

1. Kapitel 1-3

Aufgabe 1

Die folgende Tabelle enthält die Merkmalswerte zweier Merkmale X und Y an 5 Merkmalsträgern. Bestimmen Sie (ohne einen Taschenrechner zu benutzen!) die arithmetischen Mittelwerte, Varianzen, Mediane und oberen Quartile von X und Y.

i	1	2	3	4	5
x_i	0	1	2	1	2
y_i	1	0	1	2	3

Aufgabe 2

Am 11. Juli 2012 gab das Statistische Bundesamt bekannt, dass 20% der Bevölkerung in Deutschland allein im Haushalt leben, also einen Single-Haushalt bilden. Andererseits seien 40,4% aller Haushalte derartige Single-Haushalte. Bestimmen Sie aus diesen Informationen die durchschnittliche Haushaltsgröße!

Aufgabe 3

Für die Messung von (wirtschaftlicher) Armut gibt es mehrere Definitionen. Häufig wird ein Haushalt (vereinfacht) als arm definiert, wenn das Haushaltseinkommen weniger als 60% des Medianhaushaltseinkommens beträgt. Alternativ könnte man einen Haushalt als arm definieren, wenn das Haushaltseinkommen weniger als 60% des arithmetischen Mittels aller Haushaltseinkommen beträgt. Erläutern Sie, nach welcher dieser beiden Armutsdefinitionen in Deutschland mehr Haushalte als arm bezeichnet würden!

Aufgabe 4

Gegeben sei ein Merkmal X in einer Gesamtheit vom Umfang n mit arithmetischem Mittel \bar{x} und Standardabweichung s . Das Merkmal Y sei definiert durch: $y_i = \frac{x_i - \bar{x}}{s}$ für $i = 1, \dots, n$. Zeigen Sie: das arithmetische Mittel von Y beträgt 0, und die Varianz von Y beträgt 1.

Aufgabe 5

Aufgabe 5 Zeigen Sie: Für eine gegebene Häufigkeitsverteilung eines metrischen Merkmals X mit Ausprägungen x_1, x_2, \dots, x_k und relativen Häufigkeiten f_1, f_2, \dots, f_k besitzt die Funktion:

$$f : \mathbb{R} \rightarrow \mathbb{R}; t \rightarrow f(t) = \sum_{i=1}^k (x_i - t)^2 * f_i$$

in $t_0 = \bar{x}$ (arithmetisches Mittel von X) das globale Minimum.

- a; Geben Sie die tatsächlich vorkommenden Merkmalsausprägungen und die jeweiligen relativen Häufigkeiten an!
- b; Welcher Anteil der Mütter ist älter als 26 Jahre?

Aufgabe 6

Die folgende empirische Verteilungsfunktion beschreibt die Verteilung des Merkmals Älterin der aktuellen Gesamtheit aller Mütter auf der Entbindungsstation einer großen Klinik.

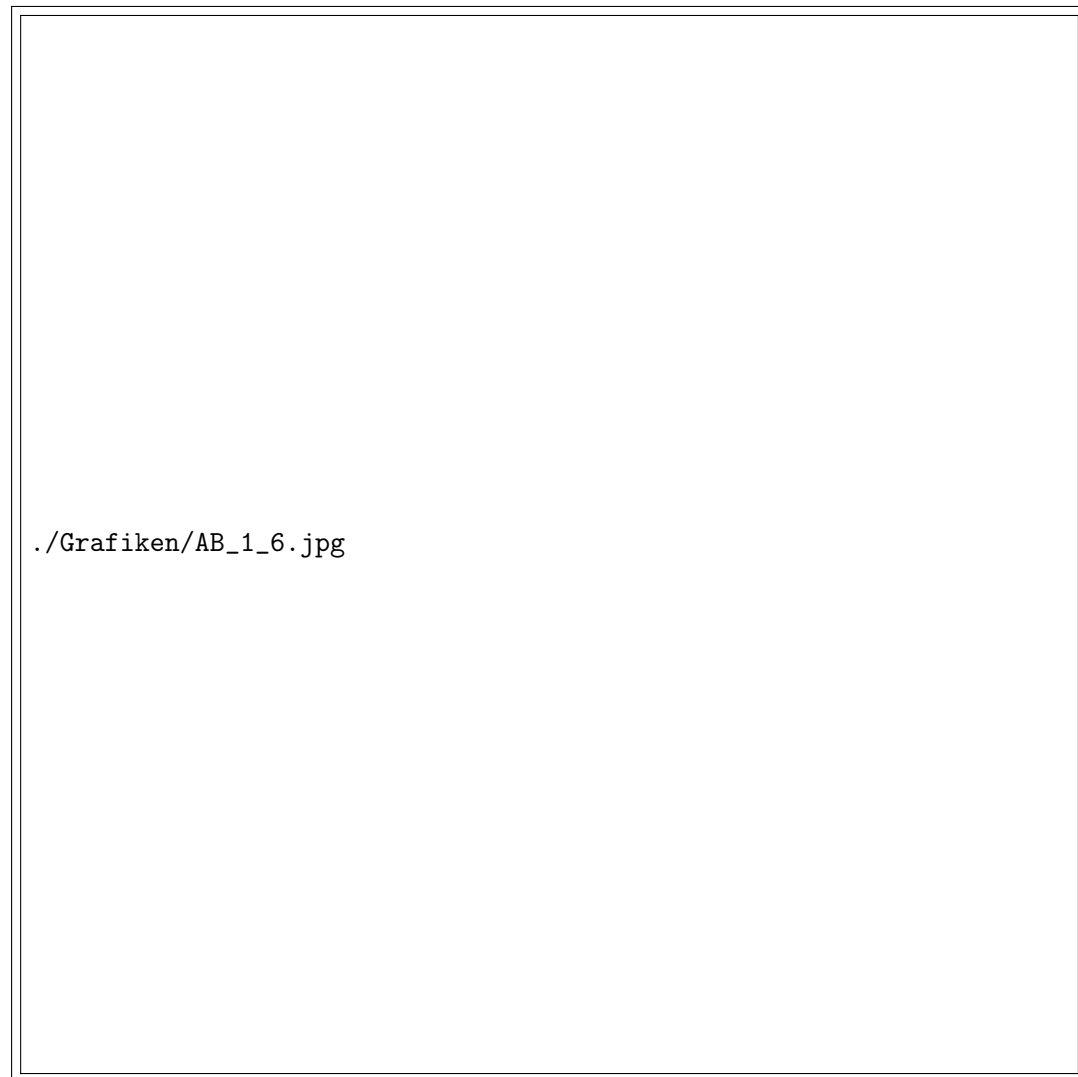


Abbildung 1: Erklärung

- a) Geben Sie die tatsächlich vorkommenden Merkmalsausprägungen und die jeweiligen relativen Häufigkeiten an!
- b) Welcher Anteil der Mütter ist älter als 26 Jahre?

Aufgabe 7

Ein Merkmal X besitze nur die beiden Ausprägungen 0 und 1. Die relative Häufigkeit der Ausprägung 1 wird in diesem Fall auch mit $\pi(\text{Anteil, Anteilswert})$ bezeichnet. Zeigen Sie, dass in diesem Fall für die Varianz von X gilt:

Aufgabe 8

Das folgende Säulendiagramm zeigt die Verteilung des Merkmals "Kinderzahl" für alle Haushalte eines bestimmten Wohnviertels.

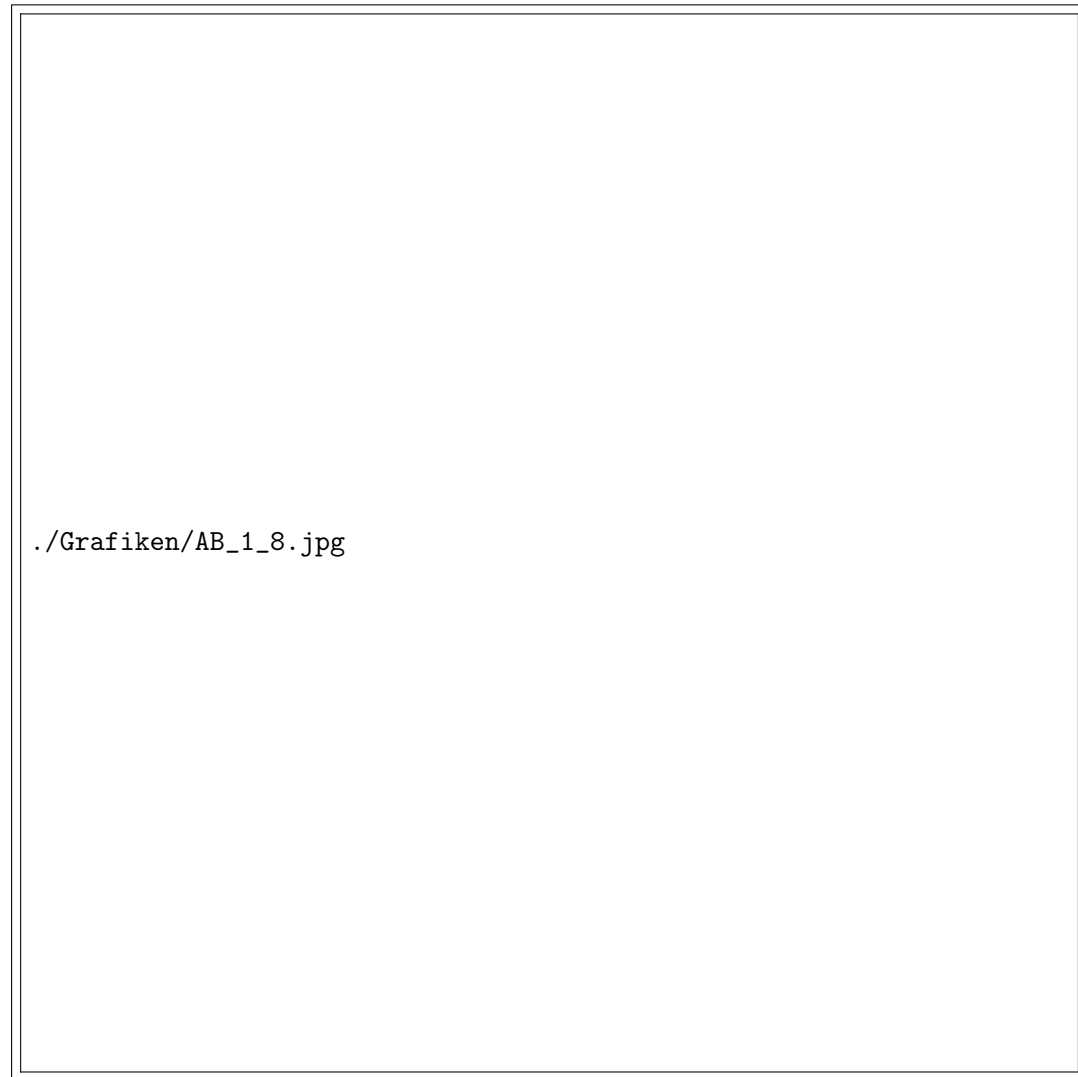


Abbildung 2: Erklärung

- a) Berechnen Sie für diese Verteilung das arithmetische Mittel, die Varianz, die Spannweite, den Median und den Quartilsabstand! Skizzieren Sie den zugehörigen Boxplot!
- b) Im daneben liegenden Wohnviertel mit 100 Haushalten gibt es insgesamt 250 Kinder; die Varianz des Merkmals "Kinderzahl" beträgt unter diesen Haushalten 3. Bestimmen Sie das arithmetische Mittel und die Varianz dieses Merkmals in der Gesamtheit aller Haushalte, wenn man die beiden Wohnviertel zusammen betrachtet!

Aufgabe 9

Sechs Schüler haben eine Prüfung absolviert; jede der (ganzzahligen) Noten von 1 bis 6 wurde dabei vergeben. Bestimmen Sie mit dem Taschenrechner oder einer Software Ihrer Wahl die Varianz des Merkmals "Prüfungsnote". Versuchen Sie durch Ausprobieren herauszufinden, bei welcher Notenverteilung die Varianz maximiert worden wäre!

Aufgabe 10

Die so genannte Lebenserwartung einer Personengruppe ist (vereinfacht) das arithmetische Mittel der (prognostizierten) vollendeten Lebensjahre zum Sterbezeitpunkt der einzelnen Personen (Sterbealter in ganzen Jahren). Nach einem Zeitungsartikel lag 1871 die Lebenserwartung eines Neugeborenen (im damaligen Deutschen Reich) bei 38 Jahren; 230 von 1000 Säuglingen starben dabei schon während des ersten Lebensjahres. Im Jahr 2011 lag die Lebenserwartung eines Neugeborenen in Deutschland bei 80 Jahren; nur 4 von 1000 Säuglingen starben während des ersten Lebensjahres. Wie hat sich die Lebenserwartung von Personen, die das erste Lebensjahr überlebt haben, von 1871 bis 2011 verändert?

Aufgabe 11

Die folgende approximierende Verteilungsfunktion beschreibt die Verteilung des Merkmals Älterin der aktuellen Gesamtheit aller Mitglieder des Kreisverbands einer Partei. Zur Bestimmung dieser Verteilungsfunktion wurde das Merkmal in sechs Klassen eingeteilt.

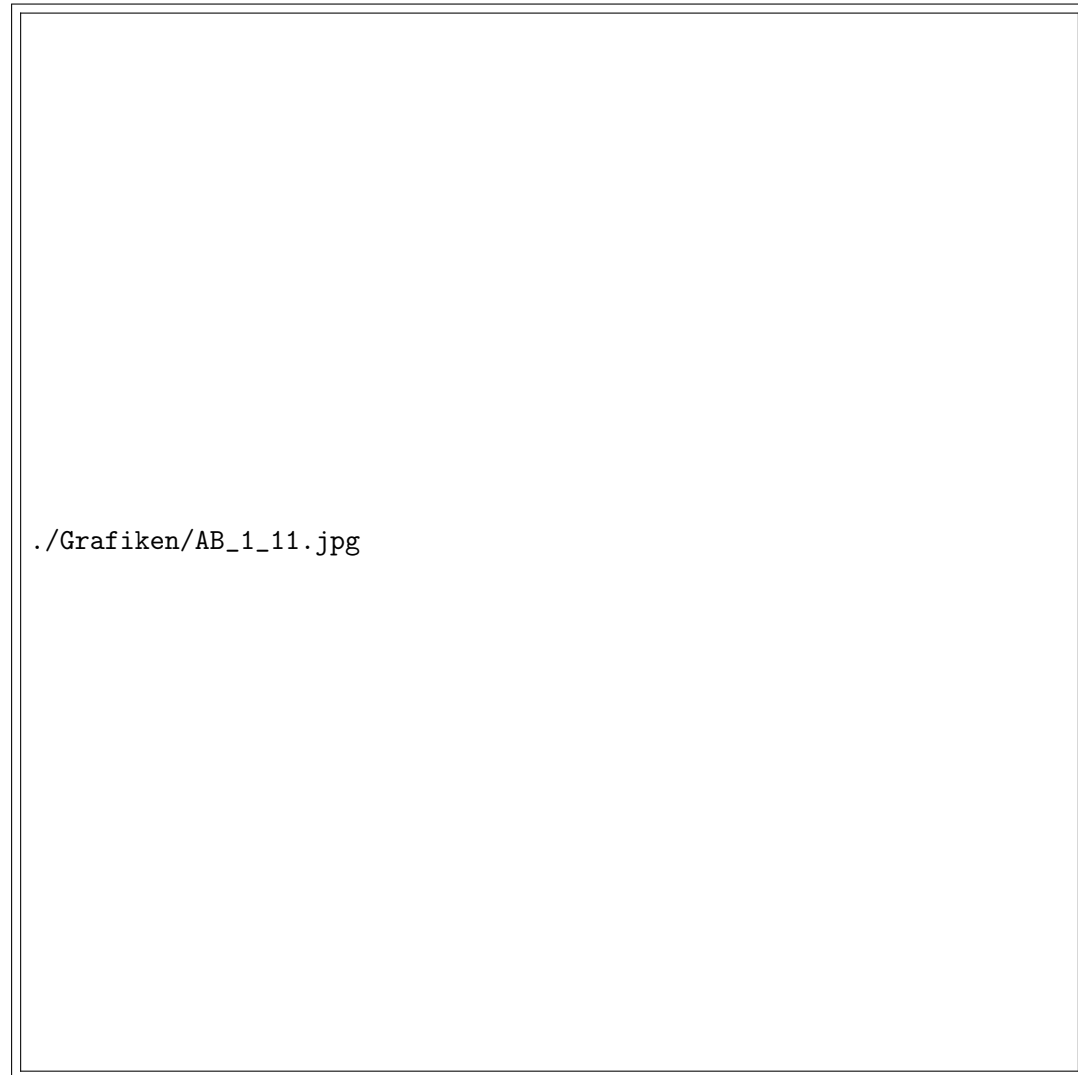


Abbildung 3: Erklärung

- a) Geben Sie die Klassengrenzen und die jeweiligen relativen Häufigkeiten innerhalb der Klassen an!
- b) Bestimmen Sie approximativ anhand der Grafik (keine Rechnung nötig!) den Median dieser Verteilung!
- c) Welcher (approximative) Anteil der Parteimitglieder ist 70 Jahre oder älter? (Keine Rechnung nötig!)

Aufgabe 12

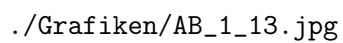
Ein metrisches Merkmal X wird in zwei verschiedenen Teilgruppen beobachtet. In beiden Teilgruppen wurden Minimum, Maximum, Median und arithmetisches Mittel berechnet:

Minimum	Maximum	Median	arithmetisches Mittel	Größe der Teilgruppe
0	4	3	2	20
3	9	5	6	30

Berechnen Sie, falls möglich, aus den angegebenen Werten das gesamte arithmetische Mittel, den gesamten Median, die gesamte Spannweite und die gesamte Varianz von X , wenn die 50 betrachteten Merkmalsträger als Gesamtheit angesehen werden. Falls Ihnen die Berechnung einiger oder aller Maßzahlen nicht möglich ist, geben Sie an, welche zusätzlichen Informationen Sie dazu jeweils bräuchten. Geben Sie für alle Maßzahlen, die Sie nicht berechnen können, möglichst kleine Intervalle an, in denen die Maßzahlen mit Sicherheit liegen.

Aufgabe 13

Das folgende Stabdiagramm zeigt die Verteilung des Merkmals $X = \text{Anzahl der Ausfälle in einem Netzwerk pro Woche}$ innerhalb eines bestimmten Zeitraums.



./Grafiken/AB_1_13.jpg

Abbildung 4: Entity-Relation Diagramm der Datenbank

- a) Bestimmen Sie die beiden Quartile und den Median von X !
- b) Insgesamt sind im betrachteten Zeitraum 217 Ausfälle beobachtet worden. Die Summe der quadrierten Ausfallzahlen (pro Woche) beträgt 1155. Bestimmen Sie daraus die Varianz von X !

Aufgabe 14

Ein kleiner Betrieb stellt 5 verschiedene Artikel her. Die folgende Tabelle zeigt die Jahresumsätze der jeweiligen Artikel. Skizzieren Sie die Lorenzkurve, die die Konzentration der Umsätze veranschaulicht.

Minimum	Maximum
A	100
B	80
C	150
D	30
E	30

2. Kapitel 4-5

Aufgabe 1

Die folgenden Tabellen zeigen Alter und Gewicht von sechs beobachteten Personen. Das Alter ist für alle sechs Personen bekannt, das Gewicht nur für zwei der sechs Personen. Ergänzen Sie die vier unbekannten Gewichtswerte jeweils so, dass

- a) die Merkmale Alter und Gewicht negativ korreliert sind,
- b) die Merkmale Alter und Gewicht unabhängig sind,
- c) die Merkmale Alter und Gewicht einen Korrelationskoeffizienten von 1 aufweisen,
- d) die Merkmale Alter und Gewicht unkorreliert, aber nicht unabhängig sind.

Es ist nur die Ergänzung der Tabellen verlangt, keine Begründungen oder Berechnungen!

a)

Nummer der Person	1	2	3	4	5	6
Alter in Jahren	20	20	30	30	40	40
Gewicht in kg	60				70	

b)

Nummer der Person	1	2	3	4	5	6
Alter in Jahren	20	20	30	30	40	40
Gewicht in kg	60				70	

c)

Nummer der Person	1	2	3	4	5	6
Alter in Jahren	20	20	30	30	40	40
Gewicht in kg	60				70	

d)

Nummer der Person	1	2	3	4	5	6
Alter in Jahren	20	20	30	30	40	40
Gewicht in kg	60				70	

Aufgabe 2

Aus langjähriger Beobachtung sei bekannt, dass nur 10% aller Studierenden, die eine einführende Statistikvorlesung besuchen, regelmäßig während des Semesters den Vorlesungsstoff nachbereiten. Von diesen Fleißigen bestehen erfahrungsgemäß 5% die Klausur am Ende des Semesters nicht; außerdem ist bekannt, dass im Durchschnitt nur 2% der Studierenden, die die Klausur nicht bestehen, zu diesen Fleißigen gehören. Welcher Anteil der Studierenden, die nicht regelmäßig den Vorlesungsstoff nachbereiten, besteht die Klausur?