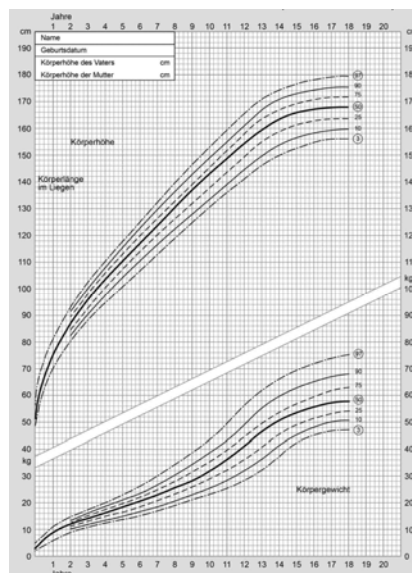
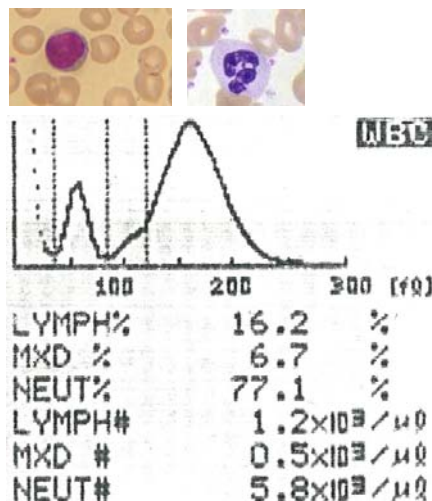


Deskriptive Statistik2



KAD 2018.09.19

Lageparameter. Charakterisierung des Zentrums der Daten

Durchschnittswert (der arithmetische Mittelwert)

=average(...)
 =Mittelwert(...)

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n}$$

Modus (Modalwert, Dichtemittel): der Wert mit der größten Wahrscheinlichkeit; der häufigste Wert einer Häufigkeitsverteilung

=mode(...)
 =Modalwert(...)

Median (Zentralwert): halbiert eine Stichprobe.

Anzahl der Daten der Stichprobe kleiner als Median =
 = Anzahl der Daten der Stichprobe größer als Median

$$x_{\text{med}} = \begin{cases} x_{(n+1)/2} & \text{falls } n \text{ ungerade} \\ (x_{n/2} + x_{(n/2+1)})/2 & \text{falls } n \text{ gerade} \end{cases}$$

=median(...)
 =Median(...)
 2

Durchschnittswert (der arithmetische Mittelwert)

$$x_1 + x_2 + x_3 =$$



$$\sum_{i=1}^n (x_i - \bar{x}) = \sum x_i - \sum \bar{x} = \sum x_i - n\bar{x} = 0$$

Die Summe der Abweichungen der Daten von diesem Wert ist gleich Null.

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n}$$

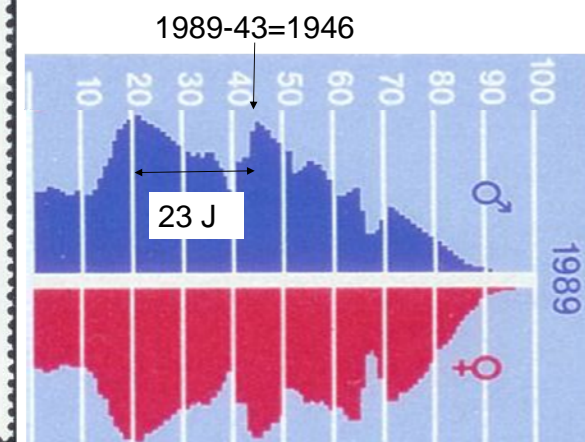
=average(...)
 =Mittelwert(...)

3

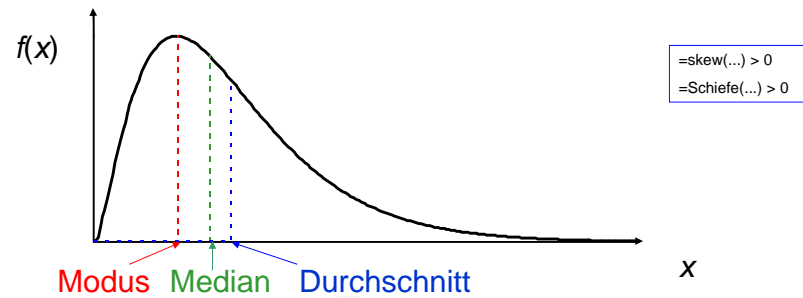
Altersaufbau der deutschen Bevölkerung



Unimodal: die Verteilung hat nur einen Gipfel
 Bimodal: die Verteilung hat zwei Gipfel.
 Multimodal: die Verteilung hat mehrere Gipfel.



Linkssteile bzw. rechtsschiefe Verteilung



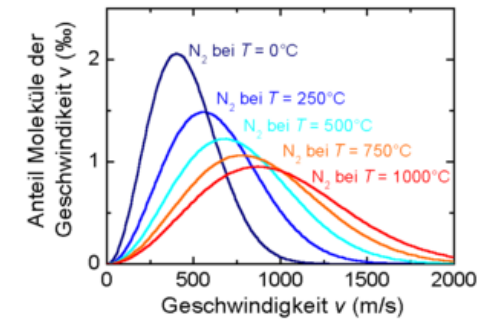
z.B. Einkommensverteilungen in einem Land:

Der Großteil der Bevölkerung verdient relativ wenig, während es nur wenig Leute gibt, die sehr viel verdienen.

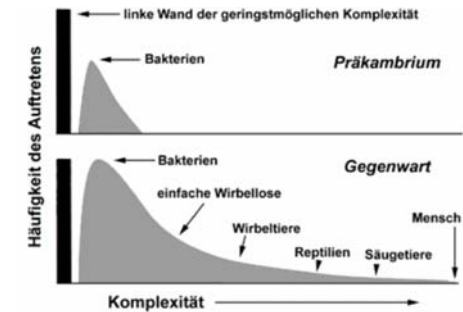
5

Weitere Beispiele

Maxwell-Boltzmann-Verteilung



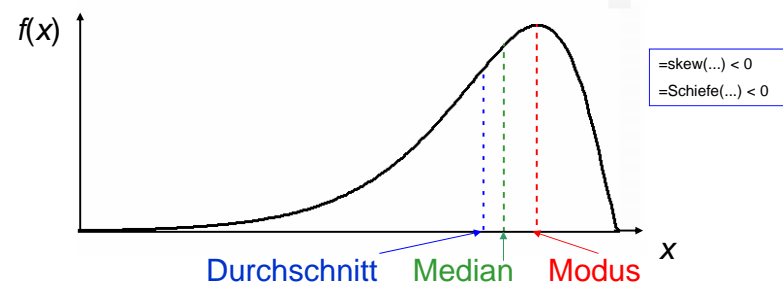
Komplexität der Tiere



www.vordenker.de/if_gould/images/verteilung.gif

6

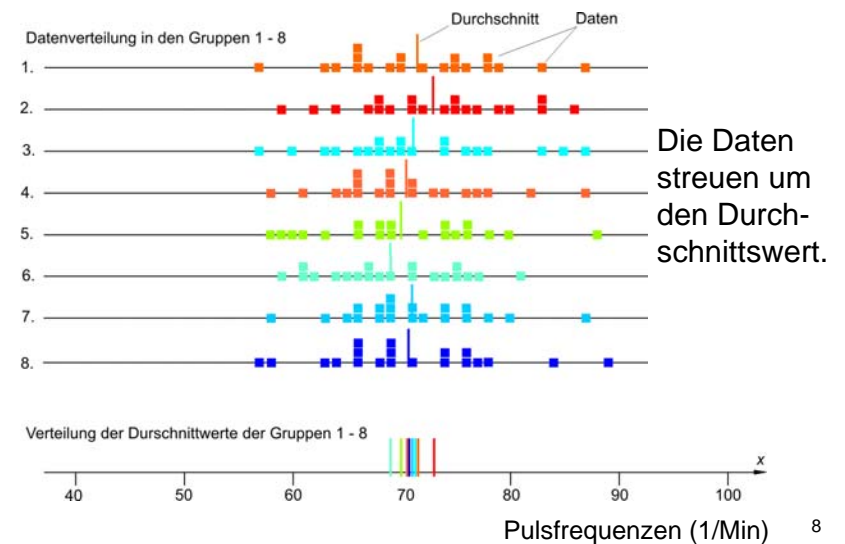
Linksschiefe bzw. rechtssteile Verteilung



z.B. Dauer einer Schwangerschaft



Daten und ihre Durchschnittswerte



Pr.Buch Abb. 10

8

Streuungsparameter. Charakterisierung der Variation der Daten

Standardabweichung

(Streuung der
Messdaten, s):
die mittlere Abweichung
vom Durchschnitt:

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

=stdev(...)
=Stabw(...)

das Quadrat der Streuung,
die mittlere quadratische
Abweichung, auch als

Varianz bezeichnet:

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

=var(...)
=Varianz(...)

Spannweite: $x_{\max} - x_{\min}$

=max(...)-min(...)

9

α -Quantil

$$0 < \alpha < 1$$

(seien dazu die x_i aufsteigend sortiert):

$$x_\alpha = \begin{cases} x_{[n\alpha]+1} & \text{falls } n\alpha \text{ keine ganze Zahl ist} \\ (x_{n\alpha} + x_{n\alpha+1})/2 & \text{falls } n\alpha \text{ ganzzahlig ist} \end{cases}$$

$x_{1/4}$ – unteres Quartil $x_{3/4}$ – oberes Quartil

$x_{1/10}$ – unteres Dezil $x_{9/10}$ – oberes Dezil

=Quantil(...)

halber Quartilabstand : $(x_{3/4} - x_{1/4})/2$

mit Wörter: z.B. **Dezile**

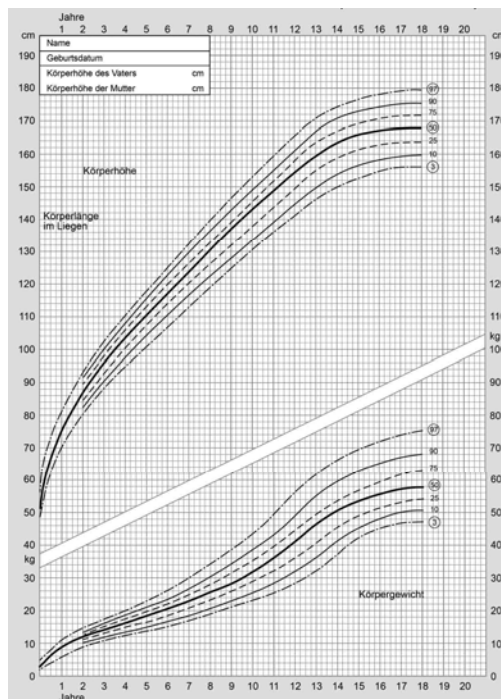
Durch Dezile (lat. „Zehntelwerte“) wird die Verteilung in
10 gleich große Teile zerlegt. Unterhalb des dritten
Dezils liegen 30 % der Verteilung.

10

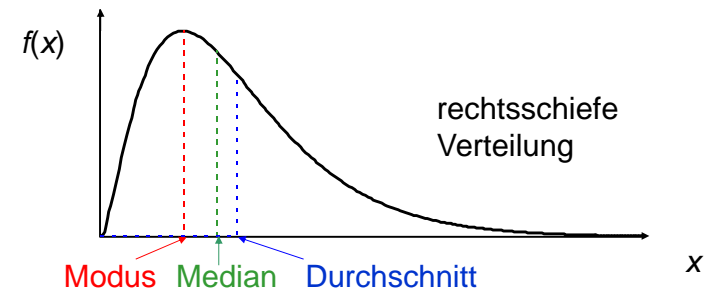
Perzentilenkurven
sind ein Werkzeug
für den Arzt.

Wachstums- und
Gewichtskurven
für Mädchen

=percentile(...)
=Quantil(...)



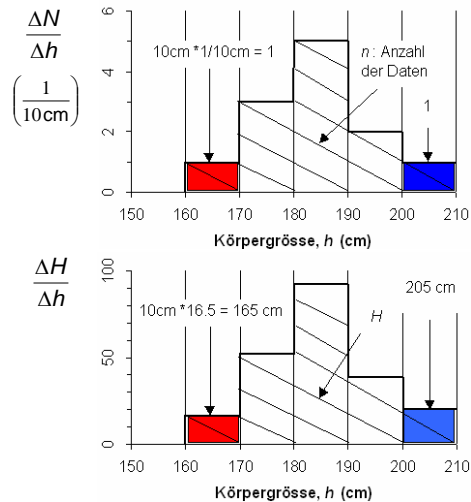
11



Skalentypen	zulässige Lage-Parameter	zulässige Streuungs-Parameter
Nominalskala	Modus	–
Ordinalskala	Modus, Median	–
numerische Skalen	Modus, Median, Durchschnittswert	Spannweite, Quartilabstand, Standardabweichung

12

Häufigkeitsverteilung



h : Körperhöhe

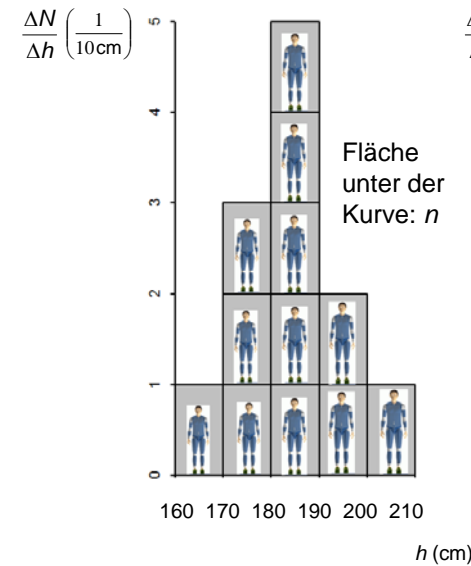
H : kollektive Höhe, Gesamthöhe



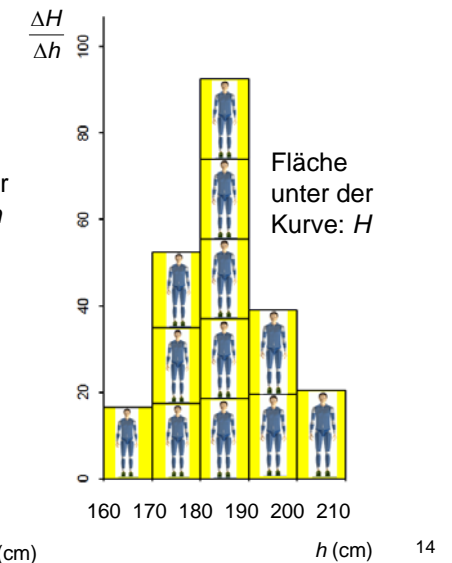
Spektrum als eine spezielle Häufigkeitsverteilung

13

Häufigkeitsdichte



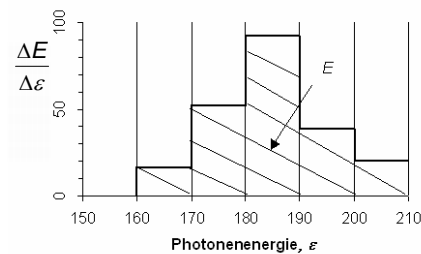
Spektrum



14

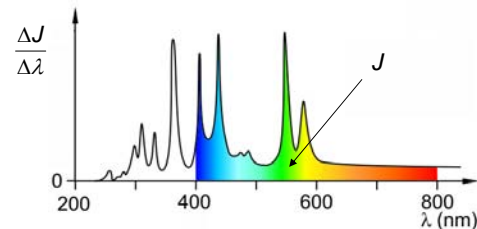
Emissionsspektrum:

wie verteilt sich die emittierte Energie über die Photonenenergien



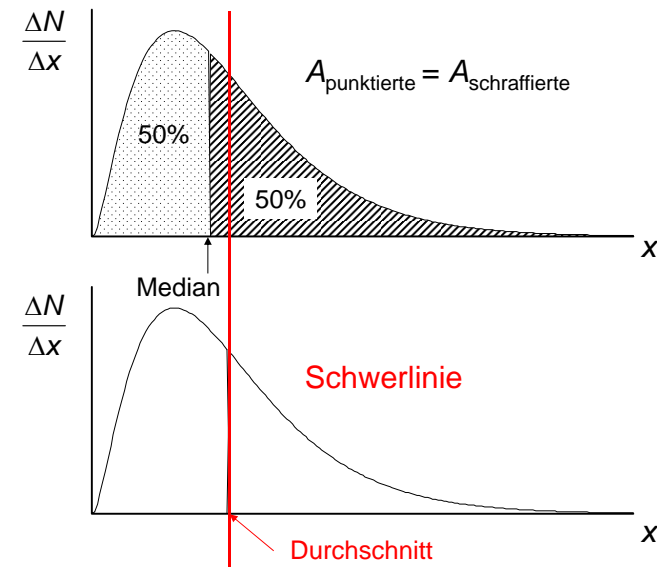
charakteristische Größe des Energietransports:
Intensität

Benützung der **Wellenlänge** ist bequemer als die der Photonenenergie



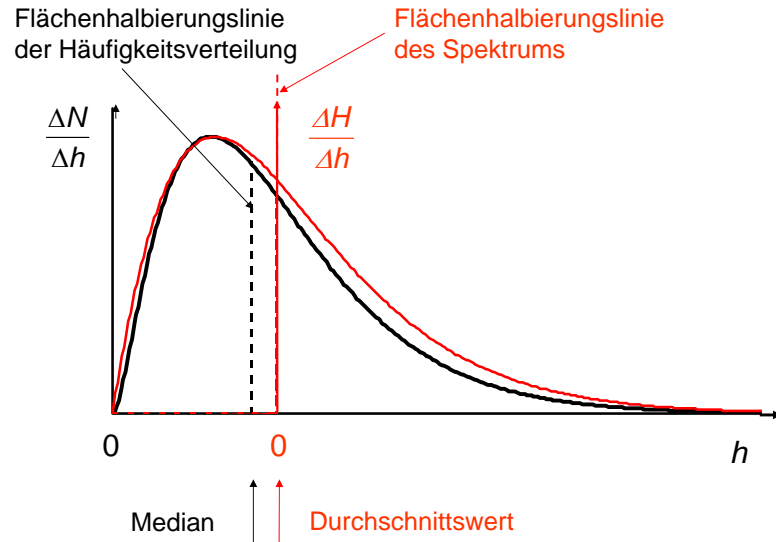
15

Position des Medians und des Durchschnitts einer Verteilung (1)



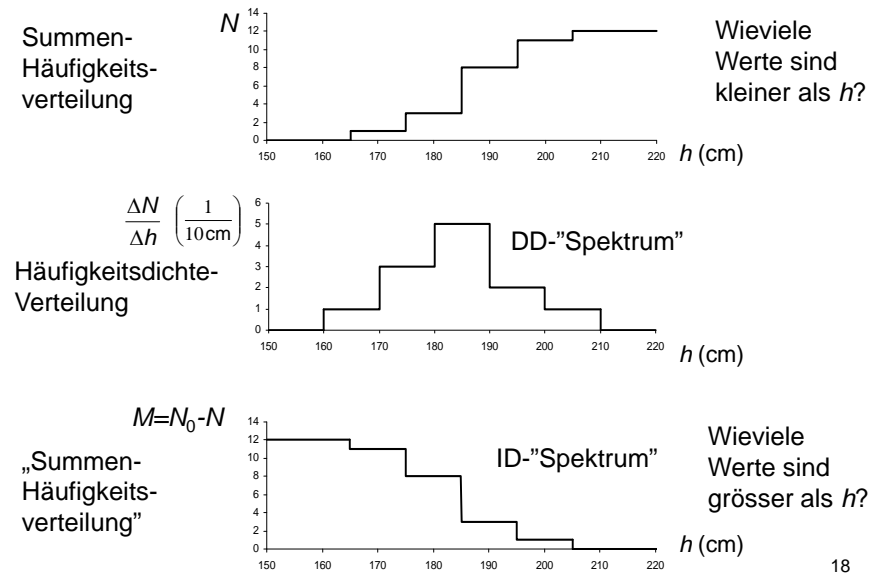
16

Position des Medians und des Durchschnitts einer Verteilung (2)



17

Summen- (kumulierte/kumulative) Häufigkeitsverteilung

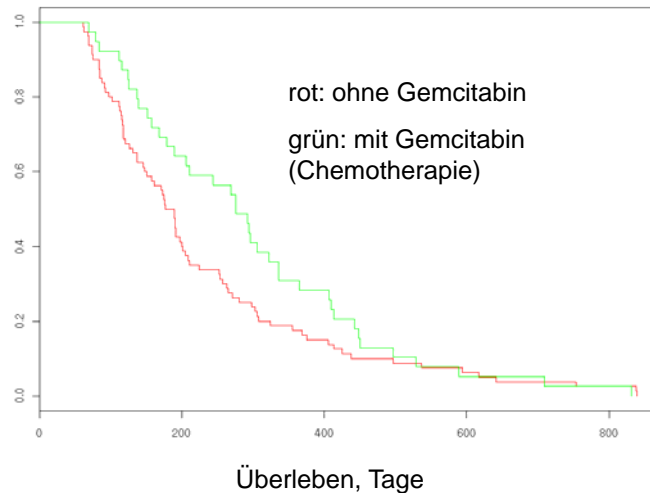


18

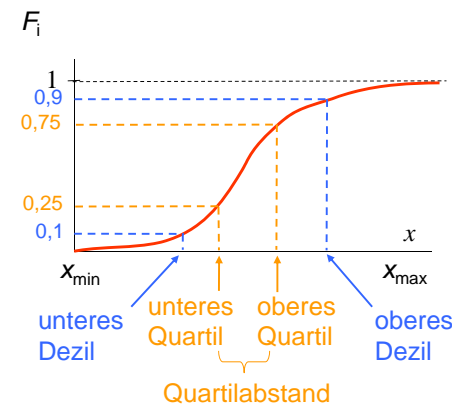
relative
„Summen-
Häufigkeits-
verteilung“

Überlebenskurven

Wirkung der Chemotherapie. Pankreaskarzinom

kumulatives
Überleben
nach der
Operation

Quantile und die relative Summenhäufigkeitsverteilung



Coulter Zähler

