1 Aufgabe E1

Es soll eine neugezüchtete Kartoffelsorte getestet werden. Dazu wird ein Testfeld ausgewählt, das auf der gesamten Fläche gleiche Wachstumsvoraussetzungen (Bodenqualität, Sonneneinstrahlung etc.) bietet. Nachdem eine gleichmäßige Pflege der Pflanzen erfolgte (Bewässerung, Düngung etc.), wurde von der ersten Ernte auf dem Versuchsfeld eine Stichprobe entnommen, die die folgende Urliste (für das Gewicht der Kartoffeln in g) lieferte:

j	1	2	3	4	5	6	7	8	9	10
x_j	132	145	172	151	152	136	143	112	159	152

a) Bestimmen Sie die geordnete Stichprobe.

j	1	2	3	4	5	6	7	8	9	10
x_j	112	132	136	143	145	151	152	152	159	172

b) Klassieren Sie die obigen Daten (wählen Sie 4 gleichlange Intervalle) und zeichnen Sie das dazugehörige Histogramm.

Intervalle	Anzahl
$I_1 = [112, 127[$	1
$I_2 = [127, 142[$	2
$I_3 = [142, 157[$	5
$I_4 = [157, 172[$	2

c) Skizzieren Sie die empirische Verteilungsfunktion der unklassierten Daten (F(x) = Anteil der Kartoffeln mit einem Gewicht kleiner gleich x).

Lösung:
fehlt

d) Zeichnen Sie die empirische Verteilungsfunktion der klassierten Daten (wie bei unklassierten Daten außer, dass innerhalb einer Klasse der Anteil als gleichmäßig wachsend betrachtet wird).

Lösung:
fehlt

e) Berechnen Sie den Mittelwert (=arithmetisches Mittel) der Daten.

Lösung:
$\overline{x} = \frac{1}{10} \sum_{n=0}^{N} x_k$

f) Wie gross ist der Modalwert (der am häufigsten vorkommende Wert) der Stichprobe. Ist diese Information brauchbar?

Lösung:
Modalwert $= 152$

g) Bestimmen Sie den Median (der Wert, so dass mind. 50% der Daten kleiner gleich und mind. 50% der Daten größer gleich diesem sind). Ist dieser Wert eindeutig?

Lösung:
$x_{0,5} \in [x_{(5)}, x_{(6)}] = [145, 151]$
$x_{0,5} = \dfrac{x_{(5)} + x_{(6)}}{2}$

h) Bestimmen Sie das 0,25 Quantil (der Wert, so dass mind. 25% der Daten kleiner gleich und mind. 75% der Daten größer gleich diesem sind). Ist dieser Wert eindeutig?

Lösung:
$x_{0,25} = x_3 = 136$

i) Bestimmen Sie das 0,75 Quantil (der Wert, so dass mind. 75% der Daten kleiner gleich und mind. 25% der Daten größer gleich diesem sind).

Lösung:
$x_{0,75} = x_{(8)} = 152$

j) Bestimmen Sie den Quartilsabstand (die Differenz zwischen dem 0,75 und dem 0,25 Quantil)

Lösung:
Quartilsabstand $x_{0,75} - x_{0,25} = 152 - 136 = 16$

k) Bestimmen Sie das 0,3 Quantil (der Wert, so dass mind. 30% der Daten kleiner gleich und mind. 70% der Daten größer gleich diesem sind). Ist dieser Wert eindeutig?

Lösung:
$x_{0,3} \in [x_{(3)}, x_{(5)}]$

l) Bestimmen Sie die empirische Varianz und daraus die empirische Standardabweichung.

Lösung:
empirische Varianz $s = \dfrac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})$
empireische Standartabweichung $\sqrt{s^2} = \sqrt{\dfrac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x}}$

Index	Gewicht	$x(i) - \overline{x}$
1	112	$[112 - 145,4]^2$
2	132	$[132 - 145,4]^2$
3	136	$[136 - 145,4]^2$
4	143	$[143 - 145,4]^2$
5	145	$[145 - 145,4]^2$
6	151	$[151 - 145,4]^2$
7	152	$[152 - 145,4]^2$
8	152	$[152 - 145,4]^2$
9	159	$[159 - 145,4]^2$
10	172	$[172 - 145,4]^2$
Summe	1454	-
Mittelwert	145,4	-

Index	Gewicht	$x(i) - \overline{x}$
1	112	$1115,56$
2	132	$179,56$
3	136	$88,36$
4	143	$5,76$
5	145	$0,16$
6	151	$31,36$
7	152	$43,56$
8	152	$43,56$
9	159	$182,96$
10	172	$707,56$
-	Summe	$2400,4$
-	s^2	$240,04$
-	s	$15,49$

m) Zeichnen Sie den zu dieser Verteilung (unklassierte Daten) passende Boxplot.

Lösung:
Minimum $x_{min} = 112$
Maximum $x_{max} = 172$
unteres Quantil $x_{0,25} = 136$
unteres Quantil $x_{0,75} = 152$
Median $x_{0,5} = 148$
arithmetisches Mittel $\overline{x} = 1145,4$

Grafik fehlt

n) Wenn man den Wert 136 durch 316 ersetzt (Schreibfehler), wie wirkt das dann auf den Mittelwert, die Standardabweichung, den Median und die Quantile aus (nur qualitativ: also wächst/fällt leicht/stark)?

Lösung:

$136 \to 316$ $x_{(0,5)}$ wächst leicht
Quantile wachsen langsam oder nicht
Quartilsabstand wächst leicht
$\overline{x}$ wächst stark
s sehr stark

2 Aufgabe E2

Ordnen Sie die Werte den passenden Histogrammen zu:

Nr.	Mittelwert	Median	Standartabweichung
1	10	6	7,9
2	10	8	3,7
3	10	10	2,2
4	10	10	6,1
5	10	10	9,2
6	10	11	6,2

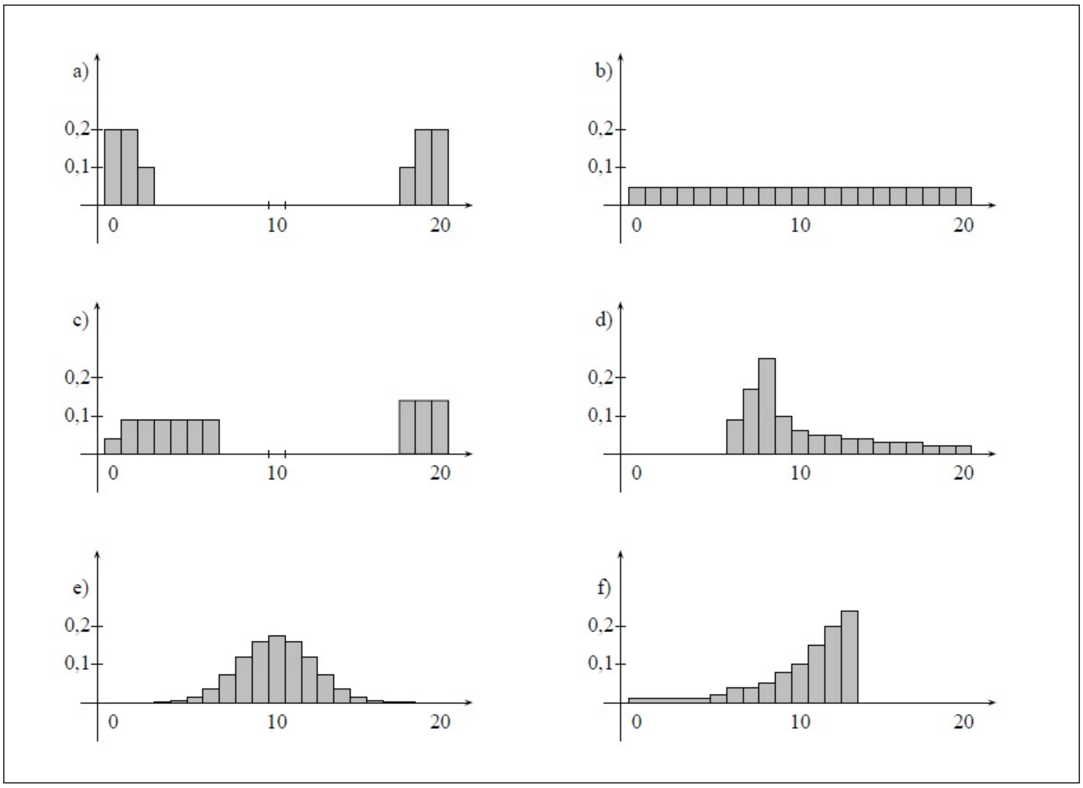

Abbildung 1: a)=5, b)=4, c)=1, d)=2, e)=3, f)=6

Lösung:

Nr.	Mittelwert	Median	Standartabweichung
a)	10	10	9,2
b)	10	10	6,1
c)	10	6	7,9
d)	10	8	3,7
e)	10	10	2,2
f)	10	11	3,2

3 Aufgabe E3

a) Fünf medizinische Untersuchungsinstitute gaben als erforderlichen Zeitaufwand für die Bestimmung der Blutbestandteile die folgenden Zeitdauern je Blutprobe in Minuten an:

Institut	1	2	3	4	5
Dauer	8	1	2	3	6

1) Berechnen Sie den mittleren Zeitaufwand, falls jedes Institut gleich viele Untersuchungen vornimmt.
2) Berechnen Sie den mittleren Zeitaufwand für die Bestimmung der Blutbestandteile, falls alle Institute über denselben Zeitraum von 8 Stunden ununterbrochen Blutproben untersuchen.

b) Man beobachtet das Wachstum von Schnittlauch, der zu Beginn des Beobachtungszeitraumes 5 cm hoch ist. Nach einer Woche ist er um 150% gewachsen. Während der zweiten Woche wächst er um 20% und während der dritten Woche um 45%.Berechnen Sie das durchschnittliche Wachstum der Pflanze pro Woche.

4 Aufgabe E4

Über eine Person sei bekannt, dass sie viel liest und zwar:

Jahr	Anz. Bücher	Median d. Seit.	M-wert d. Sz.	0,95-Quantil d. Sz.	Standartabw. d. Sz.
2008	90	350	400	710	90
2009	55	450	500	800	100

Bestimmen Sie so genau wie möglich: den Mittelwert, den Median, das 0,95-Quantil und die Standardabweichung des gesamten Zeitraumes 2008-2009.

Lösung:

$\overline{x}$, x_j, ..., für 2008
$\overline{y}$, y_j, ..., für 2009
$\overline{z}$, z_j, ..., für Gesamtzeitraum

$\overline{z} = \frac{(nx \cdot \overline{x} + n_y \cdot \overline{y})}{n_x + n_y} = \frac{1}{145} \cdot (90 \cdot 400 + 55 \cdot 500) = 438$

$\overline{z}_{0,5} \in [min(x_{0,5}, y_{0,5)}, max(x_{0,5}, y_{0,5)}] = [350, 450]$
$\overline{z}_{0,95} \in [min(x_{0,95}, y_{0,95}), max(x_{0,95}, y_{0,95}] = [710, 800]$
$s_z^2 = \frac{1}{n_x+n_y} \cdot (n_x \cdot sx^2 + n_y \cdot s_y^2) + \frac{1}{n_x+n_y} \cdot (n_x(\overline{x} - \overline{z})^2) + ny \cdot (\overline{y} - \overline{z})$
$s_z = 106 \geq min(s_x, x_y)$
$\geq$ bedeutet: immer Bestimmen Sie so genau wie möglich: den Mittelwert, den Median, das 0,95-Quantil und die Standardabweichung des gesamten Zeitraumes 2008-2009.

5 Aufgabe E5

Gegeben sei die Kontingenztafel für die Merkmale X Baumart (Ulme, Kiefer, Fichte) und Y Schädlingsbefall durch Borkenkäfer (kein, gering, mittel, groß)

Tabelle Aufgabe E5

a) Berechnen Sie die Randverteilungen in Wahrscheinlichkeiten

b) Man bestimme die bedingte Verteilung von Y bedingt auf X =Ulme.

c) Sind die Merkmale unabhängig?

6 Aufgabe E6

Drei Stichproben mit den Umfängen 10, 30 und 60 werden zu einer Gesamtstichprobe vom Umfang 100 zusammengefasst. Die Mittelwerte dieser Stichproben seien 64, 50 und 58. Konstruieren Sie ein möglichst einfaches Beispiel, für das der empirische Median der Gesamtstichprobe in obiger Situation gleich 0 ist.

7 Aufgabe E7

Um den absoluten Nullpunkt der Temperatur zu bestimmen, kann man in einem eingeschlossenen Gas bei konstantem Volumen die Temperatur und den Druck messen. Die nachfolgende Tabelle gibt die Messwerte an. Es gilt: Wenn der Druck auf Null gesenkt wird, ist der absolute Nullpunkt der Temperatur erreicht.

a) Bestimmen Sie den Korrelationskoeffizient.

b) Bestimmen Sie eine zur Frage passende Regressionsgerade und bestimmen Sie damit einen Näherungswert für die Temperatur am absoluten Nullpunktes in °C.

Tabelle Aufgabe E7

Hilfe: Sie dürfen verwenden, dass:

$$\frac{1}{11}\sum_{i=1}^{11} t_i = 59,8; \frac{1}{11}\sum_{i=1}^{11} t_i = 4156,7; \frac{1}{11}\sum_{i=1}^{11} t_i = 1100,8; \frac{1}{11}\sum_{i=1}^{11} t_i = 1218115; \frac{1}{11}\sum_{i=1}^{11} t_i = 67742;$$

8 Aufgabe E8

In einer Studie soll der Zusammenhang zwischen den Merkmalen Farbenblindheit und Geschlecht untersucht werden. Der Anteil der M¨ anner an der Bevölkerung beträgt ungefähr 48.6Farbenblinden unter den Männern 7

a) Wie groß ist die Wahrscheinlichkeit, dass eine rein zufällig ausgewählte Person farbenblind ist? (Kurz: Wieviel Prozent der Bevölkerung sind farbenblind?)

b) Sind die Ereignisse Person ist weiblichen Geschlechts und Person ist farbenblind unabhängig?

c) Wie groß ist die Wahrscheinlichkeit, dass eine zufällig ausgewählte Person männlichen Geschlechts ist, unter der Bedingung, dass die Person farbenblind ist? (Kurz: Wie groß ist der Anteil an Männern unter den Farbenblinden?)

9 Aufgabe E9

Der Tutor kennt die Arbeitsgruppe der Studenten A,B,C schon seit längerem und weiß, dass Student A 80%, Student B 15% und Student C lediglich 5% der Aufgaben bearbeitet und die Studenten so organisiert sind, dass keine Aufgabe doppelt bearbeitet wird. Aufgrund ihrer unterschiedlicher Erfahrung lösen sie die Aufgaben mit einer Wahrscheinlichkeit von 90%, 50% bzw. lediglich 10% richtig. Der Tutor hat von der Arbeitsgruppe eine fehlerhafte Lösung bekommen. Mit welcher Wahrscheinlichkeit stammt die Lösung von welchem der drei Studenten?

10 Aufgabe E10

In einer Bevölkerungsgruppe werden Personen zufällig ausgewählt und ihre Cholesterinmenge im Blut ermittelt. Es sei X das Alter der untersuchten Person, Y die Cholesterinmenge (in mg/l) im Blut und Z gebe an, ob die Person Diabetiker ist ($Z = 1$, falls Diabetiker; $Z = 0$, sonst). Es sind folgende Wahrscheinlichkeiten bekannt:

$$\begin{aligned}
&P(\{50 < X \leq 60\}) = 0.13,\\
&P(\{X \leq 60\} \cap \{Z = 1\}) = 0.05,\\
&P(\{X \leq 60\} \cap \{Z = 0\}) = 0.75,\\
&P(\{X \leq 60\} \cap \{Y \leq 150\} \cap \{Z = 1\}) = 0.01,\\
&P(\{50 < X \leq 60\} \cap \{Z = 0\}) = 0.01.
\end{aligned}$$

Berechnen Sie die Wahrscheinlichkeiten der folgenden Ereignisse:

a) $\{X > 60\}$

b) $\{X \leq 50\}$

c) $\{X \leq 50\} \cap \{Z = 0\}$

d) $\{X \leq 60\} \cap \{Y > 150\} \cap \{Z = 1\}$

e) Eine zufällig ausgewählte Person ist

   i) älter als 60 Jahre oder Diabetiker

   ii) höchstens 50 Jahre alt, oder höchstens 60 Jahre alt und nicht Diabetiker

11 Aufgabe E11

In einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ seien zwei Ereignisse A und B mit $A \cup B = \Omega$ gegeben. Zeigen Sie:

$$P(A \cap B) = P(A)P(B) - P(A^C)P(B^C)$$

Hinweis: Formen Sie die beiden Seiten so um, dass nur noch $P(A), P(B)$ vorkommen