

STATISTIK I – Übung 01

Das arithmetische Mittel

1 Kurze Wiederholung

Grundlagen

Das arithmetische Mittel gehört zu den Lagemaßen und ist somit als eine Maßzahl zu verstehen, die das Zentrum einer Verteilung beschreibt. Als das bekannteste Lagemaß wird das arithmetische Mittel häufig auch als „das Standardmittel“ oder einfach nur als „der Mittelwert“ bezeichnet. Seine Berechnung setzt voraus, dass die Daten der Verteilung mindestens metrisch skaliert sind – was in der Praxis (etwa bei Schulnoten) bedauerlicherweise häufig übersehen wird. Ist diese Vorbedingung erfüllt, berechnet sich das arithmetische Mittel wie folgt:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

mit:

\bar{x} = *Arithmetisches Mittel*

x_i = *Einzelner Wert der Verteilung*

n = *Anzahl der Werte der Verteilung*

Arithmetisches Mittel bei klassierten Daten

Neben der „einfachen“ Berechnung des arithmetischen Mittels wird für unsere Klausur noch die Berechnung des arithmetischen Mittels bei klassierten Daten von Relevanz sein. Hierfür wird auf folgende Formel zurückgegriffen:

$$\bar{x}_g = \sum_{i=1}^k m_i * f_i$$

mit:

\bar{x}_g = *Arithmetisches Mittel (bei klassierten/grupperten Daten)*

f_i = *relative Häufigkeit (der Klasse i)*

m_i = *Klassenmitte (der Klasse i)*

k = *Anzahl der Klassen*

Robustheit und getrimmtes arithmetisches Mittel

Das arithmetische Mittel wird von (univariaten) Ausreißern (besonders großen oder kleinen Werten im Datensatz) ganz erheblich beeinflusst und wird deshalb auch als „nicht robust“ bezeichnet. Diesen Effekt kann man sich leicht vor Augen führen, indem man sich klarmacht, dass das arithmetische Mittel der Verteilung [1; 2; 3; 4] bei 2,5 liegt, das arithmetische Mittel aus der Verteilung [1; 2; 3; 50] aber bei 14. Bereits ein einzelner besonders großer oder kleiner Wert (wie etwa ein exorbitantes Einkommen bei einer Vermögenserhebung), kann daher den Mittelwert deutlich nach oben oder unten verzerren.

Um diesen Effekt zu begrenzen, kann man entweder auf ein anderes Lagemaß (wie etwa den Median) ausweichen, oder das sogenannte getrimmte arithmetische Mittel berechnen. Hierbei wird der

Datensatz vor der Berechnung des arithmetischen Mittels um eine gewisse Anzahl an Werten an den Rändern der Verteilung (symmetrisch) gekürzt, um Ausreißer aus dem Datensatz zu eliminieren. Bei einem Datensatz mit 100 Werten würden bei einer Trimmung um 5% zum Beispiel die 5 größten sowie die 5 kleinsten Werte aus dem Datensatz entfernt und anschließend das arithmetische Mittel auf Basis der bereits bekannten Formel neu berechnet. Dabei ist zu beachten, dass in vielen Fällen auch Nicht-Ausreißer aus den Daten gestrichen werden, die man im Grunde behalten möchte.

Hinweis zu softwaregestützten Analysen

Wird für die Berechnung des arithmetischen Mittels eine Software wie etwa SPSS, PSPP oder PAST eingesetzt, so ist – wie bei vielen anderen Berechnungen auch – zu berücksichtigen, dass die Erfüllung von Vorbedingungen für die Analyse in der Regel nicht von der Software geprüft wird. In diesem Fall betrifft dies das Vorliegen eines metrischen Skalenniveaus. SPSS berechnet das arithmetische Mittel fälschlicherweise nicht nur für Schulnoten, sondern auch für Telefonnummern oder Geschlechter (falls diese mit Zahlen codiert sein sollten) – auch wenn die Ergebnisse vollkommen sinnbefreit sind. Als ganz besonders gefährlich dürfen übrigens solche Fehler betrachtet werden, die – zumindest oberflächlich betrachtet – sinnvolle Ergebnisse darzustellen scheinen (wie eben beim Mittel aus Schulnoten). Beim Einsatz von Software ist daher entscheidend, dass der Anwender / die Anwenderin über die Methodenkenntnisse verfügt, um selbst beurteilen zu können, wann welche Methode zulässig ist.

2 Beispielrechnungen

Arithmetisches Mittel

Für eine Gruppe von Studierenden liegt die folgende Altersverteilung vor:

Alter	Absolute Häufigkeit	Relative Häufigkeit	Kumulierte abs. Häufigkeit	Kumulierte rel. Häufigkeit
21 Jahre	5	0,25	5	0,25
22 Jahre	4	0,20	9	0,45
23 Jahre	3	0,15	12	0,60
24 Jahre	4	0,20	16	0,80
25 Jahre	4	0,20	20	1,00
Summe	20	1,00	20	1,00

Das arithmetische Mittel berechnet sich in diesem Fall wie folgt:

$$21+21+21+21+21+22+22+22+22+23+23+23+24+24+24+24+25+25+25+25 = 458$$

$$458 / 20 = \underline{22,9}$$

Alternative Vereinfachung:

$$(21*5) + (22*4) + (23*3) + (24*4) + (25*4) = 458$$

$$458 / 20 = \underline{22,9}$$

Das arithmetische Mittel liegt somit bei 22,9 Jahren.

Arithmetisches Mittel bei klassierten Daten

Für eine Gruppe von Studierenden liegt die folgende Größenverteilung vor:

Größe	Absolute Klassenhäufigkeit	Relative Klassenhäufigkeit	Absolute kum. Klassenhäufigkeit	Relative kum. Klassenhäufigkeit
[1,55 m – 1,65 m)	12	0,24	12	0,24
[1,65 m – 1,75 m)	16	0,32	28	0,56
[1,75 m – 1,85 m)	22	0,44	50	1,00
Summe	50	1,00	50	1,00

Das arithmetische Mittel berechnet sich in diesem Fall wie folgt:

$$(0,24 * 1,60) + (0,32 * 1,70) + (0,44 * 1,80) = \underline{1,72}$$

Das arithmetische Mittel liegt somit bei 1,72 Metern.

Getrimmtes arithmetisches Mittel

Eine Umfrage unter 10 Personen zum monatlichen Bruttoeinkommen erbrachte folgende Ergebnisse:

2.250 EUR	2.130 EUR
2.320 EUR	2.640 EUR
2.400 EUR	2.550 EUR
2.140 EUR	2.250 EUR
17.380 EUR	2.710 EUR

Das arithmetische Mittel berechnet sich in diesem Fall wie folgt:

$$2250 + 2320 + 2400 + 2140 + 17380 + 2130 + 2640 + 2550 + 2250 + 2710 = 38770$$

$$38770 / 10 = \underline{3877}$$

Das arithmetische Mittel liegt bei 3.877 EUR. Da es offenkundig vom Ausreißer stark beeinflusst wird (alle befragten Personen außer einer verdienen zwischen 2.100 und 2.800 EUR – trotzdem liegt der „Mittelwert“ bei fast 4.000 EUR), soll nachfolgend das um 10% getrimmte Mittel berechnet werden.

Bei einer 10%igen Trimmung sind der größte (17.380 EUR) und der kleinste (2.130 EUR) Wert aus dem Datensatz zu entfernen. Es ergibt sich die folgende neue Grundtabelle:

2.250 EUR	2.640 EUR
2.320 EUR	2.550 EUR
2.400 EUR	2.250 EUR
2.140 EUR	2.710 EUR

Das getrimmte arithmetische Mittel berechnet sich in diesem Fall wie folgt:

$$2250 + 2320 + 2400 + 2140 + 2640 + 2550 + 2250 + 2710 = 19260$$

$$19260 / 8 = \underline{2407,5}$$

Das getrimmte arithmetische Mittel liegt somit (deutlich realitätsnäher) bei 2.407,50 EUR.

3 Übungsaufgaben (Lösungen folgen in der kommenden Woche)

Arithmetisches Mittel

Aus einem Produktionslos von 1.000 Karosserieteilen wird eine Stichprobe von 20 Teilen gezogen und gewogen. Es ergeben sich die folgenden Werte:

Teil 1	1,72 kg	Teil 6	1,74 kg	Teil 11	1,76 kg	Teil 16	1,76 kg
Teil 2	1,74 kg	Teil 7	1,73 kg	Teil 12	1,77 kg	Teil 17	1,77 kg
Teil 3	1,72 kg	Teil 8	1,73 kg	Teil 13	1,71 kg	Teil 18	1,71 kg
Teil 4	1,72 kg	Teil 9	1,76 kg	Teil 14	1,72 kg	Teil 19	1,75 kg
Teil 5	1,73 kg	Teil 10	1,72 kg	Teil 15	1,77 kg	Teil 20	1,76 kg

- a) Fassen Sie diese Werte in einer kumulierten Häufigkeitstabelle (ohne Klassierung) zusammen.
- b) Berechnen Sie das arithmetische Mittel.

Arithmetisches Mittel bei klassierten Daten

Eine Gruppe von 50 Studierenden wird nach ihrem ungefähren Lernaufwand für eine Statistikklausur (in Tagen) befragt. Es ergeben sich die folgenden (klassierten) Werte:

Größe	Absolute Klassenhäufigkeit	Relative Klassenhäufigkeit	Absolute kum. Klassenhäufigkeit	Relative kum. Klassenhäufigkeit
[1 Tag – 3 Tage)	17			
[3 Tage – 5 Tage)	23			
[5 Tage – 7 Tage)	10			
Summe				

- a) Füllen Sie die restlichen Felder der kumulierten Häufigkeitstabelle aus.
- b) Berechnen Sie das arithmetische Mittel.

Getrimmtes arithmetisches Mittel

Eine Gruppe von Studierenden befragt Passantinnen und Passanten auf dem Campus. Erhoben wird dabei unter anderem das Alter (in Jahren). Hierfür ergeben sich für zwanzig Personen folgende Werte:

Person 1	23 Jahre	Person 11	22 Jahre
Person 2	24 Jahre	Person 12	23 Jahre
Person 3	19 Jahre	Person 13	37 Jahre
Person 4	21 Jahre	Person 14	30 Jahre
Person 5	23 Jahre	Person 15	85 Jahre
Person 6	26 Jahre	Person 16	21 Jahre
Person 7	31 Jahre	Person 17	12 Jahre
Person 8	27 Jahre	Person 18	11 Jahre
Person 9	34 Jahre	Person 19	27 Jahre
Person 10	23 Jahre	Person 20	28 Jahre

- a) Berechnen Sie das um 5% getrimmte arithmetische Mittel.
- b) Berechnen Sie das um 10% getrimmte arithmetische Mittel.