1 Merkmale

1.1 Qualitativ / Kategoriell

Es wird eine Ausprägung und kein Ausmass angegeben. Insbesondere gibt es nur endlich viele Ausprägungen.

1.1.1 Nominal

Keine Kategorisierung, keine Ordnung

1.1.2 Ordinal

Ordnung vorhanden, Rangierung möglich

1.2 Quantitativ / Metrisch

Es wird ein Ausmass angegeben. Das Ausmass wird mit Zahlen angegeben.

1.2.1 Quantitativ

Diskret endlich viele aber abzḧlbar unendlich viele Ausprgünen.

1.2.2 Stetig

Alle Ausprägungen in einem reelen Interval

2 Häufigkeiten

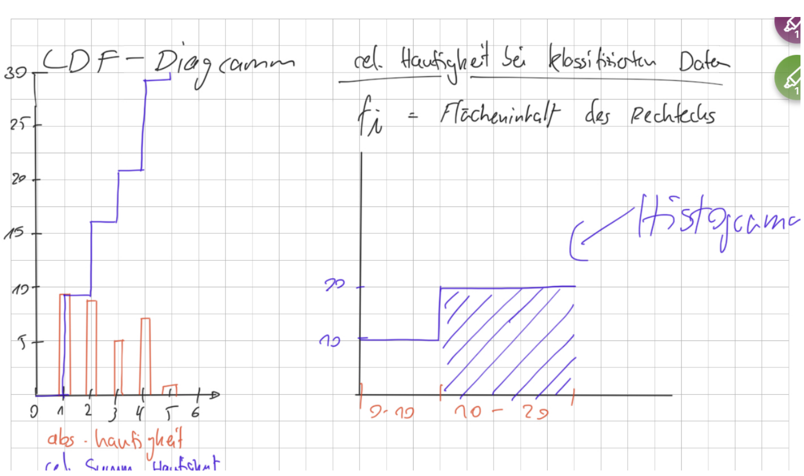
2.1 Graphische Darstellung

Kategoriell: Säulendiagramm, Skalardiagramm
Metrisch: Säulendiagramm, Skalardiaramm, Histogramm bei Klassenbildung

Beschreibung	Zeichen	Formel
Anzahl	n	
Wert / Klasse	a_i	
Absolute Häufigkeit	h_i	Anzahl
Relative Häufigkeit (PMF)	f_i	$\frac{h_i}{n}$
Kummulative rel. Häufigkeit (CDF)	F_i	Summe aller vorgehenden f_i
Spaltenbreite	d_i	
Spaltenhöhe	h	$\frac{h_i}{d_i}$
PDF-Wert	f	$\frac{f_i}{d_i}$

Auswahl von Bereichen: $F(2 \leq 4) = F(4) - F(1)$

2.2 Diagramme



3 Varianz

3.1 Arithmetisches Mittel

Das Arithmetische Mittel wird auch als Durchschnitt bzwh. empirisches Mittel bezeichnet.
$\bar{x} = \sum_{i=1}^{n} a_i * f_i$

3.1.1 Bei klassierten Daten

$\bar{x} = \sum_{i=1}^{n} x_i$ mit x_i als Klassenmitte

3.2 Empirische Varianz / Stichprobenvarianz

Beschreibt die mittlere quadratische Abweichung der einzelnen Werte vom empirischen Mittels.
$\tilde{s}^2 = \frac{1}{n} \cdot \sum_{i=1}^{n}((a_i - \bar{x})^2 * h_i)$
alternativ
$\tilde{s}^2 = \frac{1}{n} \cdot \sum_{i=1}^{n}(x_i - \bar{x})^2$

3.3 Standardabweichung

Die Standardabweichung ist ein Maß für die Streubreite der Werte eines Merkmals rund um dessen Mittelwert. Vereinfacht gesagt, ist die Standardabweichung die durchschnittliche Entfernung aller gemessenen Ausprägungen eines Merkmals vom Durchschnitt.

$\bar{s} = \sqrt{\bar{s}^2}$

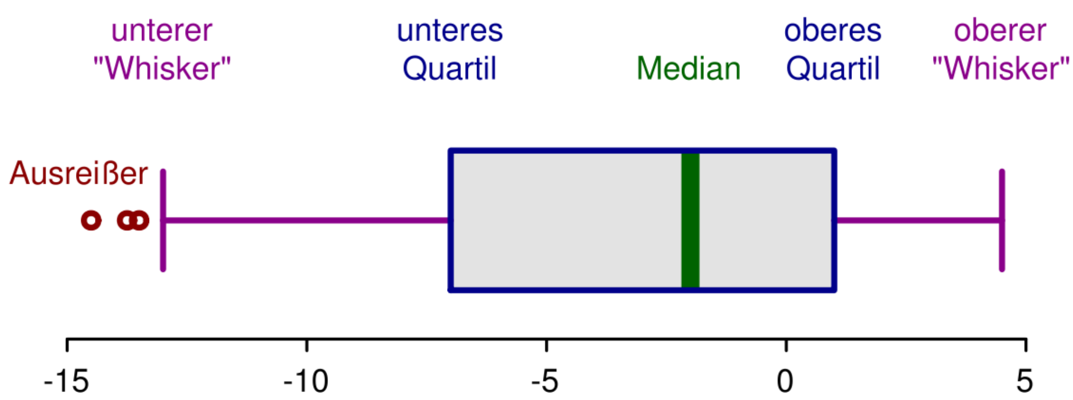
3.4 Korrigierte empirische Varianz

$\tilde{s}^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n}((a_i - \bar{x})^2 * h_i)$

3.5 Korrigierte Standardabweichung

$s = \sqrt{s^2}$

4 Boxplot



Beschreibung	Formelzeichen	Berechnung
Unterer Whisker		maximal, sonst Daten $Q_1 - 1.5 * I_{QR}$
Obrerer Whisker		maximal, sonst Daten $Q_3 + 1.5 * I_{QR}$
0.25er Quantil	Q_1	
0.75er Quantil	Q_3	
Interquartilsabstand	I_{QR}	
Median	X_{med}	
Modulo Wert	X_{mod}	Höchster Stichprobenwert (Spitze der Kurve)

4.1 Quantil

4.1.1 Wenn $n \cdot q$ eine ganze Zahl ist

$R_q = \frac{1}{2} \cdot (X_{n \cdot q} + X_{n \cdot q+1})$

4.1.2 Wenn $n \cdot q$ keine Ganze Zahl ist

$X_{|n \cdot q|}$ mit $n \cdot q$ die nächstgrösste Zahl

4.2 Boxplot in einer Klasse

1. CDF in der Verteilung berechnen
2. Klasse auswählen für die $F(a) \leq q < F(b)$

$\frac{F(b)-F(a)}{b-a} = \frac{q-F(a)}{R_q-a}$

Nach R_q umgestellt

$R_q = a + \frac{b-a}{F(b)-F(a)}[q - F(a)]$

4.3 Median

Der in der Mitte liegende Wert einer sortierten Datenmenge

4.3.1 Wenn n gerade

$X_{med} = x[\frac{n+1}{2}]$

4.3.2 Wenn n ungerade

$X_{med} = \frac{1}{2}[X_{[\frac{n}{2}]} + X_{[\frac{n}{2}+1]}]$

5 Elementare Wahrscheinlichkeitsrechnung

Eigenschaft	Symbol	Formel	Beschreibung
Ergebnisraum	Ω		
Zieldichte	ρ	$\rho : \Omega \to [0, 1]$	$\sum_{\omega \in \Omega} \rho(\omega) = 1$
Ereignis	A	Teilmenge von Ω	Leere Menge entspricht dem unmöglichen Ereignis
Ereignisraum	2^{Ω}		Menge aller möglichen Ereigniss. Ω
Wahrscheinlichkeitsmass	P	$P : 2^{\Omega} \to [0, 1]$	$sum_{\omega \in M} \rho(\omega) = 1, M \subseteq \Omega$

5.1 Kenngrössen

Mittelwert: $\mu = \frac{m}{n}$

5.2 Bedingte Wahrscheinlichkeiten

Wahrscheinlichkeit für das Eintreten des Eeignisses B unter der Bedingung oder Voraussetzung, dass das Ereignis A eintreitt.
$P(B|A)$

5.2.1 Wahrscheinlichkeitsbaum

- Alle von einem Blatt (Verzweigungspunkt) aus mit Pfeilen erreichbare Ereignisse sind paarweise disjunkt (d.h. sie schliessen sich gegenseitig aus)
- Die Summe der Übergangswahrscheinlichkeiten aller von einem Blatt ausgehenden Pfeile ist 1.

level 1=[level distance=1.5cm, sibling distance=3.5cm] level 2=[level distance=1cm, sibling distance=2cm] [end] child node[bag] A child node [bag] Z edge from parent node[left] 0.9 child node [bag] $\overline{Z}$ edge from parent node[right] 0.1 edge from parent node[left] 0.5 child node[bag] B child node [bag] Z edge from parent node[left] 0.7 child node [bag] $\overline{Z}$ edge from parent node[right] 0.1 edge from parent node[left] 0.3 child node[bag] C child node [bag] Z edge from parent node[left] 0.4 child node [bag] $\overline{Z}$ edge from parent node[right] 0.6 edge from parent node[right] 0.2
$P(A \cup Z) = P(A) * P(Z|A) = 0.5 * 0.9 = 0.45$
Allgemein gilt: Wahrscheinlichkeiten längs eines Pfades werden multipliziert.
Multiplikationssatz: $P(A \cup B) = P(A) * P(B|A) = P(B) * P(A|B)$
Satz von der Totalen Wahrscheinlichkeit: $P(B) = P(A) * P(B|A) + P(\overline{A}) * P(B|\overline{A})$
Satz von Bayes: $P(B|A) = \frac{P(A \cap B)}{P(A)}$

5.3 Stochastische Unabhängigkeit

Wenn A und B stochastisch unabhängig sind, beeinflusst das Eintreten des inene Ereignisses also das Eintreten des anderen Ereignisses nicht.

6 Diskrete und stetige Verteilung