

6. STATISTISCHE KENNGRÖSZEN VON HÄUFIGKEITSVERTEILUNGEN

Beob. x_1, \dots, x_n , Charakteristika der „Verteilung“ gesucht

6.1 Lageparameter

a) Mittelwert

$$\bar{x} = \frac{x_1 + \dots + x_n}{n} = \frac{1}{n} \sum_{i=1}^n x_i$$

Bem.: Sog. Ausreißer problematisch

b) Median einer empirischen Verteilung

Für Urdaten x_1, \dots, x_n

bei ungeradem $n = 2k+1$: $x_{(k+1)}$

bei geradem $n = 2k$: $\frac{x_{(k)} + x_{(k+1)}}{2}$

Für gruppierte Daten mittels Summenpolygon

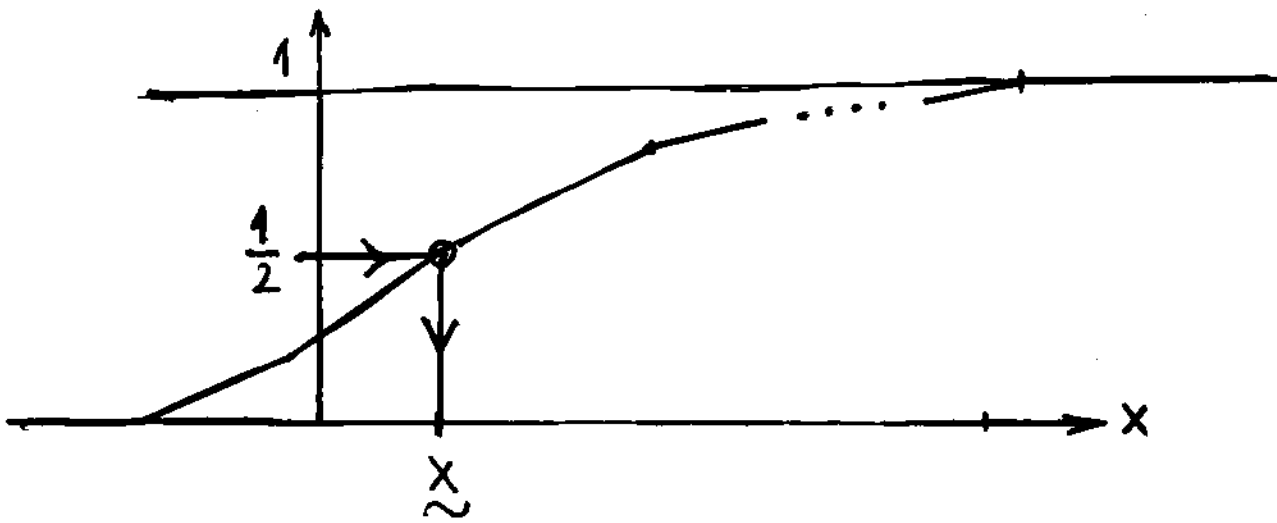
(Abb.)

Allgemeiner: Empirische Fraktile

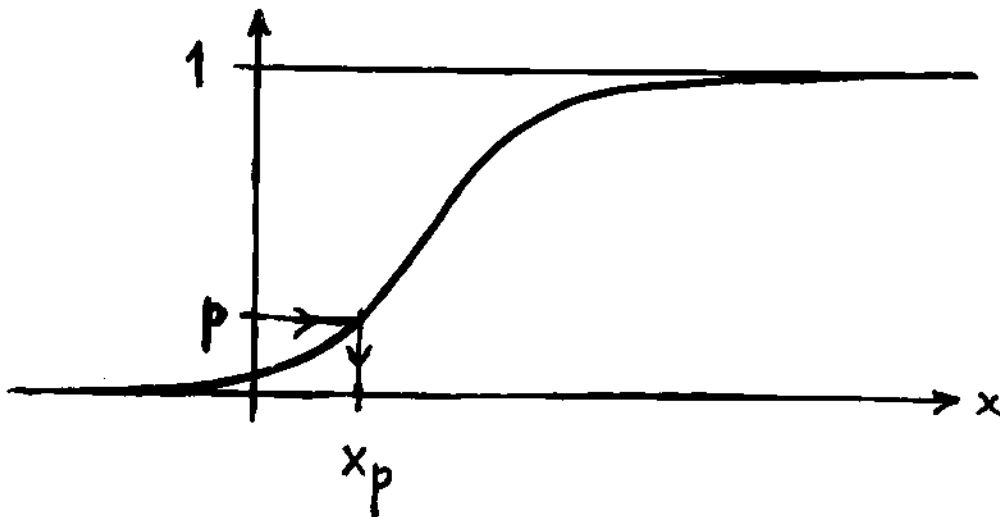
p -Fraktiles x_p : Anteil der Beobachtungen
 $\leq x_p$ ist p

(Abb.)

empirischer Median \tilde{x}

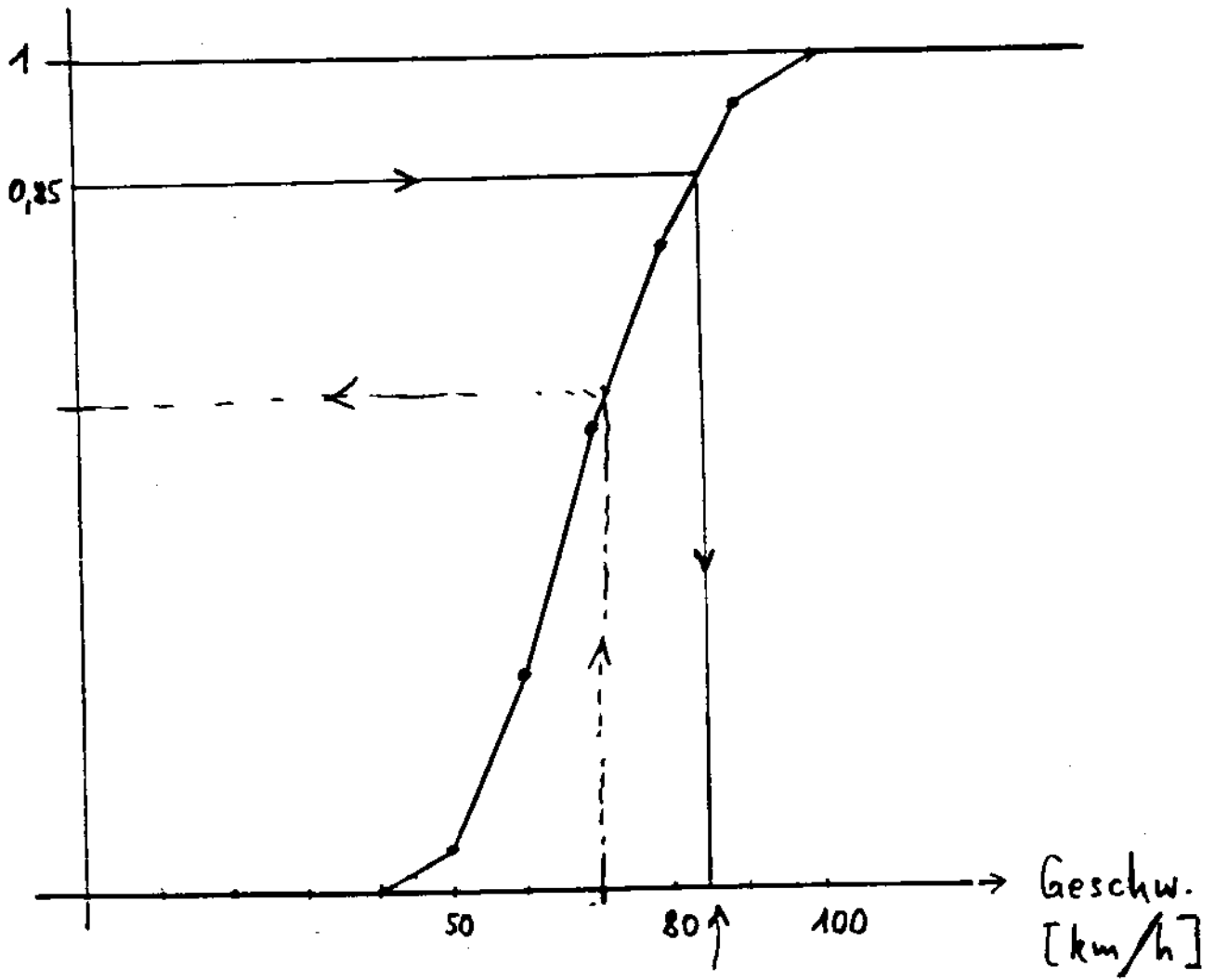


p -Fraktiles x_p



Bsp.: Beobachtete Geschwindigkeitsverteilung

● Summenhäufigkeit ●



$$x_{0,85} = 84 \text{ km/h}$$

Beschränkung 80-90 km/h

•
•
c) Empirische Fraktile

Für $p \in (0,1)$ jener Wert x_p , für den das
Summenpolygon den Funktionswert p hat

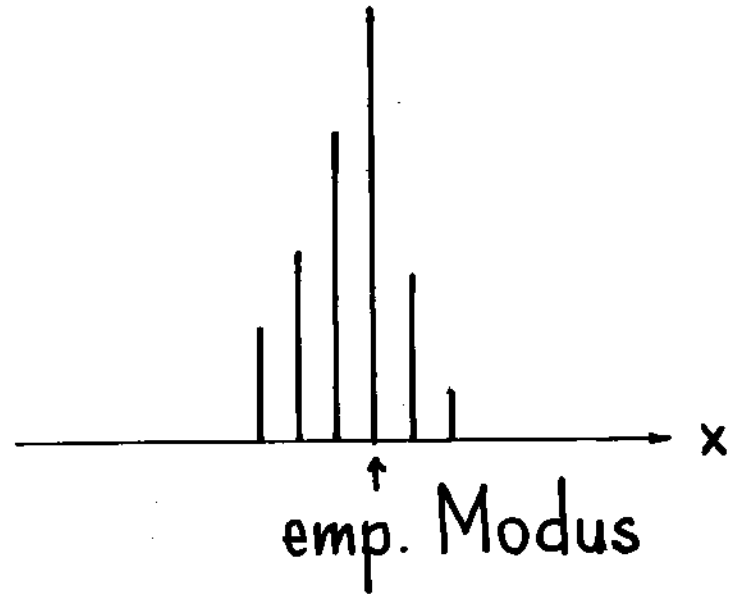
$$S_n(x_p) = p$$

Bem.: Für theoretische Verteilungen
mittels der sog. Verteilungsfunktion

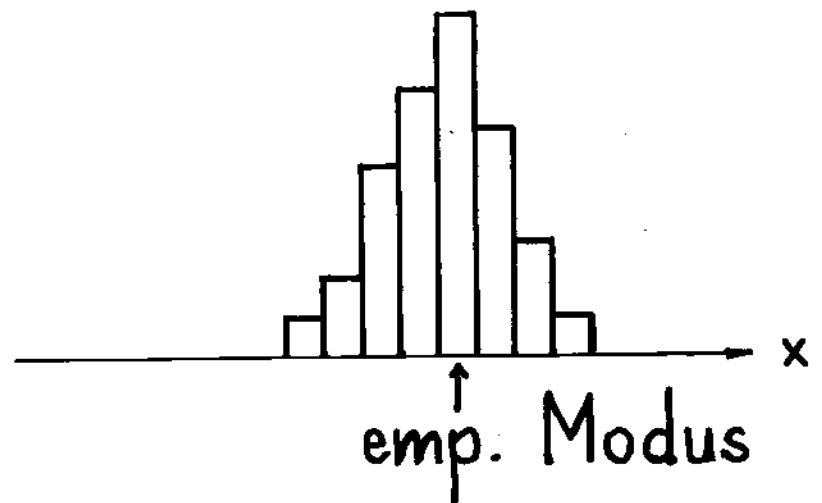
d) Modalwert (= empirischer Modus)

Jener Merkmalwert, der am häufigsten auftritt
(Abb.)

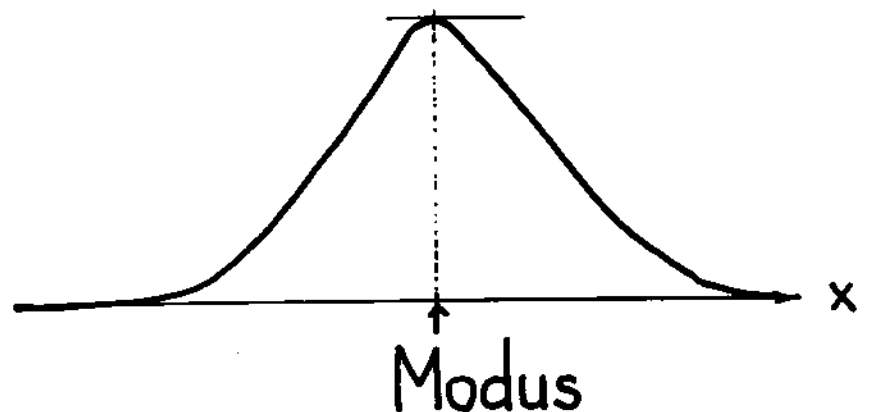
Diskrete Größen



Kontin. Größen



Bem.: Bei theoretischen kontinuierlichen Verteilungen



6.2 Streuungsparameter

Daten x_1, \dots, x_n

a) Spannweite = $x_{(n)} - x_{(1)} = \max_{i=1(1)n} x_i - \min_{i=1(1)n} x_i$

b) Quartilabstand = $x_{0,75} - x_{0,25}$

c) Mittlere absolute Abweichung
 \tilde{x} Median

$$\text{MAD} = \frac{1}{n} \sum_{i=1}^n |x_i - \tilde{x}|$$

mean absolute deviation

● ●
d) Mittlere quadratische Abweichung

(empirische Varianz)

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

e) Empirische Streuung $s := \sqrt{s^2}$

f) Empirischer Variationskoeffizient

$$VK := \frac{s}{\bar{x}}$$

Dimensionsloses Streumaß

Bsp.: Variabilität der Preise in versch. Ländern

g) Fünf-Punkt-Zusammenfassung

X_{\min} , $X_{0,25}$, \tilde{x} , $X_{0,75}$, X_{\max}

