

Serie 01: Beschreibende Statistik

Aufgabe 1

Gegeben ist die folgende Auswertung einer Qualitätskontrolle zur Anzahl der Defekte auf Platinen in einer Leiterplattenproduktion.

a_i	0	1	2	3	4	5	6
h_i	83	25	28	18	12	10	2

- Zeichne die absolute bzw. relative Häufigkeit (PMF)
- Zeichne die absolute bzw. relative Verteilungsfunktion (CDF)

Aufgabe 2

Auf die Frage nach der Anzahl Geschwister geben 25 zufällig ausgewählte Studierende folgende Antworten

1 0 1 1 2 1 0 0 2 1 1 1 1 0 0 2 2 1 1 0 2 3 1 1 2 1

- Fülle folgende Tabelle aus:

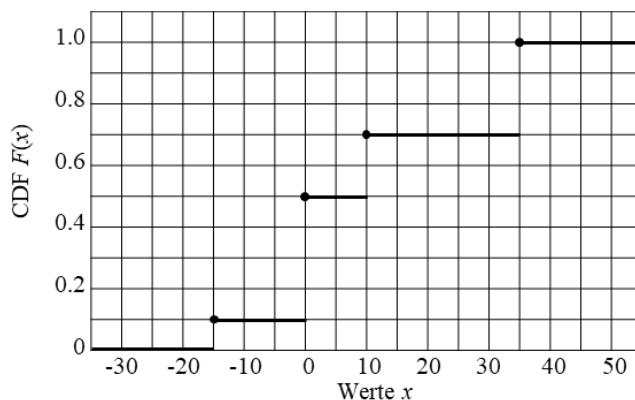
Anzahl Geschwister a_i	0	1	2	3
absolute Häufigkeit h_i	5	12	6	2
relative Häufigkeit f_i	5/25	12/25	6/25	2/25
kumulative absolute Häufigkeit H_i	5	17	23	25
kumulative relative Häufigkeit F_i	5/25	17/25	23/25	1

25

- Zeichne die relative Häufigkeitsfunktion (PMF) und die kumulative Verteilungsfunktion (CDF).

Aufgabe 3

Ein nicht in Klassen eingeteilter Datensatz hat folgende kumulative Verteilungsfunktion (CDF):



-15 0.1
 0 0.4
 10 0.7
 35 1.0

- Welche Werte kommen unter den Daten vor?
- Welche relativen Häufigkeiten haben die vorkommenden Datenwerte?
- Angenommen der Datensatz besteht aus insgesamt 40 Daten. Wie viele Daten haben dann genau den Wert 10? Wie viele höchstens den Wert 10? Und wie viele mindestens den Wert 10?

genau 10 = 8 höchstens zehn = 28
 min. zehn = 20

Klassierte Daten

Aufgabe 4

Auf die Frage «Wie viele Minuten brauchst du von zu Hause bis an die ZHAW?» geben 18 zufällig ausgewählte Studierende folgende Antworten:

~~30~~ 20 ~~10~~ ~~45~~ 15 ~~60~~ ~~20~~ 20 ~~20~~ ~~25~~ ~~10~~ 35 ~~20~~ 50 ~~10~~ 20 45 20

Wir teilen die Wegzeiten in folgende drei Klassen ein:

0 bis unter 15 Minuten, 15 bis unter 45 Minuten, 45 bis und mit 90 Minuten

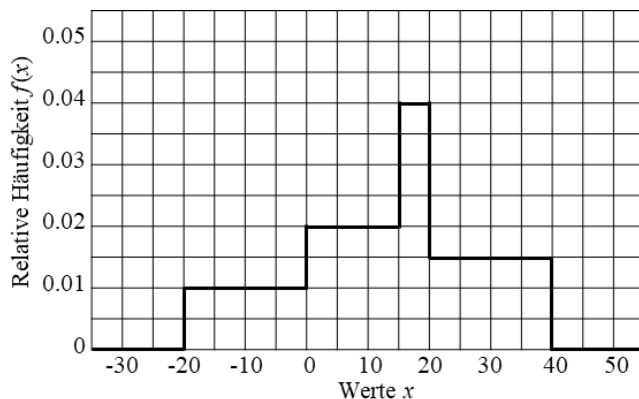
- a) Fülle nachfolgende Tabelle aus. Gib die nicht ganzzahligen Ergebnisse als Brüche und als auf 0.001 gerundete Dezimalzahlen an.

Wegzeiten in Minuten	[0, 15 [[15, 45 [[45, 90]
absolute Häufigkeit	3	9	6
relative Häufigkeit	$\frac{3}{18}$	$\frac{9}{18}$	$\frac{6}{18}$
relative Häufigkeitsdichte	$\frac{1}{90}$	$\frac{1}{60}$	$\frac{1}{135}$
kumulative relative Häufigkeit	$\frac{1}{6}$	$\frac{2}{3}$	1

- b) Zeichne die relative Häufigkeitsdichtefunktion (PDF) und die kumulative Verteilungsfunktion (CDF).

Aufgabe 5

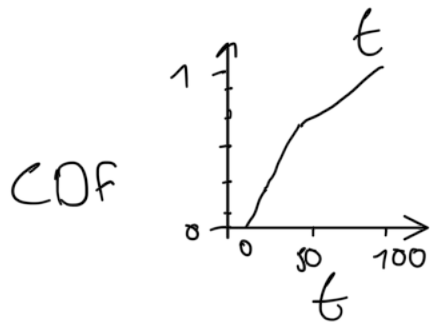
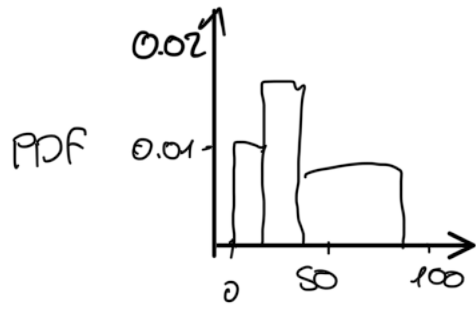
Ein in vier Klassen eingeteilter Datensatz besitzt folgende relative Häufigkeitsdichtefunktion (PDF) $f(x)$.



$[-20, 0[$, $[0, 15[$, $[15, 20[$, $[20, 40[$
 0.2 0.3 0.2 0.3

- a) Wie lauten die vier Klassen? Und welche relativen Häufigkeiten besitzen die Klassen?
 b) Zeichne die zugehörige kumulative Verteilungsfunktion (CDF) $F(x)$.
 c) Berechne mit linearer Interpolation $F(12)$ und $F(35)$.

4 b)



Aufgabe 6

Bei einer Firma werden in einem Monat 400 Lebensversicherungsverträge abgeschlossen. Nachstehend ist die klassifizierte Häufigkeitsverteilung für die Versicherungssummen angegeben.

Vers.sum. von .. bis unter ... (in kCHF)	[4,10[[10,20[[20,30[[30,40[[40,80[[80,120[
Anzahl Verträge	20	160	80	40	88	12

- Zeichne das Histogramm der relativen Häufigkeiten (PDF).
 - Zeichne die kumulative Verteilungsfunktion für klassierte Daten (CDF).
 - Simuliere die Stichprobe (rand... Befehl) und vergleiche die CDF der klassierten Daten mit der simulierten empirischen CDF der Stichprobe (optional).
-

Kenngrößen

Aufgabe 7

- Bestimme für Aufgabe 3 aus der gegebenen CDF graphisch das 0.1-Quantil und das 0.9-Quantil.
- Bestimme für Aufgabe 5 aus der selbst erstellten CDF graphisch das 0.4-Quantil und das 0.6-Quantil.

Aufgabe 8

Zwei Weitspringer führen Statistik über die im Training erzielten Leistungen:

	[700,720[[720,740[[740,760[[760,780[[780,800[[800,820[
A	19	24	26	27	10	5
B	4	8	52	40	32	24

Bei einem Wettkampf springt A 7.3m und B 7.5m weit.

- Welcher Weitspringer ist bezogen auf seine Trainingsleistung am weitesten ,unter Form'?
- Überlege, welches Mass geeignet ist, für diesen Vergleich.

Aufgabe 9

- Bestimme für Aufgabe 1 den Mittelwert, den Modus und den Median
- Bestimme für Aufgabe 6 durch Interpolation an der Summenkurve:
 - Wie viel % der Versicherten sind mit höchstens 18'000.- versichert?
 - Mit welchem Mindestbetrag sind die 20% der Personen versichert, die am höchsten versichert sind?
- Berechne für Aufgabe 6 näherungsweise den Median und den Mittelwert.

Aufgabe 10

Bei der Ermittlung der wöchentlichen Fahrzeiten verschiedener Personen im ÖV ergaben sich die folgenden Werte (in h):

2.1 2.4 2.8 3.1 4.2 4.9 5.1 6.0 6.4 7.3
10.8 12.5 13.0 13.7 14.8 17.6 19.6 23.0 25.0 35.2 39.6

- Bestimme den Median.
- Berechne den Mittelwert und die Standardabweichung, sowie die korrigierte Standardabweichung.

Aufgabe 11

Gemäss dem Statistischen Amt des Kantons Zürich versteuerten im Kanton Zürich im Jahr 2023 rund 693'000 natürliche Personen ein Einkommen zwischen CHF 20'000 und CHF 200'000, wobei die Einkommen wie folgt verteilt waren:

Steuerbares Einkommen (in kCHF)	[20,40[[40,60[[60,80[[80,100[[100,150[[150,200[
Anzahl Personen (in 1000)	148	186	141	87	98	33

- Berechne die kumulativen relativen Häufigkeiten. Runde die Ergebnisse auf 0.01.
- Berechne, welcher Anteil der Personen höchstens CHF 120'000 versteuern.
- Berechne, welches maximale Einkommen die 30% der Personen haben, welche am wenigsten versteuern.
- Welche Schiefe hat die Verteilung der steuerbaren Einkommen?
Gib ohne Rechnung an, welche Grössenbeziehung zwischen Median und Mittelwert der steuerbaren Einkommen besteht.
- Berechne näherungsweise den Median und den Mittelwert der steuerbaren Einkommen.

Boxplot

Aufgabe 12

Zeichne für das **Beispiel** Nettomieten im Skript einen Boxplot.

Aufgabe 13

Zu 36 zufällig ausgewählten Zeitpunkten wurde in einer petrochemischen Anlage die durchschnittliche Partikelkonzentration (in Mikrogramm pro Kubikmeter) der Luft gemessen.

Dabei ergaben sich folgende Messwerte:

5 18 15 7 23 220 130 85 103 25 80 7 24 6 13 65 37 25
24 65 82 95 77 15 70 110 44 28 33 81 29 14 45 92 17 53

Bestimme mit Python Histogramm, Verteilungsfunktion, Boxplot, Kennwerte (arith. Mittelwert, Q_1 , Q_2 , Q_3 , (korrigierte) Standardabweichung).