

Statistik 1, Test 1

Aufgabe 1

(a) Der Mittelwert ist

$$\begin{aligned}\mu &= \frac{16 \cdot 180 + 17 \cdot 170 + 29 \cdot 160 + 37 \cdot 150 + 22 \cdot 140 + 31 \cdot 130 + 37 \cdot 120 + 48 \cdot 110 + 31 \cdot 100 + 27 \cdot 90 + 70 \cdot 80}{365} \\ &= 120.3288\end{aligned}$$

(b) Der Median ist das mittlere Element der Messreihe, also das $\lceil \frac{365}{2} \rceil = 183$. Element. Das ist 120.

(c) Die empirische Varianz ist

$$\begin{aligned}\hat{s}^2 &= \frac{16(180 - 120.3288)^2 + 17(170 - 120.3288)^2 + \dots + 70(80 - 120.3288)^2}{365} \\ &= 945.6453\end{aligned}$$

(d) Die Stichprobenvarianz ist

$$\begin{aligned}s^2 &= \frac{16(180 - 120.3288)^2 + 17(170 - 120.3288)^2 + \dots + 70(80 - 120.3288)^2}{364} \\ &= 948.2433\end{aligned}$$

(e) Das 25%-Quantil ist das $\lceil 0.25 \cdot 365 \rceil = 92$. Element, also 90. **Im Forum steht, dass der Test aber 150 als richtigen Wert akzeptiert.**

(f) Das 75%-Quantil ist das $\lceil 0.75 \cdot 365 \rceil = 274$. Element, also 150. **Im Forum steht, dass der Test aber 90 als richtigen Wert akzeptiert.**

In K_1 sind $16 + 17 + 29 + 37 + 22 + 31 = 152$ Studenten, in K_2 sind $37 + 48 + 31 = 116$ Studenten, in K_3 sind $27 + 70 = 97$ Studenten. Die Klassenmitten sind 155, 110, 85, damit ist der Mittelwert der klassierten Daten

$$\begin{aligned}\mu &= \frac{152 \cdot 155 + 116 \cdot 110 + 97 \cdot 85}{365} \\ &= 122.0959\end{aligned}$$

Mindestens 150 Punkte haben $\frac{16+17+29+37}{365} = 0.2712329$ der Studenten erreicht.

Nur 90 oder weniger Punkte haben $\frac{29+70}{365} = 0.2657 = \hat{F}(90)$ der Studenten erreicht. **Der Test nimmt aber nur 0.808219.**

Übrigens scheinen das die Prüfungsergebnisse der letzten Statistik 1 Klausur im SS2020 gewesen zu sein: https://wwilq.wiwi.tu-dresden.de/ects_v2/index.php?action=output&jahr=2020&status=SS&typ=PF&prn=14530

Aufgabe 2

Der Mittelwert ist

$$\begin{aligned}\mu &= \frac{24 + 21 + 29 + 14 + 27 + 13 + 28 + 13}{8} \\ &= 21.125\end{aligned}$$

Die Tabelle sieht so aus

	K_1	K_2	K_3
$n(K_i)$	3	2	3
$h(K_i)$	0.375	0.25	0.375

Und der Mittelwert der klassierten Daten ist 20.

Aufgabe 3

Die Aussagen sind:

- Das Merkmal Lieblingsprimzahl ist stetig, da es unendlich viele Primzahlen gibt. FALSCH, stetig wäre das Merkmal nur, wenn es überabzählbar viele Primzahlen gibt, aber die gibt es nicht.
- Die entscheidende Information bei einem Histogramm ist nicht die Höhe der Balken, sondern die Fläche! RICHTIG
- Im Gegensatz zur induktiven Statistik beziehen sich die Ergebnisse der deskriptiven Statistik nur auf die untersuchte Datenmenge. RICHTIG
- Ein Merkmalsvektor X ist eine Abbildung $\Omega \rightarrow S$, wobei S der Merkmalsraum ist. RICHTIG
- Die empirische Verteilungsfunktion \hat{F} ist linksseitig stetig und monoton steigend. FALSCH, die Funktion ist rechtsseitig stetig.

Aufgabe 4

Richtig ist

- Sie können das arithmetische Mittel verwenden, wenn Sie die Vermögensangabe von Beate Heister und Karl Albrecht Junior aus Ihrem Datensatz entfernen.
- Sie können den Median verwenden, da dieser robust in Bezug auf Ausreißer ist.

Aufgabe 5

Die Merkmale sind:

- Adresse eines Studenten: nominal, qualitativ, diskret
- Preis eines Mensa-Essens: metrisch, quantitativ, quasistetig
- Hobby eines Studenten: nominal, qualitativ, diskret
- durchschnittliche Länge eines Stückes Kreide im Hörsaal: metrisch, quantitativ, stetig
- Datum der Klausur Statistik I: metrisch, quantitativ, diskret

Aufgabe 6

Term D, Term B, Term B, $\lambda = 0.25$, $L = 6.58 \cdot 10^{-6}$