Technische Universität Dortmund Wintersemester 2018/2019
Fakultät Statistik 15. November 2018
Jun.-Prof. Dr. Andreas Groll
M.Sc. Hendrik van der Wurp

Übungen zur Vorlesung Wahrscheinlichkeitsrechnung und Mathematische Statistik (für Informatiker)

Blatt 3 - Lösungsvorschlag

Aufgabe 8:

(a) Geben Sie alle Werte für x_3 an, für die der Variationskoeffizient den Wert 1 hat.

$$x_1 = 0, \qquad x_2 = 2, \qquad x_3 = ?$$

Lösungsvorschlag:

Seien $x_1 = 0$, $x_2 = 2$ und $x_3 = y$. Damit gilt $\bar{x} = \frac{0+2+y}{3}$. Der Variationskoeffizient $V_x = \frac{s_x}{\bar{x}}$ ist genau dann eins, wenn $s_x = \bar{x}$ gilt. Bestimme also s_x über s_x^2:

$$\begin{aligned}
& \frac{1}{3-1}\left(\left(0-\frac{2+y}{3}\right)^2+\left(2-\frac{2+y}{3}\right)^2+\left(y-\frac{2+y}{3}\right)^2\right) \\
= \; & \frac{1}{2}\left(\left(\frac{-2-y}{3}\right)^2+\left(\frac{6-2-y}{3}\right)^2+\left(\frac{3y-2-y}{3}\right)^2\right) \\
= \; & \frac{1}{2}\left(\frac{(-y-2)^2}{9}+\frac{(-y+4)^2}{9}+\frac{(2y-2)^2}{9}\right) \\
= \; & \frac{1}{2}\frac{(y^2+4y+4)+(y^2-8y+16)+(4y^2-8y+4)}{9} \\
= \; & \frac{1}{18}(6y^2-12y+24) \quad = \quad s_x^2
\end{aligned}$$

Mit $s_x = \bar{x} \quad \Leftrightarrow \quad s_x^2 = \bar{x}^2$:

$$\begin{aligned}
& \frac{1}{18}(6y^2-12y+24) &=& \left(\frac{2+y}{3}\right)^2 \\
\Leftrightarrow \; & \frac{1}{18}(6y^2-12+24) &=& \frac{1}{9}(y^2+4y+4) \\
\Leftrightarrow \; & 6y^2-12y+24 &=& 2y^2+8y+8 \\
\Leftrightarrow \; & 4y^2-20y+16 &=& 0 \\
\Leftrightarrow \; & y^2-5y+4 &=& 0 \\
\Leftrightarrow \; & \ldots \\
\Rightarrow \; & y \in \{1,4\}
\end{aligned}$$

Kurze Probe: Sei $x_3 = 1$, d.h. die Daten $(0,2,1)$. $\bar{x} = 1$ und $s_x^2 = \frac{1}{2}(1+1+0) = 1$. Somit $V_x = \frac{1}{1} = 1$.
Für $x_3 = 4$, d.h. Daten $(0,2,4)$. $\bar{x} = 2$ und $s_x^2 = \frac{1}{2}(4+4+0) = 4$. Somit $V_x = \frac{\sqrt{(4)}}{2} = \frac{2}{2} = 1$.

(b) Betrachten Sie eine beliebige Stichprobe $(x_1, x_2, x_3, x_4, x_5)$ mit $x_i \in \mathbb{R}, i = 1, \ldots, 5$ mit bekannten Werten. Wie viele Werte der Stichprobe müssten Sie verändern, um

  (i) das arithmetische Mittel,
  (ii) den Median,
  (iii) den Variationskoeffizienten,
  (iv) die mittlere absolute Medianabweichung MD bzw.
  (v) die mediane absolute Medianabweichung MAD

beliebig groß werden zu lassen? Zeigen Sie je ein möglichst einfaches Beispiel.

Lösungsvorschlag:

(i) Nur einen. $x = (1, 2, 3, 4, 10^6)$, $\bar{x} = 200002$

(ii) Drei. $x = (1, 2, 3, 4, 10^6)$, $x_{\text{median}} = 3$. Und mit $x = (1, 2, 3, 10^6, 10^6)$ noch immer $x_{\text{median}} = 3$. Erst $x = (1, 2, 10^6, 10^6, 10^6)$ führt zu explodierendem Median.

(iii) Nur einen, da $\bar{x}$ bereits bei einer einzelnen Änderung explodieren kann und V_x direkt von $\bar{x}$ abhängt.

(iv) Nur einen, da MD die Distanz von allen Werten zum Median berechnet. $x = (1, 2, 3, 4, 10^6)$ mit $x_{\text{median}} = 3$ führt zu einem MD $= \frac{1}{5}(|1-3| + |2-3| + |3-3| + |4-3| + |10^6 - 3|) = \frac{1}{5}(1 + 10^6)$

(v) Drei, mit äquivalenter Begründung wie in (ii). Sobald der Median explodiert, explodiert folglich auch der MAD.

Aufgabe 9: (per Hand)

Betrachten Sie folgende Umfrageergebnisse zur Landtagswahl 2018 in Hessen:

Wählergruppe	CDU	SPD	Grüne	Linke	FDP	AfD
Frauen	270	200	230	60	60	90
Männer	260	190	170	70	80	170

Tabelle 1: Umfrageergebnisse zur Landtagswahl in Hessen

(Um den Faktor 10 abgewandelt aus: Forschungsgruppe Wahlen, ZDF. `https://www.zdf.de/nachrichten/heute/landtagswahl-in-hessen-im-ueberblick-100.html`, Slides 23-24)

(a) Berechnen Sie Randsummen und stellen Sie die obenstehenden Daten als Kontingenztafeln (mit Randsummen) mit absoluten und relativen Werten dar.

(b) Berechnen Sie bedingten Verteilungen auf *Geschlecht* sowie auf *Partei* und stellen Sie diese als Kontingenztafeln dar.

Lösungsvorschlag:

Zuerst mit absoluten Werten:

	Partei						
	CDU	SPD	Grüne	Linke	FDP	AfD	Σ
Frauen	270	200	230	60	60	90	910
Männer	260	190	170	70	80	170	940
Σ	530	390	400	130	140	260	1850

Und anschließend mit relativen, d.h. jeweils von 1850. Bspw. $\frac{270}{1850} = 0.146$

	Partei						
	CDU	SPD	Grüne	Linke	FDP	AfD	Σ
Frauen	0.146	0.108	0.124	0.032	0.032	0.049	0.492
Männer	0.141	0.103	0.092	0.038	0.043	0.092	0.508
Σ	0.286	0.211	0.216	0.070	0.076	0.141	1

Bedinge zuerst auf das Geschlecht: Bspw: $\frac{0.124}{0.492} = 0.253$

	Partei						
	CDU	SPD	Grüne	Linke	FDP	AfD	Σ
Frauen	0.297	0.220	0.253	0.066	0.066	0.099	1
Männer	0.277	0.202	0.181	0.074	0.085	0.181	1

Bedinge dann auf die Parteien: Bspw: $\frac{0.038}{0.070} = 0.538$

	Partei					
	CDU	SPD	Grüne	Linke	FDP	AfD
Frauen	0.509	0.513	0.575	0.462	0.429	0.346
Männer	0.491	0.487	0.425	0.538	0.571	0.654
Σ	1	1	1	1	1	1