♦

26.3.3 Momentenmethode

Diese einfach zu handhabende Methode wurde bereits öfters angewandt. Das Prinzip der *Momentenmethode* besteht darin, einen Parameter, der sich durch ein *theoretisches Moment* ausdrücken läßt (siehe Kap.19), durch den entsprechenden Ausdruck des *empirischen Moments* zu schätzen. Als bekannteste Fälle hierzu haben wir

* *empirischen Mittelwert*

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

* *empirische Varianz/Streuung*

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

als *Schätzwerte* für *Erwartungswert* bzw. *Varianz/Streuung* kennengelernt, wobei die zugehörigen Punktschätzfunktionen (*Stichprobenmittel* bzw. *Stichprobenvarianz*) konsistent und mindestens asymptotisch erwartungstreu sind.

☞

Da die Systeme MATHCAD und MATLAB zur Berechnung von empirischen Mittelwert und Varianz/Streuung vordefinierte Funktionen zur Verfügung stellen, können beide erfolgreich für die Momentenmethode herangezogen werden.

♦