

I Daten und ihre Aufbereitung

1 Grundbegriffe der Datenerhebung

1 Auf Autobahnen wird laufend die Verkehrsdichte (Anzahl der Autos je Stunde) gemessen. Fig. 1 zeigt die Verkehrsdichte im Verlaufe eines Tages auf den drei Spuren des Kölner Autobahnringes. Erläutern Sie die dargestellten Informationen.

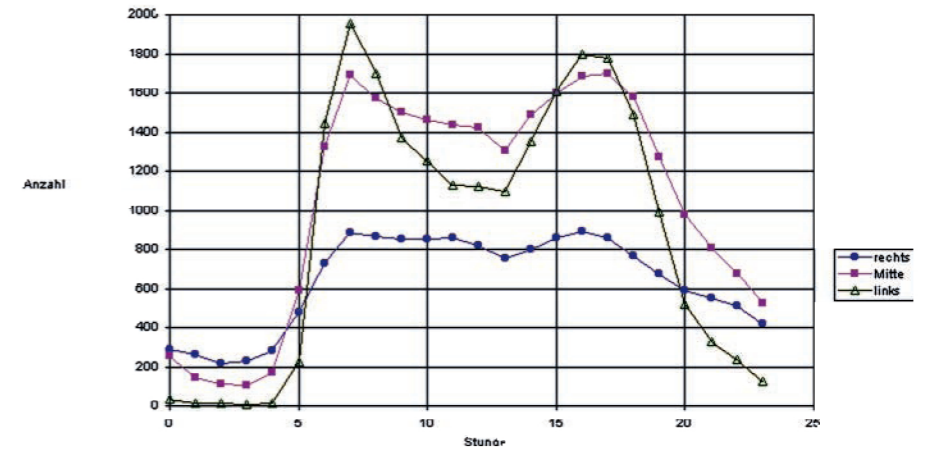


Fig. 1

Befragungen von Personen oder das Zählen von Gegenständen sind Beispiele **statistischer Erhebungen**. Bei einer solchen Erhebung wird an **Merkmalsträgern** (z. B. Personen oder Autos) ein bestimmtes **Merkmal** (z. B. Körpergröße oder gefahrene Geschwindigkeit) untersucht. Alle Merkmalsträger zusammen nennt man **Grundgesamtheit**. Ausgewählte Merkmalsträger bilden eine **Stichprobe**. Die Anzahl n der Merkmalsträger in der Stichprobe heißt **Stichprobenumfang**.

Bei einer statistischen Erhebung werden die Daten in einer **Urliste** wie in Fig. 2 festgehalten. Die einzelnen Beobachtungswerte der Urliste heißen **Stichprobenwerte**; sie werden mit x_1, x_2, \dots, x_n bezeichnet. Sie können mehrfach vorkommen. Davon zu unterscheiden sind die **Merkmalsausprägungen**; sie werden mit a_1, a_2, \dots, a_k bezeichnet. Diese sind alle voneinander verschieden.

In der Urliste von Fig. 2 sind die Merkmalsträger Autos, bei denen die Merkmale Zeit (T) und Geschwindigkeit (v) gemessen wurden. Der Stichprobenumfang ist $n = 2584$. Merkmalsausprägungen sind hier die ganzzahligen Geschwindigkeiten. Natürlich müssen bei 2584 Stichprobenwerten etliche mehrfach auftreten. Beispielsweise gilt $v_{0002} = v_{0125} = 52 \frac{\text{km}}{\text{h}}$.

Nr.	Uhrzeit T	Geschwindigkeit v (in $\frac{\text{km}}{\text{h}}$)	2 Stichprobenwerte zur selben Merkmalsausprägung	
			Merkmalsausprägung	Stichprobenwert
0001	09:45:13	44		
0002	09:46:04	52	←	
...				
0123	11:09:21	33		
0124	11:09:47	62		
0125	11:10:04	52	←	
0126	11:10:21	41		
...				
2583	23:58:47	83		
2584	23:59:13	53		

Fig. 2

Vergleicht man Merkmalsausprägungen verschiedener Merkmale, stellt man Unterschiede fest.

