

Einführung, grafische Darstellung univariater Datensätze

Aufgabe 1

Vor einer Bürgermeisterwahl, bei der fünf Kandidaten (A bis E) zur Auswahl stehen, wurden 160 Wahlberechtigte nach ihrer Wahlabsicht befragt. Die Befragung lieferte folgende Tabelle:

Kandidat	A	B	C	D	E
Stimmen	12	40	60	20	28

- Bestimmen Sie die relativen Häufigkeiten und zeichnen Sie das dazugehörige Stabdiagramm.
- Zeichnen Sie das dazugehörige Kreisdiagramm.

Aufgabe 2

Bei einer Untersuchung wurde die Körpergröße in *cm* von 20 Personen bestimmt. Dabei ergaben sich folgende Werte:

172 164 160 162 173 180 158 185 158 192
 171 181 162 184 177 175 177 174 151 177

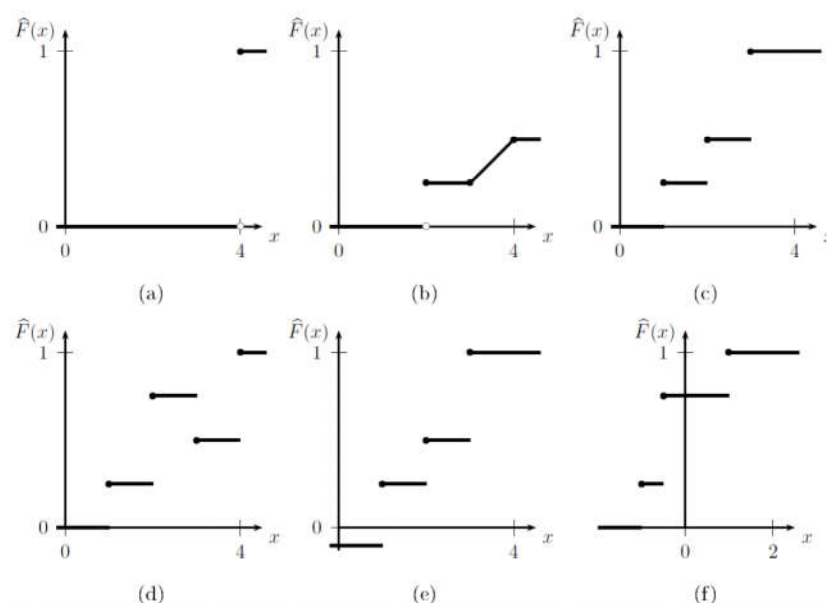
- Stellen Sie die relativen Häufigkeiten in einem Histogramm dar. Teilen Sie dafür die Stichprobe in die fünf Klassen $[150; 160)$; $[160; 170)$; $[170; 180)$; $[180; 190)$; $[190; 200)$ ein.
- Stellen Sie die relativen Häufigkeiten in einem Histogramm dar. Teilen Sie dafür die Stichprobe in die vier Klassen $[140; 160)$; $[160; 165)$; $[165; 190)$; $[190; 200)$ ein.

Aufgabe 3

Eine Stichprobe vom Umfang $n = 8$ lieferte folgende Ergebnisse:

26 23 16 32 23 18 27 18

- Zeichnen Sie die empirische Verteilungsfunktion.
- Bei welchen der folgenden Abbildungen handelt es sich nicht um eine empirische Verteilungsfunktion?



Beschreibung univariater Datensätze I

Aufgabe 1

Gegeben seien die Merkmalsausprägungen x_i mit:

i	1	2	3	4	5	6	7	8	9
x_i	3,75	4,43	5,5	3,5	3	3	6,5	6,25	2,5

- Bestimmen Sie den Modus, den Median und das arithmetische Mittel.
- Bestimmen Sie die Varianz und die Standardabweichung.
- Geben Sie an, ob die empirische Verteilung linkssteil, rechtssteil oder symmetrisch ist. Begründen Sie kurz Ihre Antwort.
- Wie ändern sich der Modus, der Median, das arithmetische Mittel und die Varianz, wenn die x_i alle um 20% steigen?
- Wie ändern sich der Modus, der Median und das arithmetische Mittel, wenn sich x_7 verdoppelt? Eine Tendenz genügt als Antwort.
- Gegeben sei eine zweite Stichprobe vom Umfang $n = 10$ mit $\bar{y} = 50$ und $\sum (y_i - \bar{y})^2 = 250$. Welche Stichprobe weist die größere Streuung auf?

Aufgabe 2

Gegeben seien die Merkmalsausprägungen x_i mit:

i	1	2	3	4	5	6	7	8	9	10
x_i	5	2	4	9	8	4	2	6	4	8

Zusätzlich seien bekannt:

$$\widehat{M}_3 = \frac{1}{10} \sum_{i=1}^{10} (x_i - \bar{x})^3 = 2,856 \quad \widehat{M}_4 = \frac{1}{10} \sum_{i=1}^{10} (x_i - \bar{x})^4 = 54,7792$$

- Bestimmen Sie den Modus und den Median.
- Bestimmen Sie die rechnerisch die Schiefe und die Wölbung.

Aufgabe 3

Gegeben seien nachfolgende Beobachtungswerte:

5,5	6	8	8,5	10,6	12,5	13,2
13,7	15	15,2	16,6	16,8	18,9	19,4

- Bestimmen Sie das arithmetische Mittel $\bar{x}$ aus den Beobachtungswerten der Urliste.
- Bilden Sie die Klassen über 5 bis 10, über 10 bis 15, über 15 bis 20 und berechnen Sie die Durchschnitte $\bar{x}_k$ pro Klasse und daraus erneut das arithmetische Mittel $\bar{x}$.
- Bestimmen Sie mit Hilfe der Klassenmitten eine Näherung $\bar{x}'$ für das arithmetische Mittel und vergleichen Sie die bisher ermittelten Ergebnisse.

Aufgabe 4

Der Absatz eines Gutes ist in nachfolgender Tabelle gegeben. Berechnen Sie die mittlere jährliche Absatzsteigerung.

Jahr i	1	2	3	4	5	6	7	8	9	10	11
Absatz x_i	9	20	40	75	90	140	244	410	705	1036	1718

Beschreibung univariater Datensätze II

Aufgabe 1

Gegeben seien nachfolgende 17 Werte:

0 0 1 2 2 3 3 3 3 4 5 6 6 7 8 9 15

- Zeichnen Sie den dazugehörigen Box-Plot.
- Treffen Sie eine Aussage zur Schiefe. Begründen Sie Ihre Aussage.

Aufgabe 2

Es wurden sechs Personen zu ihrem Jahreseinkommen in 1000 € befragt. Dabei ergab sich folgende Tabelle:

Person	1	2	3	4	5	6
Einkommen	5	100	15	25	40	15

- Zeichnen Sie die dazugehörige Konzentrationskurve.
- Das Gesamteinkommen beträgt 200.000 €. Eine der Personen verdient alleine 100.000 €, also genau die Hälfte. Kann man demnach von einer hohen Konzentration sprechen? Begründen Sie Ihre Antwort.

Aufgabe 3

Ein Markt mit zehn Unternehmen setzt insgesamt 15 Mio. € um und kann nach dem Umsatz in drei Klassen (klein, mittel, groß) unterteilt werden. Auf dem Markt agieren fünf kleine, vier mittlere Unternehmen und ein Marktführer. Die mittleren Unternehmen setzen dabei je 1,5 Mio. € um. Der Marktführer erzielt einen Umsatz von 6 Mio. €.

- Zeichnen Sie die Lorenzkurve und bestimmen Sie den Gini-Koeffizienten. Beurteilen Sie die Konzentration.
- Auf einem zweiten Markt agieren vier Unternehmen. Dabei beträgt $G_2 = 0,39$. Welcher der beiden Märkte weist eine höhere Konzentration auf?

Beschreibung bivariater Datensätze I: Abhängigkeitsmaße

Aufgabe 1

Bei einer Verkehrskontrolle wurden 10 Personen erfasst. Für die beiden Merkmale Alter und Alkoholgehalt ergab sich folgende Tabelle:

Alter	20	21	20	22	20	22	21	20	21	21
Promille	0,2	0,1	0	0,2	0,1	0,1	0,2	0	0,1	0

- Stellen Sie in einer Kontingenztafel die absoluten und relativen Häufigkeiten dar.
- Bestimmen Sie den χ^2 -Koeffizienten und den normierten Kontingenzkoeffizienten und interpretieren Sie Ihr Ergebnis.
- Wie ändern sich die beiden Koeffizienten, wenn sich der Stichprobenumfang verdoppelt, aber die relativen Häufigkeiten gleich bleiben? Ändert eine Umordnung der Spalten und/oder Zeilen das Ergebnis?

Aufgabe 2

Gegeben sei nachfolgende Tabelle:

x_i	16	19	17	15	20	19	20
y_i	100	150	120	90	150	140	160

- Zeichnen Sie das zugehörige Streudiagramm.
- Bestimmen Sie den Korrelationskoeffizienten von Bravais-Pearson. Interpretieren Sie Ihr Ergebnis.

Aufgabe 3

Gegeben sei nachfolgende Tabelle:

x_i	-2	-1	0	1	2
y_i	4	1	0	1	4

- Zeichnen Sie das zugehörige Streudiagramm.
- Bestimmen Sie den Korrelationskoeffizient von Bravais-Pearson. Nehmen Sie zur Aussage, es bestehe kein Zusammenhang, Stellung.

Aufgabe 4

Es wurden sechs Abiturienten befragt, welche Note sie bei ihrer Prüfung erzielt haben ((g)ut, (b)efriedigend, (a)usreichend, (m)angelhaft). Zusätzlich wurden sie befragt, ob sie eine vorherige Probeklausur bestanden haben ((j)a, (n)ein). Dabei ergaben sich folgende Antworten:

(b, n) (b, j) (b, j) (a, n) (m, n) (g, j)

Kann ein Zusammenhang zwischen der erzielten Note und dem Abschneiden bei der Probeklausur festgestellt werden?

Beschreibung bivariater Datensätze II: Das lineare Regressionsmodell

Aufgabe 1

In einem Schwellenland wurde eine Studie zum Zusammenhang zwischen dem Einkommen der Eltern und dem Geburtsgewicht des Kindes durchgeführt. Dabei wurden das monatliche Einkommen x_i in 1000 GE und das Geburtsgewicht y_i in Pfund betrachtet:

x_i	2,7	1,9	3,1	3,9	4	3,4	2,1	2,9
y_i	5	6	9	8	7	6	7	8

- Stellen Sie eine sinnvolle Regressionsbeziehung auf.
- Zeichnen Sie das zugehörige Streudiagramm.
- Bestimmen Sie die Regressionsgerade.
- Das Einkommen einer Familie beträgt 3000 GE. Welches Gewicht wird dann prognostiziert?
- Ist die gewählte Regression geeignet? Nutzen Sie das Bestimmtheitsmaß für Ihre Stellungnahme.

Aufgabe 2

Gegeben sei folgende Regressionsgerade:

$$\widehat{schlaf} = 3586,4 - 0,151 \cdot arbeit$$

Dabei ist *schlaf* die Zeit in Minuten, die pro Woche zum Schlafen genutzt wird. *arbeit* gibt die Minuten, die pro Woche gearbeitet werden, an.

- Interpretieren Sie den Achsenabschnitt.
- Angenommen *arbeit* steigt um 2 Stunden. Wie wirkt sich diese Erhöhung auf *schlaf* aus? Handelt es sich um einen großen Effekt?

Aufgabe 3

Gegeben sei folgende Tabelle:

P	11	9	12	13	14	16	17	30	28	28	42	55
Q	67	55	56	58	51	57	46	38	39	69	36	35

Dabei ist P der beobachtete Preis und Q die nachgefragte Menge eines Gutes in den USA zwischen den Jahren 1958 und 1969.

- Bestimmen Sie die zugehörige Regressionsgerade und unterstellen Sie, dass P der Regressor ist. Interpretieren Sie außerdem die Parameter der Regressionsgerade.
- Führen Sie eine log-log-Regression durch. Interpretieren Sie außerdem den Steigungsparameter der Regressionsgerade.

Indizes und Zeitreihen

Aufgabe 1

Für 5 Güter sind die Preise $p_{I,i}$ und $p_{II,i}$ ($i = 1, \dots, 5$) und die Umsätze $u_{I,i}$ und $u_{II,i}$ für die Perioden I und II in nachfolgender Tabelle gegeben:

i	$p_{I,i}$	$p_{II,i}$	$u_{I,i}$	$u_{II,i}$
1	5	6	50	48
2	4	6	200	240
3	10	9	80	144
4	8	10	48	40
5	3	3	30	24

- Bestimmen Sie die Preisindizes nach Laspeyres und nach Paasche mit I als Basis- und II als Berichtsperiode.
- Bestimmen Sie die Preisindizes nach Laspeyres und nach Paasche mit II als Basis- und I als Berichtsperiode.

Aufgabe 2

Zeigen Sie, dass der Fisher-Preisindex folgende Fisher-Forderung erfüllt:

- Zeitumkehrbarkeit.
- Faktorumkehrbarkeit

Aufgabe 3

Bestimmen Sie den Fisher-Preisindex für die in Aufgabe 1 gegebenen Werte mit:

- Basisperiode I und Berichtsperiode II.
- Basisperiode II und Berichtsperiode I.

Aufgabe 4

Ein Unternehmen hat in den letzten 12 Monaten folgende Umsätze in 1000 Euro erzielt:

Monat t	1	2	3	4	5	6	7	8	9	10	11	12
Umsatz y_t	80	90	87	100	76	68	92	110	95	72	102	90

- Bestimmen Sie für die gegebene Zeitreihe die geglätteten Werte mittels der Methode des exponentiellen Glättens. Verwenden Sie hierfür $\alpha = 0,6$.
- Stellen Sie graphisch den Verlauf der Zeitreihe und der geglätteten Zeitreihe dar.
- Wie verändert sich die Graphik unter Verwendung von $\alpha = 0$ bzw. $\alpha = 1$?

Einführung in die Wahrscheinlichkeitstheorie I

Aufgabe 1

Für die Ereignisse A , B und C aus einem geeigneten Wahrscheinlichkeitsraum gilt:

$$\begin{array}{lll} P(A) = 0,5 & P(B) = 0,2 & P(C) = 0,3 \\ P(A \cup B) = 0,7 & P(A \cup C) = 0,6 & P(B \cap C) = 0,1 \end{array}$$

Berechnen Sie nachfolgende Wahrscheinlichkeiten:

- a) $P(\overline{A})$ und $P(A \cap B)$
- b) $P(\overline{A} \cap \overline{C})$ und $P(\overline{A} \cup \overline{C})$
- c) $P(A \cap B \cap C)$
- d) $P(\overline{B} \cap C)$ und $P(\overline{C} \cap B)$

Aufgabe 2

In einem Semester werden die Vorlesungen Mathematik, Statistik und Volkswirtschaftslehre angeboten. Von den Studenten besuchen 59% Statistik, 65,8% besuchen Mathe und 37,2% besuchen VWL. Außerdem ist bekannt, dass 43,4% sowohl Mathe als auch Statistik besuchen, 16,6% Mathe und VWL und 12,6% besuchen Statistik und VWL.

- a) Skizzieren Sie ein geeignetes Venn-Diagramm.
- b) Wie groß ist der Anteil der Studierenden, der alle drei Vorlesungen besucht?
- c) Wie groß ist der Anteil der Studierenden, der zwar Mathe, aber nicht Statistik besucht?
- d) Wie groß ist der Anteil der Studierenden, der Mathe oder Statistik besucht, aber nicht VWL?
- e) Wie groß ist der Anteil der Studierenden, der nur Statistik besucht?

Aufgabe 3

Entscheiden Sie für die nachfolgenden Funktionen, ob diese die Eigenschaften einer Wahrscheinlichkeits-/Dichtefunktion bzw. einer Verteilungsfunktion besitzen.

$$g_1(x) = \begin{cases} \frac{x}{2} & \text{für } 0 \leq x \leq 2 \\ 0 & \text{sonst} \end{cases}$$

$$g_2(x) = e^x \quad \text{für } x \in \mathbb{R}$$

$$g_3(x) = \begin{cases} 0 & \text{für } x < 0 \\ \frac{1}{9}x^2 & \text{für } 0 \leq x \leq 3 \\ 1 & \text{für } x > 3 \end{cases}$$

$$g_4(x) = \begin{cases} \frac{x}{28} & \text{für } x \in \{1, 2, 3, 4, 5, 6, 7\} \\ 0 & \text{sonst} \end{cases}$$

$$g_5(x) = \begin{cases} 0 & \text{für } x < 0 \\ 1 & \text{für } x \geq 0 \end{cases}$$

$$g_6(x) = \begin{cases} 0 & \text{für } x \leq 0 \\ 1 & \text{für } x > 0 \end{cases}$$

Einführung in die Wahrscheinlichkeitstheorie II

Aufgabe 1

Gegeben sei folgende Wahrscheinlichkeitstabelle:

$Y \backslash X$	-1	3	5
1	0,06	0,3	0,24
2	0,04	0,2	0,16

- Bestimmen Sie die Wahrscheinlichkeit für $Y = 2$ unter der Bedingung, dass X positiv ist.
- Bestimmen Sie den Erwartungswert und die Varianz von X .

Aufgabe 2

Gegeben sei eine Zufallsvariable X mit nachfolgender Dichtefunktion:

$$f_X(x) := \begin{cases} 0,1x & \text{für } 0 \leq x \leq 4 \\ 2 - 0,4x & \text{für } 4 < x \leq 5 \\ 0 & \text{sonst} \end{cases}$$

- Bestimmen Sie die zugehörige Verteilungsfunktion.
- Bestimmen Sie den Erwartungswert und die Varianz.

Aufgabe 3

An einer Studie zum Auftreten von Farbenblindheit nimmt eine Gruppe von Personen teil, die sich zu 45% aus Männern (A) und zu 55% aus Frauen ($\bar{A}$) zusammensetzt. Man weiß, dass im Allgemeinen 6% der Männer farbenblind (B) sind. Dagegen sind nur 0,5% der Frauen farbenblind.

- Formulieren Sie die obigen Angaben als Wahrscheinlichkeiten.
- Bestimmen Sie die Wahrscheinlichkeit, eine Person zu wählen, die eine farbenblinde Frau ist.
- Beschreiben Sie die folgenden Ereignisse in Worten und bestimmen Sie deren Wahrscheinlichkeit.
 - $P(A \cap B)$
 - $P(\bar{A} \cap \bar{B})$
 - $P(B)$
 - $P(\bar{A}|B)$

Aufgabe 4

In Hamburg gibt es im Mittel 10% Schwarzfahrer. 70% von ihnen haben keine Fahrkarte. Die übrigen sind mit einer Fälschung unterwegs. Von den ehrlichen Fahrgästen haben im Mittel 5% ihre Karte vergessen. Bestimmen Sie die Wahrscheinlichkeit, dass ein Fahrgast ein Schwarzfahrer ist, wenn der Fahrgast keine Karte hat.

Spezielle diskrete Verteilungen

Aufgabe 1

In einer Urne befinden sich fünf weiße und drei schwarze Kugeln. Es werden fünf Kugeln mit Zurücklegen gezogen.

- Durch welche Ihnen bekannte Verteilung könnte das Zufallsexperiment beschrieben werden? Welche Parameter besitzt diese Verteilung? Bestimmen Sie deren Werte für das angegebene Zufallsexperiment.
- Bestimmen Sie den Erwartungswert und die Varianz.
- Wie groß sind die Wahrscheinlichkeiten 0, 1, 2, 3, 4 oder 5 schwarze Kugeln zu ziehen?
- Wie groß ist die Wahrscheinlichkeit, mindestens drei weiße Kugeln zu ziehen?
- Wie groß ist die Wahrscheinlichkeit, mindestens eine, aber höchstens zwei schwarze Kugeln zu ziehen?

Aufgabe 2

Das Präsidium eines Vereins besteht aus 5 männlichen und 3 weiblichen Mitgliedern. Aus diesem Kreis sollen durch eine Zufallsauswahl der 1. und 2. Vorsitzende und der Schatzmeister ausgewählt werden, wobei niemand mehr als einen Posten bekleiden darf.

- Wieviele Posten in dem Präsidium werden im Mittel von Frauen bekleidet?
- Wie hoch ist die Wahrscheinlichkeit, dass für alle Ämter Frauen ausgewählt werden?
- Wie hoch ist die Wahrscheinlichkeit, dass für alle Ämter Männer ausgewählt werden?
- Man berechne die Wahrscheinlichkeit, dass in der Dreiergruppe die Frauen in der Überzahl sind.

Aufgabe 3

In einer sehr großen technischen Anlage fallen im Mittel drei Module pro Tag aus.

- Welche Ihnen bekannte Verteilung könnte dem zu Grunde liegen? Bestimmen Sie alle Parameter, den Erwartungswert und die Varianz.
- Wie hoch ist die Wahrscheinlichkeit, dass an einem Tag 5 Module ausfallen?
- Wie hoch ist die Wahrscheinlichkeit, dass an einem Tag mehr Module ausfallen als im täglichen Mittel?

Aufgabe 4

Ein Glücksrad hat 25 Felder, von denen vier Felder Gewinnfelder sind. Ein Spieler darf 5-mal drehen.

- Wie hoch ist die Wahrscheinlichkeit, spätestens nach dem dritten Versuch zu gewinnen?
- Wie hoch ist die Wahrscheinlichkeit, nach fünf Versuchen noch nicht gewonnen zu haben?
- Bestimmen Sie den Erwartungswert und die Varianz für die Anzahl der Versuche, um zu gewinnen.

Spezielle stetige Verteilungen

Aufgabe 1

An einem Bahnhof fährt ein bestimmter Bus exakt alle 20 Minuten ab. Unterstellen Sie dabei, dass die Wartezeit X stetig gleichverteilt ist.

- Wie groß ist die Wahrscheinlichkeit, dass ein zufällig eintreffender Fahrgast mehr als 15 Minuten auf diesen Bus warten muss?
- Bestimmen Sie die Wartezeit, die im Mittel zu erwarten ist.

Aufgabe 2

Die Lebensdauer einer Maschine beträgt im Mittel 25000 Stunden.

- Welche Ihnen bekannte Verteilung könnte hier zu Grunde liegen? Bestimmen Sie deren Parameter.
- Wie hoch ist die Wahrscheinlichkeit, dass die Maschine
 - höchstens 30000 Stunden hält.
 - mindestens 20000 Stunden hält.
 - zwischen 20000 und 30000 Stunden hält.
- Wie hoch ist die Wahrscheinlichkeit, dass die Maschine ihre mittlere Lebensdauer um mehr als zwei Standardabweichungen übersteigt.

Aufgabe 3

Es sei $X \sim \mathcal{N}(35, 144)$. Bestimmen Sie mittels geeigneter Tabelle nachfolgende Wahrscheinlichkeiten.

- $P(X \leq 41)$, $P(X \leq 50)$, $P(X \leq 23)$
- $P(X > 41)$, $P(X > 20)$
- $P(23 < X \leq 41)$, $P(38 < X \leq 50)$, $P(20 < X \leq 23)$

Quantile, Ungleichungen, momenterzeugende Funktionen

Aufgabe 1

Gegeben sei nachfolgende Verteilungsfunktion:

$$F_X(x) = \begin{cases} 0 & \text{für } x < 0 \\ 0,05x^2 & \text{für } 0 \leq x \leq 4 \\ -0,2x^2 + 2x - 4 & \text{für } 4 < x \leq 5 \\ 1 & \text{für } x > 5 \end{cases}$$

Bestimmen Sie das 10% und das 95%-Quantil.

Aufgabe 2

Die Zufallsvariable X ist normalverteilt mit den Parametern $\mu = 8$ und $\sigma^2 = 4$. Bestimmen Sie das Intervall mit unterer Grenze 8, dem eine Wahrscheinlichkeit von 0,475 zukommt.

Aufgabe 3

Von einer nichtnegativen Zufallsvariablen X sei bekannt, dass $E(X) = 10$ und $Var(X) = 2$ gilt. Schätzen Sie nachfolgende Wahrscheinlichkeiten ab: $P(8 < X < 12)$, $P(5 < X < 15)$, $P(X \geq 13 \vee X \leq 7)$, $P(X \geq 18 \vee X \leq 2)$

Aufgabe 4

Gegeben sei eine Zufallsvariable X mit nachfolgender momenterzeugenden Funktion ($t < 1$):

$$M_X(t) := \frac{1}{(t-1)^2}$$

Bestimmen Sie $E(X)$ und $Var(X)$.