

Das Histogramm

– Artikel Nr. 4 der Statistik-Serie in der DMW –

The Histogram

Autoren

S. Lange¹ R. Bender¹

Institut

¹ Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, Köln

Das Histogramm ist eine einfache Möglichkeit, einen Datensatz grafisch darzustellen. Hierbei werden die absoluten oder relativen Häufigkeiten, mit denen eine Merkmalsausprägung beobachtet wurde, als Rechtecke abgebildet. Liegen viele verschiedene Merkmalsausprägungen vor, ist es sinnvoll, mehrere Ausprägungen in Klassen zusammenzufassen. Wenngleich es für die Breite und Anzahl der Klassen grobe Empfehlungen gibt (zum Beispiel für die Anzahl der Klassen die Wurzel aus dem Stichprobenumfang: „ \sqrt{n} -Regel“), genügen im Allgemeinen, abgesehen von sehr großen Stichprobenumfängen, 8 bis 15 Klassen. Wichtiger als solche Empfehlungen sind auf jeden Fall inhaltliche Überlegungen, insbesondere auch im Hinblick auf die Wahl der Klassengrenzen. Bei der Darstellung von Altersverteilungen ist es beispielsweise üblich, Dekaden oder 5-Jahres-Intervalle zu verwenden, und die Klasse bei 5er oder 10er Jahreszahlen zu beginnen. Klassenbreiten von 7,74 Jahren, die sich bei strikter Anwendung der „ \sqrt{n} -Regel“ bei einem Stichprobenumfang von 60 ergäben, sind eher nicht adäquat.

Tab. 1 Altersverteilung von Personen, die in einem Londoner Stadtbezirk einen häuslichen Unfall erlitten haben.

Altersgruppe (Jahre)	Häufigkeit
[0–5)	206
[5–15)	154
[15–45)	247
[45–65)	111
[65–90]	95
Σ	813

Die eckigen Klammern bei den Intervallangaben bedeuten, dass der entsprechende Wert noch zum Intervall gehört, die runden Klammern bedeuten, dass der Wert nicht mehr zum Intervall gehört.

Tab. 2 Übersetzungen (deutsch – englisch)

Histogramm	histogram
Absolute (relative) Häufigkeit	absolute (relative) frequency

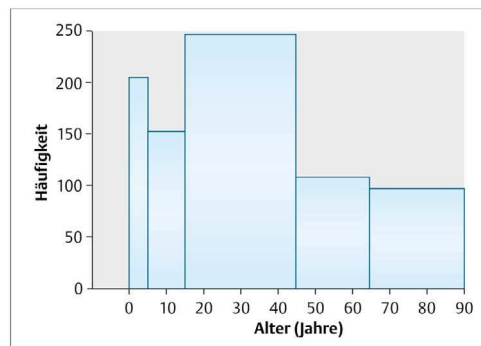


Abb. 1 Fälschliches „Histogramm“ bei ungleichen Klassenbreiten für die Daten aus Tabelle 1: Die Höhe der Rechtecke wird als Maß für die Häufigkeit verwendet.

In einem Histogramm haben die Klassen zumeist die gleiche Breite. Es kann aber auch sinnvoll sein, nicht oder nur schwach besetzte Klassen zusammenzufassen. In dieser Situation mit unterschiedlich breiten Klassen muss allerdings beachtet werden, dass bei einem Histogramm nicht die Höhe, sondern die Fläche der Rechtecke der beobachteten (absoluten oder relativen) Häufigkeit der jeweiligen Klasse entspricht. Um die richtigen Relationen herzustellen, muss bei unterschiedlichen Klassenbreiten dann die Häufigkeit jeder Klasse durch die Breite dieser Klasse geteilt werden.

In **Tab. 1** ist die Altersverteilung von Personen angegeben, die in einem Londoner Stadtbezirk einen häuslichen Unfall erlitten haben. Die meisten Opfer sind Erwachsene mit der größten Zahl von Unfallopfern im Altersbereich zwischen 15 und 44 Jahren. Die Breite der Altersklassen variiert zwischen 5 und 30 Jahren. Eine Darstellung, in der die Höhe der Rechtecke als Maß für die Häufigkeit verwendet wird, ergibt den Eindruck, dass Kleinkinder im Alter bis 5 Jahre bei den Opfern nur wenig vertreten sind, was eigentlich nicht zu erwarten ist (▶ **Abb. 1**). Die richtige

Schlüsselwörter

- ▶ Histogramm
- ▶ Statistische Grafik
- ▶ Verteilung
- ▶ Relative Häufigkeit

Key words

- ▶ Histogram
- ▶ Statistical graphs
- ▶ Distribution
- ▶ Relative frequency

Bibliografie

DOI 10.1055/s-2007-959027
 Dtsch Med Wochenschr 2007; 132: e7–e8 · © Georg Thieme Verlag KG Stuttgart · New York · ISSN 0012-0472

Korrespondenz

Privatdozent Dr. Stefan Lange
 Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG)
 Dillenburger Straße 27
 51105 Köln
 eMail stefan.lange@iqwig.de

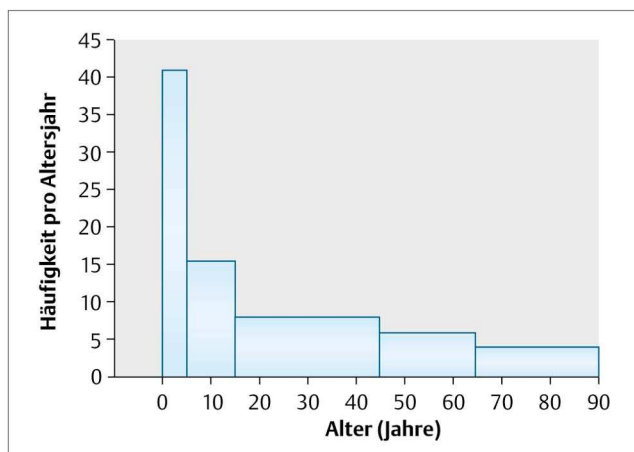


Abb. 2 Richtiges Histogramm bei ungleichen Klassenbreiten für die Daten aus Tabelle 1: Die Fläche der Rechtecke wird als Maß für die Häufigkeit verwendet.

Darstellung, bei der die Fläche der Häufigkeit entspricht, und die Zahl von Unfallopfern pro Altersjahr (Höhe entspricht Häufigkeit dividiert durch Klassenbreite) angegeben ist, korrigiert den fälschlichen Eindruck: Das größte Risiko haben Kleinkinder

► **Abb. 2.**

Tab. 2 enthält die englischsprachigen Übersetzungen der in diesem Beitrag besprochenen Begriffe.

kurzgefasst

In einem Histogramm werden Häufigkeiten als Rechtecke abgebildet. Zumeist haben hierbei die Rechtecke der verschiedenen Klassen die gleiche Breite. Sollte dies nicht der Fall sein, muss man die Häufigkeit jeder Klasse durch die Breite teilen, um richtig interpretieren zu können.

Dieser Beitrag ist eine überarbeitete Fassung aus dem Supplement Statistik aus dem Jahr 2001.