

Empirisches Quantil

Ein **empirisches (p-)Quantil**, auch kurz einfach nur **Quantil** genannt, ist ein Begriff aus der Statistik und eine Kennzahl einer Stichprobe. Vereinfacht teilt ein empirisches p-Quantil die Stichprobe so, dass ein Anteil der Stichprobe von p kleiner als das empirische p-Quantil ist und ein Anteil von $1-p$ der Stichprobe größer als das empirische p-Quantil ist. Ist beispielsweise eine Stichprobe von Schuhgrößen gegeben, so ist das empirische 0,35-Quantil diejenige Schuhgröße s , so dass 35 % der Schuhgrößen in der Stichprobe kleiner als s sind und 65 % größer als s sind.

Einige empirische p-Quantile tragen Eigennamen. Zu ihnen gehören der Median ($p = 0,5$), das **obere Quantil** und das **untere Quantil** sowie die **Terzile**, **Quintile**, **Dezile** und die **Perzentile**.

Von den hier besprochenen empirischen Quantilen sind die Quantile (im Sinne der Wahrscheinlichkeitstheorie) zu unterscheiden. Diese sind Kennzahlen einer Wahrscheinlichkeitsverteilung und damit einer abstrakten (Mengen-)Funktion (ähnlich dem Erwartungswert), während die empirischen Quantile Kennzahlen einer Stichprobe sind (ähnlich dem arithmetischen Mittel).

Inhaltsverzeichnis

Definition

Beispiel

Spezielle Quantile

Median

Terzile

Quartile

Quintile

Dezile

Perzentile

Abgeleitete Begriffe

Darstellung

Einzelnachweise

Definition

Es bezeichne $\lfloor x \rfloor$ die Abrundungsfunktion. Sie rundet jede Zahl x auf die nächste kleinere ganze Zahl ab. Es gilt also beispielsweise $\lfloor 1,2 \rfloor = 1$ und $\lfloor 3,99 \rfloor = 3$.

Gegeben sei eine Stichprobe (x_1, x_2, \dots, x_n) der Größe n , bei der die Elemente der Größe nach geordnet sind. Dies bedeutet, es gilt

$$x_1 \leq x_2 \leq \dots \leq x_n.$$

Dann heißt für eine Zahl $p \in (0, 1)$

$$x_p = \begin{cases} \frac{1}{2}(x_{n \cdot p} + x_{(n \cdot p)+1}), & \text{wenn } n \cdot p \text{ ganzzahlig,} \\ x_{\lfloor n \cdot p + 1 \rfloor}, & \text{wenn } n \cdot p \text{ nicht ganzzahlig.} \end{cases}$$

das empirische p-Quantil von x_1, x_2, \dots, x_n .^[1]

Es existieren einige von der hier angegebenen Definition abweichende Definitionen.^[2]

Beispiel

Die folgende Stichprobe besteht aus zehn zufälligen ganzen Zahlen (gezogen aus den Zahlen zwischen null und hundert, versehen mit der diskreten Gleichverteilung):

82; 91; 12; 92; 63; 9; 28; 55; 96; 97

Sortieren liefert die Stichprobe

$x_1 = 9; x_2 = 12; x_3 = 28; x_4 = 55; x_5 = 63; x_6 = 82; x_7 = 91; x_8 = 92; x_9 = 96; x_{10} = 97$
 .

Es ist $n = 10$.

Für $p = 0,5$ erhält man $p \cdot n = 5$. Da dies ganzzahlig ist, erhält man über die Definition

$$x_{0,5} = \frac{1}{2}(x_5 + x_{5+1}) = \frac{1}{2}(63 + 82) = 72,5$$

Für $p = 0,25$ erhält man $p \cdot n + 1 = 0,25 \cdot 10 + 1 = 2,5 + 1$. Die Abrundungsfunktion liefert dann $\lfloor 3,5 \rfloor = 3$ und damit

$$x_{0,25} = x_3 = 28.$$

Analog erhält man für $p = 0,75$ direkt $p \cdot n + 1 = 0,75 \cdot 10 + 1 = 8,5$ und damit $\lfloor 8,5 \rfloor = 8$, also ist

$$x_{0,75} = x_8 = 92.$$

Das empirische Quantil ist im Gegensatz zum arithmetischen Mittel robust gegenüber Ausreißern. Dies bedeutet, dass wenn man Werte einer Stichprobe oberhalb (oder unterhalb) eines bestimmten Quantils durch einen Wert oberhalb (oder unterhalb) des Quantils ersetzt, sich das Quantil selbst nicht verändert. Dies beruht darauf, dass Quantile nur durch ihre Ordnung und damit ihre Lage zueinander bestimmt werden und nicht durch die konkreten Zahlenwerte der Stichprobe. So wäre im Fall der obigen Stichprobe das arithmetische Mittel $\bar{x} = 62,2$. Modifiziert man nun aber den größten Wert der Stichprobe, setzt beispielsweise

$$x_{10} = 1000,$$

so ist $\bar{x} = 152,8$, wohingegen der Median sowie das untere und das obere Quartil unverändert bleiben, da sich die Reihenfolge der Stichprobe nicht verändert hat.

Spezielle Quantile

Für gewisse p-Werte tragen die zugehörigen Quantile Eigennamen. Sie sind hier im Folgenden kurz vorgestellt. Zu beachten ist, dass auch die entsprechenden Quantile von Wahrscheinlichkeitsverteilungen teils mit denselben Eigennamen bezeichnet werden.

Median

→ Hauptartikel: Median

Der Median ist das **0,5**-Quantil und teilt somit die Stichprobe in zwei Hälften: Eine Hälfte ist kleiner als der Median, die andere größer als der Median. Er ist mit dem Modus und dem arithmetischen Mittel ein wichtiger Lageparameter in der deskriptiven Statistik.

Terzile

Als Terzile werden die beiden p-Quantile für $p = \frac{1}{3}$ und $p = \frac{2}{3}$ bezeichnet. Sie teilen die Stichprobe in drei gleich große Teile: ein Teil ist kleiner als das untere Terzil ($=\frac{1}{3}$ -Quantil), ein Teil ist größer als das obere Terzil ($=\frac{2}{3}$ -Quantil), und ein Teil liegt zwischen den Terzilen.

Quartile

Als Quartile werden die beiden Quantile mit $p = 0,25$ und $p = 0,75$ bezeichnet. Dabei heißt das **0,25**-Quantil das untere Quartil und das **0,75**-Quantil das obere Quartil. Zwischen oberem und unterem Quartil liegt die Hälfte der Stichprobe, unterhalb des unteren Quartils und oberhalb des oberen Quartils jeweils ein Viertel der Stichprobe. Auf Basis der Quartile wird der Interquartilsabstand definiert, ein Streuungsmaß.

Quintile

Als Quintile werden die vier Quantile mit $p = 0,2; 0,4; 0,6; 0,8$ bezeichnet. Demnach befinden sich 20 % der Stichprobe unter dem ersten Quintil und 80 % darüber, 40 % der Stichprobe unter dem zweiten Quintil und 60 % darüber etc.

Dezile

Die Quantile für vielfache von **0,1**, also für $p = 0,1; 0,2; \dots; 0,9$ werden Dezile genannt. Dabei heißt das **0,1**-Quantil das erste Dezil, das **0,2**-Quantil das zweite Dezil etc. Unterhalb des ersten Dezils liegen 10 % der Stichprobe, oberhalb entsprechend 90 % der Stichprobe. Ebenso liegen 40 % der Stichprobe unterhalb des vierten Dezils und 60 % oberhalb.

Perzentile

Als Perzentile werden die Quantile von **0,01** bis **0,99** in Schritten von **0,01** bezeichnet.

Abgeleitete Begriffe

Aus den Quantilen lassen sich noch gewisse Streuungsmaße ableiten. Das wichtigste ist der Interquartilabstand (englisch *interquartile range*)

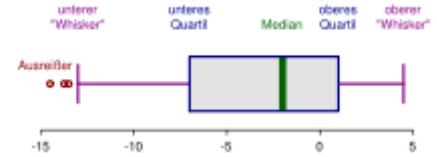
$$IQR := x_{0,75} - x_{0,25}.$$

Er gibt an, wie weit das obere und das untere Quartil auseinanderliegen und damit auch, wie breit der Bereich ist, in dem die mittleren 50 % der Stichprobe liegen.^[3] Etwas allgemeiner kann der **(Inter-)quantilabstand** definiert werden als $x_{1-p} - x_p$ für $p \in (0; 0,5)$. Er gibt an, wie breit der Bereich ist, in dem die mittleren $200 \cdot p$ % der Stichprobe liegen. Für $p = 0,25$ entspricht er dem Interquartilabstand.

Ein weiteres abgeleitetes Streumaß ist die **Median-Abweichung**, englisch *median absolute deviation*. Ist eine Stichprobe x_1, x_2, \dots, x_n mit Median $x_{0,5}$ gegeben, so ist die Median-Abweichung der empirische Median der modifizierten Stichprobe $|x_1 - x_{0,5}|, |x_2 - x_{0,5}|, \dots, |x_n - x_{0,5}|$.^[4]

Darstellung

Eine Möglichkeit, Quantile darzustellen ist der Box-Plot. Dabei wird die gesamte Stichprobe durch einen Kasten – versehen mit zwei Antennen – dargestellt. Die äußere Begrenzung des Kastens sind jeweils das obere und das untere Quartil. Somit befindet sich die Hälfte der Stichprobe im Kasten. Der Kasten selbst ist nochmals unterteilt, der unterteilende Strich ist dabei der Median der Stichprobe. Die Antennen sind nicht einheitlich definiert. Eine Möglichkeit ist, als Begrenzung der Antennen das erste und das neunte Dezil zu wählen.



Box-Plot einer Stichprobe

Einzelnachweise

1. Norbert Henze: *Stochastik für Einsteiger*. Eine Einführung in die faszinierende Welt des Zufalls. 10. Auflage. Springer Spektrum, Wiesbaden 2013, ISBN 978-3-658-03076-6, S. 30, doi:10.1007/978-3-658-03077-3 (<https://doi.org/10.1007/978-3-658-03077-3>).
2. Eric W. Weisstein: *Quantile*. (<http://mathworld.wolfram.com/Quantile.html>) In: *MathWorld* (englisch).
3. Eric W. Weisstein: *Interquartile Range*. (<http://mathworld.wolfram.com/InterquartileRange.html>) In: *MathWorld* (englisch).
4. Norbert Henze: *Stochastik für Einsteiger*. Eine Einführung in die faszinierende Welt des Zufalls. 10. Auflage. Springer Spektrum, Wiesbaden 2013, ISBN 978-3-658-03076-6, S. 32, doi:10.1007/978-3-658-03077-3 (<https://doi.org/10.1007/978-3-658-03077-3>).

Abgerufen von „https://de.wikipedia.org/w/index.php?title=Empirisches_Quantil&oldid=189697116“

Diese Seite wurde zuletzt am 19. Juni 2019 um 22:29 Uhr bearbeitet.

Der Text ist unter der Lizenz „Creative Commons Attribution/Share Alike“ verfügbar; Informationen zu den Urhebern und zum Lizenzstatus eingebundener Mediendateien (etwa Bilder oder Videos) können im Regelfall durch Anklicken dieser abgerufen werden. Möglicherweise unterliegen die Inhalte jeweils zusätzlichen Bedingungen. Durch die Nutzung dieser Website erklären Sie sich mit den Nutzungsbedingungen und der Datenschutzrichtlinie einverstanden.

Wikipedia® ist eine eingetragene Marke der Wikimedia Foundation Inc.