STATISTIK I – Übung 03
Statistische Lagemaße

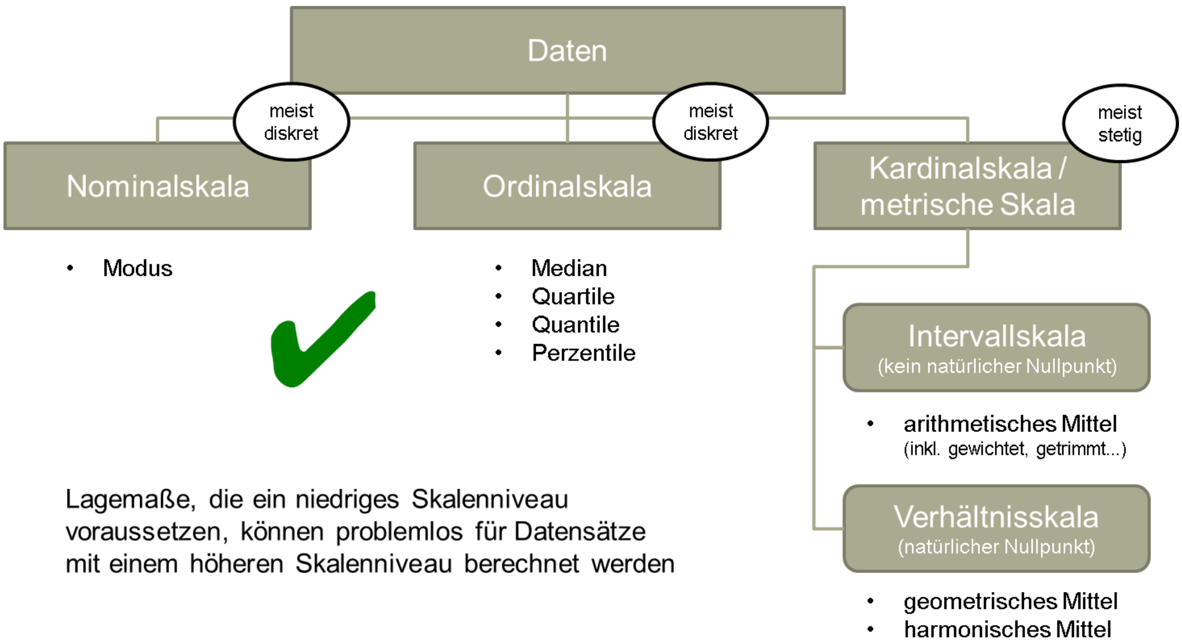
1 Kurze Wiederholung

Was sind statistische Lagemaße?

Die statistischen Lagemaße werden auch als Mittelwerte oder Maße der zentralen Tendenz bezeichnet. Sie geben Auskunft über das Zentrum einer Verteilung und sind am besten geeignet, wenn es gilt, eine Verteilung mit nur einem Parameter zusammenzufassen – wie etwa die Einkommensverteilung mit der Angabe des Durchschnittseinkommens. (Warum nur ein Parameter für die Darstellung einer Verteilung in der Regel eben nicht ausreicht, wird im nächsten Übungsabschnitt zum Thema Streuungsmaße erläutert.) Im Rahmen der Vorlesung haben wir uns insbesondere mit drei Lagemaßen – dem arithmetischen Mittel (Durchschnittswert der Verteilung), dem Median (Wert genau in der Mitte der geordneten Verteilung) und dem Modus (in der Verteilung am häufigsten auftretender Wert) – beschäftigt. Die beiden kurz betrachteten Sonderfälle des geometrischen Mittels (für relative Veränderungen) sowie des harmonischen Mittels (für Quotienten) sind nicht klausurrelevant, dafür allerdings die Berechnung beliebiger Perzentile als Verallgemeinerung des Modus. In dieser Übungseinheit wollen wir uns noch einmal allen Lagemaßen gemeinsam widmen, bevor es in der nächsten Übungseinheit dann um zwei Streuungsmaße (Interquartilsabstand und Spannweite) gehen soll.

Lagemaße und Skalenniveaus

Wie wir ebenfalls gelernt haben, entscheidet sich die Frage, welches Lagemaß für eine beliebige Verteilung berechnet werden kann, am Skalenniveau der Daten. Dabei ist zu beachten, dass die Lagemaße zwar „aufwärtskomptibel“, nicht aber „abwärtskompatibel“ sind. Liegen metrisch skalierte Daten vor, kann neben dem arithmetischen Mittel also auch der Median, oder (falls die Verteilung ein eindeutiges Maximum aufweist) der Modus berechnet werden – liegen dagegen ordinalskalierte Daten vor, ist die Berechnung des arithmetischen Mittels definitiv nicht möglich. Statistische Lagemaße, die ein niedrigeres Skalenniveau voraussetzen, können also auch auf Daten eines höheren Skalenniveaus angewandt werden – nicht jedoch umgekehrt. Die nachfolgende Grafik verdeutlicht noch einmal, welches statistische Lagemaß ab welchem Skalenniveau zum Einsatz kommen kann.



2 Beispielrechnungen

Im Rahmen einer Erhebung unter 100 Mieterinnen und Mietern wurde unter anderem nach der Größe des Badezimmers (gerundet auf ganze Zahlen) gefragt. Es ergaben sich folgende Werte:

Badgröße	Probanden/innen	Badgröße	Probanden/innen
7 m²	13	12 m²	7
8 m²	8	13 m²	23
9 m²	12	14 m²	5
10 m²	6	15 m²	6
11 m²	12	16 m²	8

Nachfolgend sollen das arithmetische Mittel, das um 5% getrimmte arithmetische Mittel, die Quartile (oberes/unteres Quartil und Median) sowie der Modus berechnet werden.

<u>Berechnung des arithmetischen Mittels</u>

13 + 8 + 12 + 6 + 12 + 7 + 23 + 5 + 6 + 8 = 100

(7 * 13) + (8 * 8) + (9 * 12) + (10 * 6) + (11 * 12) + (12 * 7) + (13 * 23) + (14 * 5) + (15 * 6) + (16 * 8)
= 91 + 64 + 108 + 60 + 132 + 84 + 299 + 70 + 90 + 128 = 1126

1126 / 100 = <u>11,26</u>

Das arithmetische Mittel liegt bei 11,26 m².

<u>Berechnung des um 5% getrimmten arithmetischen Mittels</u>

Bei Trimmung eines Datensatzes der Größe n = 100 um 5% fallen jeweils 5 Werte (100 * 0,05) am oberen und unteren Ende der Verteilung weg. Im vorliegenden Fall sind also 5 x 7 m² sowie 5 x 16 m² zu streichen. Anschließend wird das arithmetische Mittel neu berechnet.

8 + 8 + 12 + 6 + 12 + 7 + 23 + 5 + 6 + 3 = 90

(7 * 8) + (8 * 8) + (9 * 12) + (10 * 6) + (11 * 12) + (12 * 7) + (13 * 23) + (14 * 5) + (15 * 6) + (16 * 3)

= 56 + 64 + 108 + 60 + 132 + 84 + 299 + 70 + 90 + 48 = 1011

1011 / 90 = <u>11,23</u>

Das um 5% getrimmte arithmetische Mittel liegt also bei 11,23 m². Die vergleichsweise geringe Distanz zum ungetrimmten Mittel lässt vermuten, dass sich im Datensatz keine Ausreißer befinden (die sich in der Tabelle auch objektiv nicht identifizieren lassen).

<u>Berechnung der Quartile</u>

Für die Berechnung des Median bzw. der Quartile empfiehlt sich die Anlage der Tabelle der kumulierten Häufigkeitsverteilung, da sich auf diesem Wege die Bildung einer geordneten Reihe mit sämtlichen 100 Werten vermeiden lässt und man schneller zum gewünschten Ergebnis kommt. Der Median ist in dieser Tabelle aus der Zeile abzulesen, in welcher der Wert der kumulierten Häufigkeitsverteilung die 0,50 überschreitet. Ebenso sind das obere und das untere Quartil aus den Zeilen abzulesen, in denen der Wert der kumulierten Häufigkeitsverteilung die 0,25 bzw. die 0,75 überschreitet.

Badgröße	Abs. Häufigkeit	Rel. Häufigkeit	Kumulierte abs. Häufigkeit	Kumulierte rel. Häufigkeit
7 m²	13	0,13	13	0,13
8 m²	8	0,08	21	0,21
9 m²	12	0,12	33	0,33
10 m²	6	0,06	39	0,39
11 m²	12	0,12	51	0,51
12 m²	7	0,07	58	0,58
13 m²	23	0,23	81	0,81
14 m²	5	0,05	86	0,86
15 m²	6	0,06	92	0,92
16 m²	8	0,08	100	1,00
Summe	100	1,00	100	1,00

Aus der Tabelle lassen sich die drei Quartile nun wie folgt direkt ablesen:

Unteres Quartil: 9 m²
Mittleres Quartil: 11 m²
Oberes Quartil: 13 m²

Das mittlere Quartil ist mit dem Median identisch.

<u>Berechnung des Modus</u>

Der Modus ist als der Wert definiert, der in der Verteilung am häufigsten auftritt: 13 m². Da es sich um eine unimodale Verteilung handelt (der Wert 13 m² tritt 23 x auf, der nächsthäufigere Wert lediglich 12 x), kann der Modus in diesem Fall sinnvoll gebildet werden.

3 Übungsaufgaben (Lösungen folgen in der kommenden Woche)

Die 100 Personen aus der vorangegangenen Aufgabe wurden auch zur Größe (ebenfalls gerundet) anderer Räume in ihren Wohnungen befragt. Für das Schlafzimmer ergaben sich folgende Werte:

Schlafzimmergröße	Probanden/innen	Schlafzimmergröße	Probanden/innen
12 m²	11	17 m²	22
13 m²	21	18 m²	6
14 m²	12	19 m²	5
15 m²	7	20 m²	3
16 m²	8	21 m²	5

a) Bestimmen Sie das arithmetische Mittel.

b) Bestimmen Sie das um 10% getrimmte arithmetische Mittel.

c) Bestimmen Sie die drei Quartile.

d) Bestimmen Sie den Modus.