

Einführung in die Stochastik

2. Übungsblatt



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Fachbereich Mathematik
M. Kohler
A. Fromkorth
D. Furer

SS 2011
29.04.2011

Gruppen und Hausübung

Aufgabe 5

(4 Punkte)

Im Darmstädter Echo vom 15.01.2009 wurde unter der Überschrift „IM ALLTAG BESTIMMT STRESS DAS ERNÄHRungsverhalten“ folgendes berichtet:

85 Prozent der Deutschen ernähren sich anders als sie es für richtig halten. Das ist ein Ergebnis einer Umfrage zum Ernährungsverhalten der Deutschen, die vom Lebensmittelkonzern Nestlé in Auftrag gegeben wurde. Etwa 4000 Deutsche wurden dazu befragt, Experten des Instituts für Demoskopie in Allensbach und der Unternehmensberatung Boston Consulting Group werteten die Ergebnisse schließlich aus und analysierten sie.

Fast jeder zweite Deutsche hält zwar eine ausgewogene Ernährung für einen wichtigen Aspekt von Lebensqualität. Für zwei Drittel spielt eine gute Ernährung sogar eine große oder sogar sehr große Rolle in ihrem Leben, zumal auch das Gesundheitsbewusstsein zugenommen hat. Doch im Alltag dominieren eine zu einseitige Ernährung, zu viel und unkontrolliertes Essen, Heißhungerattacken und Frustessen.

- Erläutern Sie, was man bei einer Umfrage unter dem sampling bias („Verzerrung durch Auswahl“) und dem non-response bias („Verzerrung durch Nicht-Antwort“) versteht.
- Inwiefern können Sie aufgrund des obigen Artikels eine Aussage über den sampling bias, bzw. den non-response bias machen? Begründen Sie Ihre Antwort. Falls Sie über eine dieser Verzerrungen keine Aussage machen können, so geben Sie bitte an, welche Angabe diesbezüglich in obigem Artikel fehlt.

Lösung:

- Ein sampling bias liegt vor, wenn die Umfrage nicht repräsentativ ist, d.h. das bestimmte Gruppen, deren Antwortverhalten (bezogen auf die Umfrage) vom Rest abweicht, in der Umfrage überrepräsentiert sind. Ein non-response bias liegt vor, wenn ein Teil der Befragten auf einzelne oder alle Fragen nicht antwortet und deren Verhalten sich vom Verhalten der restlichen Befragten unterscheidet.
- Nach den Angaben des Artikels lässt sich über das Vorliegen (und über das nicht Vorliegen) von non-response oder sampling bias nichts aussagen. Um non-response bias belegen oder ausschließen zu können, müsste angegeben werden, wieviele Personen geantwortet haben. Im Falle von sampling bias fehlt die Angabe, wie die Umfrage durchgeführt wurde.

Aufgabe 6

(4 Punkte)

Gegeben sei die Messreihe

4, 5, 1, 3.7, 2.3, 4, 1.3, 2.7, 5, 3.3

- Bestimmen sie das Mittel, den Median, die Spannweite, die Varianz, die Standardabweichung und den Interquartilabstand dieser Messreihe.
- Zeichnen sie ein Histogramm dieser Messreihe bzgl. der Partition $\{[1, 2.5), [2.5, 5]\}$ des Intervalls $[1, 5]$.

Lösung: Datenpunkte der Größe nach aufsteigend geordnet (hier $n = 10 \leftrightarrow$ gerade):

1, 1.3, 2.3, 2.7, 3.3, 3.7, 4, 4, 5, 5



- Mittel: $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \rightarrow \bar{x} = 3.23$
- Median: $\tilde{x} = \begin{cases} x_{\frac{n+1}{2}}, & \text{falls } n \text{ ungerade,} \\ \frac{1}{2}(x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}), & \text{falls } n \text{ gerade.} \end{cases} \rightarrow \tilde{x} = 3.5$
- (a) Spannweite: $r := x_{max} - x_{min} \rightarrow r = 4$
- Varianz: $s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \rightarrow s^2 = 1.9468$
- Standardabweichung: $s = \sqrt{s^2} \rightarrow s = 1.3953$
- Interquartilabstand: $IQR = x_{(\lceil \frac{3}{4}n \rceil)} - x_{(\lfloor \frac{1}{4}n \rfloor)} \rightarrow IQR = 1.7$

Lösung zu Aufgabe 6 b)

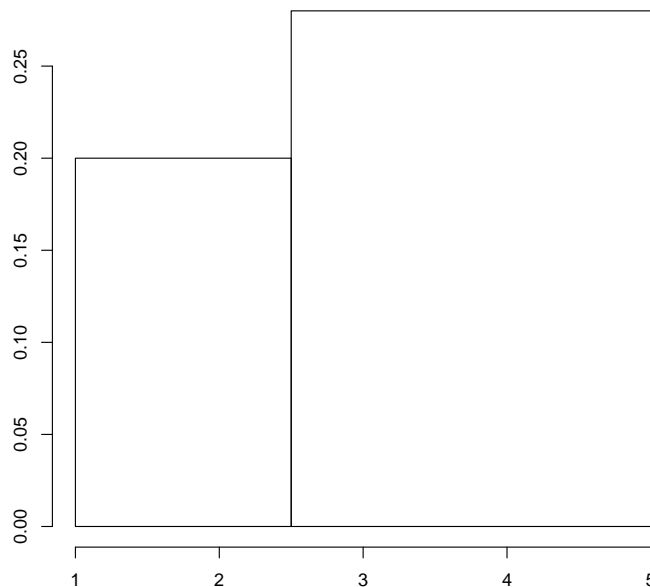


Abbildung 1: Lösung von Aufgabe 6 b)

Aufgabe 7

(4 Punkte)

Am Fachbereich Mathematik der TU Darmstadt ist es üblich, dass bei Veranstaltungen ab einer bestimmten Teilnehmerzahl eine Vorlesungsevaluation durchgeführt wird. Dabei wird folgendermaßen vorgegangen. In einer der Vorlesungen einer festgelegten Woche (üblicherweise gegen Semesterende) werden Fragebögen an die Studierenden verteilt. Diese erhalten danach einige Minuten Zeit zum Ausfüllen der Bögen. Danach wird alles eingesammelt und von der Hochschuldidaktischen Arbeitsstelle ausgewertet. Die Befragung erfolgt anonym.

Bei der Veranstaltung „Mathematik 1 für Bauwesener“ machten im Wintersemester 2010/2011 insgesamt 583 Teilnehmer Angaben zur Häufigkeit des Vorlesungsbesuchs. Die folgende Tabelle gibt an, wieviel Prozent der Teilnehmer die Vorlesung zu 0, 20, 40, 60, 80 bzw. 100 Prozent besuchten

Besuch der Vorlesung zu	0%	20%	40%	60%	80%	100%
Anteil der Studenten	0.7%	0.9%	1%	2.4%	21.6%	73.4%

- (a) Nehmen Sie (kurz) Stellung zu der These: „Das Ergebnis der Vorlesungsevaluation zeigt ganz klar, dass der Vorlesungsbesuch bei der Veranstaltung Mathematik 1 für Bauwesen im WS 2010 / 2011 sehr gut war“.
- (b) Kann bei einer wie oben beschrieben durchgeführten Vorlesungsevaluation ein sampling bias auftreten? Begründen Sie Ihre Antwort.

Lösung:

- (a) Abgesehen davon, dass nicht klar ist, ab wann man von einem sehr guten Vorlesungsbesuch reden kann, konnten natürlich nur die anwesenden Studenten die Evaluationsbögen ausfüllen. Es ist anzunehmen, dass das Ergebnis anders ausgesehen hätte, wenn man auch Personen befragt hätte, die nicht anwesend waren.
- (b) Ja, da nur anwesende Personen befragt werden. Das Antwortverhalten von nicht anwesenden Personen könnte sich vom Antwortverhalten der Anwesenden unterscheiden.

Aufgabe 8

(4 Punkte)

- (a) Der Boxplot in Abbildung 2 beschreibt die Wohnfläche (in Quadratmetern) von 100 zufällig ausgewählten Einwohnern Hessens im Jahr 2005. Wie groß ist bei diesen Daten der Median und der Interquartilsabstand (wegen evt. Problemen beim genauen Ablesen genügt hier eine ungefähre Angabe der Größen)?
- (b) Das Histogramm in Abbildung 3 beschreibt die monatliche Kaltmiete (d.h. die Miete ohne Nebenkosten) von 100 zufällig ausgewählten Einwohnern Hessens im Jahr 2005. Wieviele dieser 100 Einwohner Hessens haben im Jahr 2005 zwischen 200 und 500 Euro pro Monat für die Kaltmiete ihrer Wohnung gezahlt? Wegen eventuellen Problemen beim Ablesen von Werten in dem unten stehenden Histogramm genügt hier wieder eine ungefähre Angabe dieses Wertes.

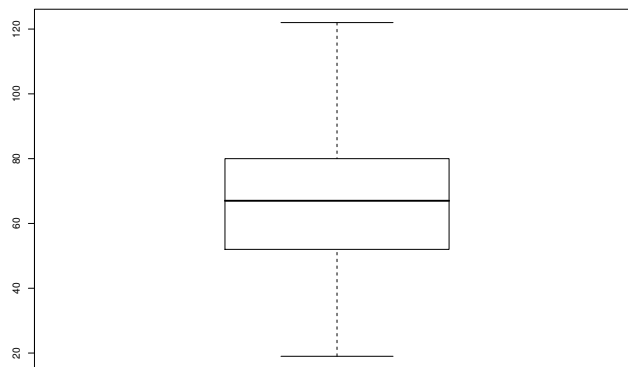


Abbildung 2: Wohnfläche von 100 zufällig ausgewählten Einwohnern Hessens

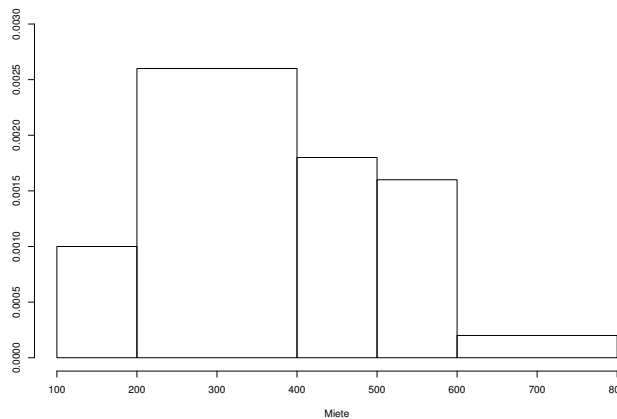


Abbildung 3: Kaltmiete von 100 zufällig ausgewählten Einwohnern Hessens im Jahr 2005

Lösung:

(a)

$$\text{Median} \approx 65$$

$$\text{IQR} \approx 80 - 55 = 25.$$

(b) Seien $I_2 = [200, 400)$, $I_3 = [400, 500)$. Gesucht ist $n_2 + n_3$, wobei n_2 die Anzahl der Datenpunkte über dem Intervall I_2 und n_3 die Anzahl der Datenpunkte über dem Intervall I_3 bezeichne. Ablesen am Histogramm ergibt einen Wert von 0.0026 über dem Intervall I_2 bzw. 0.0018 über I_1 . Es folgt

$$0.0026 = \frac{n_2}{100 \cdot \lambda(I_2)} = \frac{n_2}{100 \cdot 200}$$

$$0.0018 = \frac{n_3}{100 \cdot \lambda(I_3)} = \frac{n_3}{100 \cdot 100}$$

und somit

$$n_2 + n_3 = 0.0026 \cdot 100 \cdot 200 + 0.0018 \cdot 100 \cdot 100 = 52 + 18 = 70.$$

Von den 100 Befragten zahlen also ca. 70 Befragte zwischen 200 und 500 Euro im Monat für die Kaltmiete ihrer Wohnung.

Dieses Übungsblatt wird im Rahmen der Übungen am 02. bzw. 03.05.2011 besprochen. Ihre Ausarbeitungen geben Sie am 09. bzw. 10.05.2011 in Ihrer Übungsgruppe ab. Sie erhalten diese am 16. bzw. 17.05.2011 korrigiert zurück.