Prof. Dr. U. Kamps
C. Empacher, D. Rauwolf

RWTH Aachen, SoSe 2022

Ausgabe: 7. April 2022 —— Präsentation der Lösungen: 14. April 2022

Einführung in die angewandte Stochastik

1. Globalübung

Aufgabe 1

Die von einer Bäckerei angebotenen Brotsorten unterscheiden sich u. a. in den Ausprägungen folgender Merkmale:

Preis pro Stück, Backtemperatur, Zuckerzusatz (ja/nein), Haltbarkeit, Produktionszahl pro Tag, Name, interne Produktnummer.

Geben Sie für jedes Merkmal mögliche Ausprägungen an, und ordnen Sie es jeweils den folgenden Merkmalstypen zu:

qualitativ/nominal, qualitativ/ordinal, quantitativ/diskret, quantitativ/stetig.

Geben Sie darüber hinaus weiter an, welche der Merkmale (speziell) dichotom und welche (speziell) verhältnisskaliert sind.

Aufgabe 2

Bei einer Umfrage in einem Verein werden 50 Mitglieder zu ihrem Familienstand befragt. Die verschiedenen Antwortmöglichkeiten hierbei sind *ledig (l)*, *verheiratet (vh)*, *geschieden (g)* sowie *verwitwet (vw)*. Die Umfrage ergibt die folgenden Antworten:

$$vh\ ,\ g\ ,\ l\ ,\ g\ ,\ vw\ ,\ vh\ ,\ vh\ ,\ l\ ,\ vw\ ,\ g\ ,\ g\ ,\ l\ ,\ vh$$
$$g\ ,\ l\ ,\ g\ ,\ l\ ,\ vh\ ,\ vh\ ,\ vh\ ,\ vh\ ,\ l\ ,\ vw\ ,\ g\ ,\ vh\ ,\ g$$
$$vh\ ,\ g\ ,\ vh\ ,\ vh\ ,\ l\ ,\ vw\ ,\ vw\ ,\ vh\ ,\ l\ ,\ vw\ ,\ g\ ,\ vh\ ,\ g$$
$$vh\ ,\ g\ ,\ vh\ ,\ g\ ,\ vw\ ,\ vh\ ,\ vw\ ,\ vh\ ,\ g\ ,\ vh\ ,\ vh\ .$$

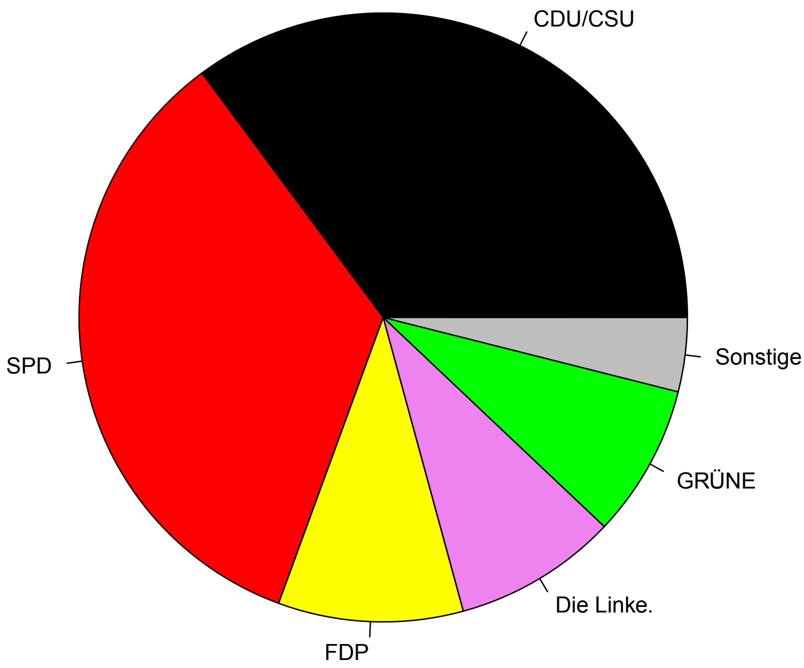
(a) Bestimmen Sie die absoluten und relativen Häufigkeiten der einzelnen Ausprägungen des Merkmals Familienstand.

(b) Stellen Sie die in (a) berechneten absoluten Häufigkeiten in einem Stabdiagramm dar.

(c) Stellen Sie die in (a) berechneten relativen Häufigkeiten in einem Kreisdiagramm dar.

Aufgabe 3

Im folgenden Kreisdiagramm ist die Häufigkeitsverteilung der gültigen Zweitstimmen bei der Bundestagswahl 2005 dargestellt:



Messen Sie die Winkel der dargestellten Kreissektoren, und berechnen Sie hieraus (näherungsweise) die prozentualen Stimmenanteile der einzelnen Parteien.

Aufgabe 4

Bei einem Quiz mussten die 15 Teilnehmer(innen) jeweils 8 Fragen beantworten. Hierbei ergaben sich für die einzelnen Personen die folgenden Anzahlen richtiger Antworten:

$$x_1 = 3\,,\ x_2 = 5\,,\ x_3 = 5\,,\ x_4 = 7\,,\ x_5 = 4\,,\ x_6 = 8\,,\ x_7 = 2\,,\ x_8 = 4\,,$$
$$x_9 = 4\,,\ x_{10} = 6\,,\ x_{11} = 2\,,\ x_{12} = 7\,,\ x_{13} = 6\,,\ x_{14} = 4\,,\ x_{15} = 6\,.$$

(a) Bestimmen Sie die zugehörige Rangwertreihe sowie die Ränge der einzelnen Beobachtungswerte $x_1, \ldots, x_{15}$.

(b) Berechnen Sie zu dem gegebenen Datensatz $x_1, \ldots, x_{15}$ die absoluten und relativen Häufigkeiten der *verschiedenen* Beobachtungswerte, und bestimmen Sie die zugehörigen Modalwerte (d.h. die Ausprägungen mit maximaler (relativer) Häufigkeit unter den verschiedenen Beobachtungswerten).

Aufgabe 5

Gegeben sei die Situation aus der vorherigen Aufgabe 4.

(a) Bestimmen Sie die empirische Verteilungsfunktion F_{15} zu den gegebenen Beobachtungswerten $x_1, \ldots, x_{15}$, und stellen Sie diese graphisch dar.

(b) Kann man den bereits in Aufgabe 4, (b) bestimmten Modalwert direkt aus dem Graphen der empirischen Verteilungsfunktion ablesen?

(c) Berechnen Sie mit Hilfe der empirischen Verteilungsfunktion den Anteil der Quiz-Teilnehmer(innen), die

   (i) höchstens vier Fragen,
   (ii) mehr als zwei Fragen,
   (iii) mindestens drei, aber höchstens fünf Fragen

   richtig beantwortet haben.

(d) Bestimmen Sie den Median zu den gegebenen Beobachtungswerten $x_1, \ldots, x_{15}$, und vergleichen Sie ihn mit dem Modalwert.

(e) Berechnen Sie das untere und obere Quartil sowie das 2-te Dezentil zu den gegebenen Beobachtungswerten $x_1, \ldots, x_{15}$.

Aufgabe 6

Gegeben sei ein metrischer Datensatz $x_1, \ldots, x_n \in \mathbb{R}$.

(a) Beweisen Sie die Steiner-Regel (A 3.31). Zeigen Sie somit, dass für ein beliebiges $a \in \mathbb{R}$ die empirische Varianz s^2 der Beobachtungswerte $x_1, \ldots, x_n$ die folgende Gleichung erfüllt:

$$s^2 \;=\; \left(\frac{1}{n} \sum_{i=1}^{n} (x_i - a)^2 \right) - (\overline{x} - a)^2 \,.$$

   Welche Gleichung ergibt sich hieraus speziell für $a = 0$?

(b) Es seien $a, b \in \mathbb{R}$ und $y_1, \ldots, y_n$ gegeben durch die (affin-) lineare Transformation

$$y_i \;=\; a + b\, x_i \;\;,\;\; i \in \{1, \ldots, n\} \;.$$

   Weiter bezeichnen d_x, d_y die mittleren absoluten Abweichungen, s_x, s_y die empirischen Standardabweichungen sowie

$$V_x \;=\; \frac{s_x}{\overline{x}} \quad \text{und} \quad V_y \;=\; \frac{s_y}{\overline{y}}$$

   die *Variationskoeffizienten* zu $x_1, \ldots, x_n$ bzw. $y_1, \ldots, y_n$. Zeigen Sie:

   (i) $d_y \;=\; |b|\, d_x$,
   (ii) $s_y \;=\; |b|\, s_x$,
   (iii) $V_y \;=\; V_x$ unter den zusätzlichen Annahmen $x_1, \ldots, x_n > 0$, $a = 0$ und $b > 0$.

Hinweis: Verwenden Sie in Aufgabenteil (b), (i) (ohne eigenen Nachweis), dass gemäß Regel A 3.40 gilt:

$$\tilde{y} \;=\; a + b\, \tilde{x} \;.$$

Aufgabe 7

Bei einem Test, der mit den 32 Schülerinnen und Schülern einer Klasse durchgeführt wurde, ergaben sich folgende Messwerte für die zum Lösen einer Rechenaufgabe benötigte Zeit (in Sekunden):

$$\begin{gathered}x_1 = 7.2\ ,\ x_2 = 5.6\ ,\ x_3 = 2.9\ ,\ x_4 = 4.1\ ,\ x_5 = 3.9\ ,\ x_6 = 9.1\ ,\ x_7 = 5.2\ ,\ x_8 = 5.2\ ,\\ x_9 = 2.8\ ,\ x_{10} = 5.1\ ,\ x_{11} = 6.7\ ,\ x_{12} = 8.1\ ,\ x_{13} = 4.1\ ,\ x_{14} = 4.8\ ,\ x_{15} = 8.9\ ,\ x_{16} = 3.0\ ,\\ x_{17} = 7.7\ ,\ x_{18} = 6.7\ ,\ x_{19} = 8.0\ ,\ x_{20} = 4.6\ ,\ x_{21} = 4.8\ ,\ x_{22} = 5.0\ ,\ x_{23} = 7.0\ ,\ x_{24} = 5.3\ ,\\ x_{25} = 6.6\ ,\ x_{26} = 8.8\ ,\ x_{27} = 5.9\ ,\ x_{28} = 6.8\ ,\ x_{29} = 8.9\ ,\ x_{30} = 6.1\ ,\ x_{31} = 5.4\ ,\ x_{32} = 4.7\ .\end{gathered}$$

(a) Berechnen Sie zu diesem Datensatz die zugehörigen Modalwerte, das 5%-Quantil, das 90%-Quantil, den arithmetischen Mittelwert, die Spannweite, den Quartilsabstand, die mittlere absolute Abweichung, die empirische Varianz, die empirische Standardabweichung und den Variationskoeffizienten (vgl. Aufgabe 6).

(b) Erstellen Sie ein Box-Plot zu dem gegebenen Datensatz.

(c) Erstellen Sie ein Histogramm der gemessenenen Zeiten zur folgenden Klasseneinteilung:

$$\begin{gathered}K_1 = [1.5, 2.5]\,,\ K_2 = (2.5, 3.5]\,,\ K_3 = (3.5, 4.5]\,,\ K_4 = (4.5, 5.5]\,,\ K_5 = (5.5, 6.5]\,,\\ K_6 = (6.5, 7]\,,\ K_7 = (7, 7.5]\,,\ K_8 = (7.5, 8.5]\,,\ K_9 = (8.5, 9.5]\,.\end{gathered}$$