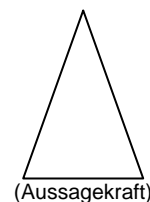


Deskriptive Statistik

Skalierungen:

Nominalskala	Namen, keine Reihenfolge
Ordinalskala	Bezeichnung, auf- oder absteigende Folge
Metrische- / Kardinalskala	zahlenmässig ausdrückbar (quantitativ)
Intervallskala	Intervalle (einfache Abstände)
Verhältnisskala	Verhältnis und Intervalle



Häufigkeiten:

n	Gesamtzahl der Merkmalswerte	
v	Anzahl verschiedener Merkmalswerte (bzw. Merkmalsträger)	
x	Wert	
h_i	absolute einfache Häufigkeit	
f_i	relative einfache Häufigkeit	$= \frac{h_i}{n}$
H_i	absolute <i>kumulierte</i> Häufigkeit	$= \sum_{k=1}^i h_k$ ($= h_1 + h_2 + h_3 \dots$)
F_i	relative <i>kumulierte</i> Häufigkeit	$= \sum_{k=1}^i f_k = \frac{H_i}{n}$ ($= f_1 + f_2 + f_3 \dots$)
HR_i	(Resthäufigkeit)	$HR_i = n - H_i$
FR_i		$FR_i = 1 - F_i$
$F(x)$	(relative kumulierte Häufigkeit) für einen Wert x aus der j-ten Klasse	$F(x) = F_{j-x} + \frac{x - x_j^u}{x_j^o - x_j^u} * (F_j - F_{j-1})$
d_j	Häufigkeitsdichte	$d_j = \frac{h_j}{x_j^o - x_j^u} = \frac{h_j}{\text{Klassenbreite}}$

Klassen:

j	Laufindex für Klassen
x_j^u	Untergrenze der Klasse j
x_j^o	Obergrenze der Klasse j

Darstellungsarten: (Diagramme)

Säulen	vertikale Balken
Balken	horizontale Balken
Treppenfunktion	(für Ordinalskalen) Stufen, jeweils mit H_i
Kreis	Flächenproportional (Kreis: $r^2 * \pi$) doppelte Menge = $2 * (r^2 * \pi)$
Histogramm	Balkenhöhe gem. Häufigkeitsdichte (Klassenbreite!)
Polygonzug	Linie analog Histogramm (Punkte d_j bei Klassenmitte)
	(linker Rand: $\frac{d_j}{2}$ der ersten Klasse; rechter Rand: $\frac{d_j}{2}$ der letzten Klasse)
Summenpolygon	linker Rand: 0; Linie mit Punkten am Ende der Klassen; Punkte gem. H_i
Ogive	analog Summenpolygon aber Punkte gem. F_i empirische Verteilungsfunktion

Lageparameter:**Mo = Modus**

(häufigster Wert)

einzigster möglicher Mittelwert für Nominalskalen, ggf. zwei Modi möglich
Klasse mit grösster Klassenhäufigkeit h_j ,
bzw. bei unterschiedlichen Klassenbreiten mit grösster d_j

$$Mo = x_m^u + \left[(x_m^o - x_m^u) * \frac{h_m - h_{m-1}}{(h_m - h_{m-1}) + (h_m - h_{m+1})} \right]$$

(Mo = Klassenuntergrenze + Klassenbreite * Prozentsatz)

Me = Median

(zentraler Wert)

ohne Klassen:**Ordinalskala zwingend**; Merkmalsträger müssen in Rangordnung seinungerade Anzahl Merkmalsträger (n):

$$Me = x_{\left[\frac{n+1}{2}\right]} \quad (\text{Wert des Merkmalsträgers in der Mitte})$$

gerade Anzahl Merkmalsträger (n):

$$Me = \frac{1}{2} * \left(x_{\left[\frac{n}{2}\right]} + x_{\left[\frac{n}{2}+1\right]} \right) \quad (\text{Durchschnitt der Werte der beiden zentralen Merkmalsträger})$$

mit Klassen:

1. Medianklasse bestimmen: $\frac{n}{2}$ (Medianklasse ist die Klasse, in der dieser Wert liegt)

(für $n = H_j$ der höchsten Klasse einsetzen; also die Totalsumme, bzw. kumm.Häufigkeit)

2. Me mit dieser Klasse bestimmen:

$$Me = x_m^u + \left(\frac{\frac{n}{2} - H_{m-1}}{H_m - H_{m-1}} \right) * (x_m^o - x_m^u)$$

Quantile

(dezimal)

Aufteilung nicht hälftig (50%) wie bei Median sondern z.B. D_x =Dezantil

oder Perzentile (Hundertstel)

1. Medianklasse bestimmen: $\frac{z}{10} * n$ (Medianklasse ist die Klasse, in der dieser Wert liegt)

(für $n = H_j$ der höchsten Klasse einsetzen – also die Totalsumme, bzw. kumm.Häufigkeit
→ also entspricht dies $z\%$ von H_j der höchsten Klasse)

$$D_z = x_m^u + \left(\frac{\left(\frac{z}{10} * n \right) - H_{m-1}}{H_m - H_{m-1}} \right) * (x_m^o - x_m^u)$$

Quartile

Aufteilung in Viertel (Medianklasse wie oben) (Beispiel für 25%, bzw. 1 Viertel)

$$Q_1 = x_m^u + \left(\frac{\left(\frac{1}{4} * n \right) - H_{m-1}}{H_m - H_{m-1}} \right) * (x_m^o - x_m^u)$$

\bar{x} = **arithmetisches Mittel** **Intervallskala zwingend**; (für $n = H_j$ der höchsten Klasse einsetzen)
ohne Klassen:

$$\bar{x} = \frac{\sum_{i=1}^v x_i * h_i}{n} = \sum_{i=1}^v x_i * f_i \quad \text{bzw.} \quad \frac{(h_1 * \text{Wert}) + (h_2 * \text{Wert}) + \dots}{n}$$

mit Klassen:

$$\bar{x} = \frac{\sum_{i=1}^v x_j^l * h_i}{n} = \sum_{i=1}^v x_j^l * f_i \quad \text{bzw.} \quad \frac{(\text{Klassenmitte}_1 * \text{Wert}) + (\dots) + \dots}{n}$$

F_{GM} = **geometrisches Mittel** **Verhältnisskala zwingend**
 geeignet für den Vergrößerungs- / Verkleinerungsfaktor (Wachstum)
 (n = Anzahl Wachstumsfaktoren = Anzahl Intervalle)

Variante 1:

$$F_{GM} = \sqrt[n]{\frac{\text{Endwert}}{\text{Anfangswert}}}$$

Variante 2:

$$MG = \sqrt[n]{F_1 * F_2 * \dots * F_n} = \sqrt[n]{\prod_{i=1}^n F_i}$$