STATISTIK I – Übung 06
Schiefe und Wölbung

1 Kurze Wiederholung

Eine dritte Form von Verteilungsparametern?

Neben den Maßen der zentralen Tendenz (Zentrum einer Verteilung) und den Dispersionsparametern (Streuung der Werte einer Verteilung um dieses Zentrum), lassen sich Verteilungen auch – wenn dies auch weniger gebräuchlich ist – über ihre Form charakterisieren. Dies kann über die Schiefe (symmetrisch, linkssteil/rechtsschief oder rechtssteil/linksschief) sowie über die Wölbung (ähnlich der einer Normalverteilung, spitzer als die einer Normalverteilung oder flacher als die einer Normalverteilung) geschehen. Die Schiefe kann über den Momentenkoeffizienten der Schiefe sowie über den Quartilskoeffizienten der Schiefe, die Wölbung über die Kurtosis / Exzeß bestimmt werden.

Momentenkoeffizient der Schiefe

Die Formel für die Berechnung des Momentenkoeffizienten der Schiefe basiert auf der bereits bekannten Formel für die Berechnung der Varianz (d.h. potenzierte durchschnittliche Abweichung der Werte einer Verteilung von deren arithmetischem Mittel). Da die Berechnung des Momentenkoeffizienten die Berechnung des arithmetischen Mittels voraussetzt, kann dieser nur für metrische Daten ermittelt werden. Für die Berechnung des Momentenkoeffizienten g_m werden drei Formeln benötigt:

$$g_m = \frac{m_3}{s^3}$$

$$m_3 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3$$

$$s^3 = \left(\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)^3$$

Liegt der Momentenkoeffizient der Schiefe nahe Null, handelt es sich um eine symmetrische Verteilung. Positive Werte deuten dagegen auf eine linkssteile, negative auf eine rechtssteile Verteilung hin. Wer die Formel für s^3 näher betrachtet, wird feststellen, dass es sich um die Formel für die Varianz handelt, aus der die Wurzel gezogen (ergibt die Standardabweichung) und die anschließend mit 3 potenziert wird. Bei s^3 handelt es sich also in der Tat um die dritte Potenz der Standardabweichung s.

Quartilskoeffizient der Schiefe

Liegen lediglich ordinalskalierte Daten vor, kann der Momentenkoeffizient der Schiefe nicht berechnet werden. Möglich ist allerdings die Berechnung des Quartilskoeffizienten der Schiefe, der anstelle des arithmetischen Mittels auf die ebenfalls bereits bekannten Quartile aufbaut.

$$g_{0,25} = \frac{(x_{0,75} - x_{med}) - (x_{med} - x_{0,25})}{x_{0,75} - x_{0,25}}$$

Die Interpretation des Quartilskoeffizienten erfolgt analog zu der des Momentenkoeffizienten: Liegt der Wert nahe Null, handelt es sich um eine symmetrische Verteilung, während positive Werte auf eine linkssteile und negative Werte auf eine rechtssteile Verteilung hindeuten. Zu beachten ist außerdem: Beträgt der Interquartilsabstand ($x_{0,75} - x_{0,25}$) Null, kann der Quartilskoeffizient der Schiefe nicht berechnet werden, da eine Division durch Null nicht möglich ist.

Kurtosis / Exzeß

Neben der Linkssteilheit/Rechtssteilheit von Verteilungen ist die Wölbung einer Verteilung ein weiteres interessantes Kriterium. Mit Hilfe der Kurtosis (auch als Exzeß bezeichnet), kann festgestellt werden, inwieweit die Wölbung einer Verteilung der Wölbung der bekannten Normalverteilung gleicht. Da die Formel voraussetzt, dass das arithmetische Mittel berechnet werden kann, lässt sich die Kurtosis – wie bereits der Momentenkoeffizient der Schiefe – nur dann berechnen, wenn metrisch skalierte Daten vorliegen.

$$g_k = \frac{m_4}{s^4} - 3$$

$$m_4 = \frac{1}{n}\sum_{j=1}^{n}(x_j - \bar{x})^4$$

$$s^4 = \left(\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)^4$$

Ergibt sich ein Wert nahe Null, entspricht die Wölbung der Verteilung der Wölbung einer Normalverteilung. Bei einem positiven Wort ist von einer „spitzeren" Form der Verteilung, bei einem negativen Wert dagegen von einer „flacheren" Form der Verteilung auszugehen. Die Subtraktion von 3 in der Hauptformel ist übrigens darauf zurückzuführen, dass die Normalverteilung eine Kurtosis von 3 aufweist – durch das Abziehen von 3 vom Ergebnis, ergibt sich bei völliger Gleichheit mit der Normalverteilung also ein Wert von Null und somit die Möglichkeit, das Ergebnis analog zum Momentenkoeffizienten der Schiefe zu interpretieren. Ebenso wie beim Momentenkoeffizienten der Schiefe ist die Interpretation der Kurtosis nur dann sinnvoll, wenn eine unimodale Verteilung vorliegt – und ebenso wie beim Momentenkoeffizienten findet sich auch hier in der Formel für s^4 die Varianz bzw. die Standardabweichung wieder, die hier anstelle mit 3 mit 4 potenziert wird. Wurden Varianz und Standardabweichung für die vorliegenden Daten bereits berechnet, lässt sich die Berechnung des Momentenkoeffizienten sowie der Kurtosis durch Rückgriff auf die Standardabweichung also abkürzen.

2 Beispielrechnungen

An einer Fertigungsanlage werden 20 Polymerbauteile als Zufallsstichprobe aus der laufenden Produktion entnommen und gewogen. Die (absoluten) Abweichungen von einem produktionstechnisch bedingten Idealgewicht in Gramm werden in einer Tabelle festgehalten.

Bauteil	Abweichung	Bauteil	Abweichung
1	0	11	1
2	6	12	2
3	3	13	0
4	2	14	3
5	2	15	5
6	4	16	7
7	3	17	2
8	1	18	2
9	1	19	1
10	0	20	3

Berechnung des Momentenkoeffizienten

Ein Blick auf die Formeln verrät, das eine Hilfstabelle zu Berechnung dreier Termini erforderlich ist.

$$g_m = \frac{m_3}{s^3}$$

$$m_3 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3$$

$$s^3 = \left(\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)^3$$

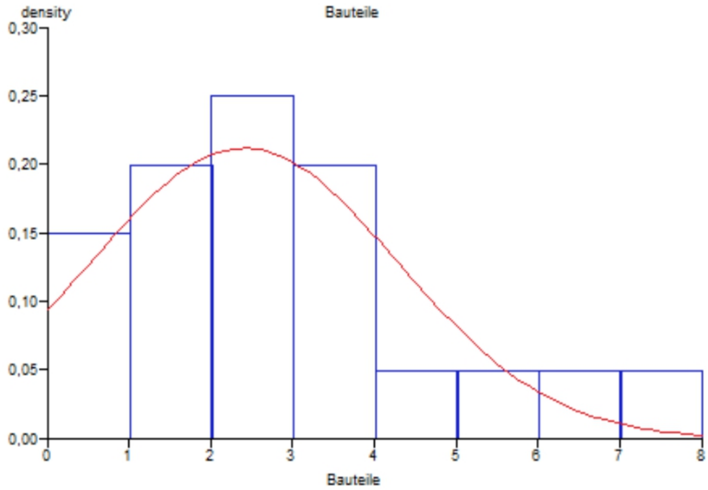
Bauteil	x_i	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^3$
1	0	5,76	-13,82
2	6	12,96	46,66
3	3	0,36	0,22
4	2	0,16	-0,06
5	2	0,16	-0,06
6	4	2,56	4,10
7	3	0,36	0,22
8	1	1,96	-2,74
9	1	1,96	-2,74
10	0	5,76	-13,82
11	1	1,96	-2,74
12	2	0,16	-0,06
13	0	5,76	-13,82
14	3	0,36	0,22
15	5	6,76	17,58
16	7	21,16	97,34
17	2	0,16	-0,06
18	2	0,16	-0,06
19	1	1,96	-2,74
20	3	0,36	0,22
∑	48	70,80	113,76
Ø	2,4	3,54	5,69

$$g_m = \frac{m_3}{s^3} = \frac{5{,}69}{6{,}66} = 0{,}85$$

$$m_3 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3 = 5{,}69$$

$$s^3 = \left(\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)^3 = \sqrt{3{,}54}^3 = 6{,}66$$

Der Momentenkoeffizient der Schiefe liegt bei 0,85 – die Verteilung ist somit leicht linkssteil. Ein Blick auf die Verteilungskurve (erstellt mit Smith's Statistical Package) bestätigt diese Interpretation.



Berechnung des Quartilskoeffizienten

Der Quartilskoeffizient der Schiefe berechnet sich aus den drei Quartilswerten.

Die geordnete Verteilung: 0; 0; 0; 1; 1; 1; 1; 2; 2; 2; 2; 2; 3; 3; 3; 3; 4; 5; 6; 7

Ergibt (n * p) einen ganzzahligen Wert (k), berechnet sich das Perzentil wie folgt:

$$x_p = \frac{1}{2}(x_{(k)} + x_{(k+1)})$$

(n*p) = (20 * 0,25) = 5 -> k = 5; k+1 = 6 -> xp = (1+1)/2 = 1
(n*p) = (20 * 0,50) = 10 -> k = 10; k+1 = 11 -> xp = (2+2)/2 = 2
(n*p) = (20 * 0,75) = 15 -> k = 15; k+1 = 16 -> xp = (3+3)/2 = 3

$$g_{0,25} = \frac{(x_{0,75} - x_{med}) - (x_{med} - x_{0,25})}{x_{0,75} - x_{0,25}} = \frac{(3-2)-(2-1)}{(3-1)} = \frac{0}{2} = 0$$

Der Quartilskoeffizient der Schiefe beträgt 0. Dies legt eine symmetrische Verteilung nahe und scheint zunächst im Widerspruch zum Momentenkoeffizient der Schiefe zu stehen. Tatsächlich lässt die Kurve aber beide Interpretationen zu – und das Beispiel verdeutlicht, wie problematisch bisweilen die Subjektivität der Ergebnisinterpretation (ohne definierte Grenzwerte) sein kann.

Berechnung der Kurtosis

Auch für die Berechnung der Kurtosis empfiehlt sich die Anlage einer Hilfstabelle.

Bauteil	x_i	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^4$
1	0	5,76	33,18
2	6	12,96	167,96
3	3	0,36	0,13
4	2	0,16	0,03
5	2	0,16	0,03
6	4	2,56	6,55
7	3	0,36	0,13
8	1	1,96	3,84
9	1	1,96	3,84
10	0	5,76	33,18
11	1	1,96	3,84

12	2	0,16	0,03
13	0	5,76	33,18
14	3	0,36	0,13
15	5	6,76	45,70
16	7	21,16	447,75
17	2	0,16	0,03
18	2	0,16	0,03
19	1	1,96	3,84
20	3	0,36	0,13
∑	48	70,80	783,50
Ø	2,4	3,54	39,18

$$g_k = \frac{m_4}{s^4} - 3 = \frac{39{,}18}{12{,}53} - 3 = 0{,}13$$

$$m_4 = \frac{1}{n}\sum_{j=1}^{n}(x_j - \bar{x})^4 = 39{,}18$$

$$s^4 = \left(\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)^4 = \sqrt{3{,}54}^4 = 12{,}53$$

Die Kurtosis liegt mit 0,13 nahe der Null – damit ist die Wölbung ähnlich der einer Normalverteilung. Auch diese Annahme wird durch die Betrachtung der mit SSP erstellten Kurve bestätigt.

3 Übungsaufgaben (Lösungen folgen in der kommenden Woche)

Der gleichen Fertigungsanlage werden erneut 20 Polymerbauteile als Zufallsstichprobe aus der laufenden Produktion entnommen und gewogen. Die diesmal deutlich größer ausfallenden (absoluten) Abweichungen von einem Idealgewicht in Gramm werden in einer Tabelle festgehalten.

Bauteil	Abweichung	Bauteil	Abweichung
1	12	11	20
2	27	12	17
3	13	13	19
4	10	14	21
5	7	15	30
6	5	16	44
7	11	17	56
8	23	18	8
9	12	19	7
10	11	20	3

a) Berechnen Sie den Momentenkoeffizienten der Schiefe.

b) Berechnen Sie den Quartilskoeffizienten der Schiefe.

c) Berechnen Sie die Kurtosis.