C.H.BECK
www.chbeck.de

Unverkäufliche Leseprobe

Christian Hesse
Wer falsch rechnet, den bestraft das Leben
Das kleine Einmaleins der Alltagsmathematik

217 Seiten mit 73 Abbildungen. Gebunden
ISBN: 978-3-406-64472-6

Weitere Informationen finden Sie hier:
http://www.chbeck.de/11253545

Nachdem das alles herausgearbeitet worden ist, wollen wir auf den Irrtum von der globalen Annäherung an einen Mittelwert zu sprechen kommen. Die Frage wurde schon aufgeworfen und war offengeblieben. Bedeutet Regression zur Mitte nicht zwingend den langfristigen Ausgleich aller Unterschiede? Werden sich im Beispiel der Körpergröße nicht alle Größen schließlich beim Mittelwert einpendeln? Warum sind wir nicht alle gleich groß? Warum haben wir nicht alle braune Haare? Warum heißen wir nicht alle Schmidt?

Das nach dem zuvor Gesagten vermeintlich Naheliegende widerspricht natürlich eindeutig unserer Anschauung. Variation war und ist bei Körpergrößen offensichtlich vorhanden. Um hier nicht diesseits der Banalitätsgrenze zu bleiben, wollen wir die Streuung genauer untersuchen. Anhand unseres Mütter-Töchter-Beispiels stellt sich das Rätsel der Streuung angesichts beständiger Regression zur Mitte folgendermaßen dar:

Große Mütter haben im Mittel weniger große Töchter. Kleine Mütter haben im Mittel weniger kleine Töchter. Beide Befunde würden, für sich genommen, tatsächlich dazu führen, dass die Streuung in der Töchtergeneration geringer ausfällt als in der Müttergeneration.

In der nächsten Generation würde eine weitere Verringerung der Streuung auftreten, wenn dann die einstigen Töchter selbst auch Mütter werden und ihrerseits Töchter haben sollten, und abermals bei den Enkelinnen, den Urenkelinnen usw. Auch dieser Gedankengang weist in Richtung abnehmender Variation. Wir müssen etwas übersehen haben.

Und in der Tat: Es gibt einen bislang noch nicht angesprochenen Vorgang, der dem Regressionseffekt entgegenwirkt. Es ist gewissermaßen ein Antiregressionseffekt, der, für sich genommen, in Richtung einer Vergrößerung der Streuung arbeitet, während der Regressionseffekt, isoliert betrachtet, die Streuung verringert.

Es ist ein Effekt, der von den Müttern mittlerer Größe ausgeht. Wir hatten ihnen bisher wenig Aufmerksamkeit geschenkt. Diese Mütter haben Töchter, die aufgrund ihrer genetischen Disposition und aufgrund weiterer Wirkfaktoren im Schnitt vom Mittelwert stärker nach unten und nach oben abweichen, als es die Mütter selbst tun. Das bringt zusätzliche Variation. Dieses Phänomen der Streuungsverstärkung hält dem Verlust an Streuung, den der Regressionseffekt von einer Generation zur nächsten bewirken würde, recht genau die Waage.

Das Symmetrieparadoxon

Wir wissen aufgrund der Daten, dass große Mütter im Schnitt weniger große Töchter haben und kleine Mütter weniger kleine Töchter. Das ist gesichertes Wissen.

Nichts hindert uns daran, die Perspektive nun zu ändern und zu fragen: Welcher Zusammenhang stellt sich ein, wenn wir, von den Töchtern ausgehend, unser Augenmerk auf deren Mütter richten, also gegengedanklich die Größenbeziehungen andersherum untersuchen? Die Richtung der Regression kehrt sich dann genau um.

Wir kehren zwar die Blickrichtung um, aber die Argumentationsweise kann unverändert bleiben. Konzentrieren wir uns zuerst auf sehr große Töchter, dann ist zu erwarten, dass ihre Mütter im Schnitt weniger groß sind. Und betrachtet man ausschließlich sehr kleine Töchter, dann ist zu erwarten, dass die dazugehörigen Mütter weniger klein sind.

Gehen wir mit den Zahlen nun ähnlich um wie beim ersten Anlauf für die andere Richtung: Heben wir jene Töchter hervor, deren Größe nach oben um mindestens eine Standardabweichung vom Töchter-Mittelwert abweicht. Das sind 36 der 278 Töchter. Im Mittel liegt die Größe ihrer Mütter bei 0,82 Standardabweichungen oberhalb des Mütter-Durchschnitts.

Ganz so wie intuitiv erwartet, sind die Mütter großer Töchter ebenfalls groß. Aber im Mittel sind sie weniger groß, in Standardabweichungen gemessen, als ihre Töchter.

Der Regressionseffekt ist also vollkommen symmetrisch. Sowohl bei der Regression von *X* in Richtung *Y* als auch bei der Regression von *Y* in Richtung *X* tritt der Regressionseffekt gleichermaßen auf.

Beides kann nicht sein, oder?

Das es offenbar doch so ist, nennt man Symmetrieparadoxon der Regression.

Knobelzone

Unsymmetrie

Tom und Jerry sind beide Turner. Tom ist klein, und Jerry ist groß. Jerry überragt Tom um fast einen Viertelmeter. Eines Tages möchten die beiden in der Küche ganz oben aus einem hohen Regal etwas herausnehmen. Dazu stellt sich einer der beiden auf den anderen. Sollte Tom auf Jerrys Schultern stehen oder Jerry auf Toms Schultern, damit die beiden möglichst hoch reichen können?

Lösung
Jerry hat wegen seiner Größe auch die längeren Arme. Wer auf den Schultern des anderen steht, sollte die längeren Arme haben, um möglichst hoch zu reichen. Also muss Jerry nach oben. Stände er unten, würden seine langen Arme gar keinen Vorteil bringen.

Sturz der Parität

Linkshänder sind bessere Rechtshänder, als Rechtshänder Linkshänder sind.

Damit wäre das Paradoxon beschrieben. Aber wie soll ich es Ihnen erklären?

Eine Möglichkeit ist die folgende: Der vorgenommene Perspektivenwechsel bei der Regression läuft im Streudiagramm darauf hinaus, statt Punkten – also Mutter-Tochter-Paaren – in einem *vertikalen* Streifen nunmehr Punkte in einem *horizontalen* Streifen zu betrachten. Das aber sind unterschiedliche Teilmengen der Gesamtmenge.

Beim Symmetrieparadoxon, den scheinbar unvereinbaren Aussagen zu Richtung und Gegenrichtung des Regressionseffekts, wird also das Verhalten von *verschiedenen* Teilmengen im Streudiagramm bewertet. Verschiedene Teilmengen können sich aber unterschiedlich verhalten, so wie wir es auch gesehen haben. In die Paradoxiefalle gerät man mit der Annahme, dass sich die beiden konträren Aussagen auf das Verhalten ein und derselben Teilmenge der Punktwolke beziehen. Das ist aber nicht der Fall. Bedenkt man dies, löst sich die Paradoxie in Luft auf.

Damit wär' auch das geklärt.

Regression ist sehr facettenreich. Die Regressionsmethode einfach nur nützlich zu nennen, wäre ungenau. Sie ist unschlagbar, wenn es darum geht, Art und Stärke der Beziehungen zwischen den Variablen dieser Welt zahlenmäßig zu erfassen. Wir zeigen weiteres Anschauungsmaterial, von heiter bis ernst. Das erste ist ein spielerisches Schulbeispiel für ihren gelungenen Einsatz in der Biologie.

Vom Grillen-Thermometer

Im Sommer hört man auf vielen Wiesen Grillen zirpen. Grilllenmännchen erzeugen dieses bis zu 50 Meter entfernt hörbare Geräusch durch Aneinanderreiben ihrer Flügel. Grillen sind Kaltblüter. Ihr gesamter Stoffwechsel und ihr Aktivitätsniveau passen sich der Umgebungstemperatur an. Die Frequenz, mit der die Grillen zirpen, macht davon keine Ausnahme.

In Nordamerika ist die *Gestreifte Bodengrille* verbreitet. Der Biologe Pieru hat diese Grillenart 1949 ausführlich studiert. Unserem

Spieltrieb folgend, interessiert uns vor allem ein kleiner Datensatz: Bei 15 verschiedenen Außentemperaturen hat Pieru die Zirpfrequenz, also die Tonhöhe, von Bodengrillen gemessen. Die Zahlen sind in der folgenden Tabelle festgehalten:

Frequenz (in Hz)	**Außentemperatur (in Grad Celsius)**
20,0	31,4
16,0	22,0
19,8	34,1
18,4	29,1
17,1	27,0
15,5	24,0
14,7	20,9
17,1	27,8
15,4	20,8
16,2	28,5
15,0	26,4
17,2	28,1
16,0	27,0

In dieser Darreichungsform ist ein Zusammenhang nur schwer erkennbar. Besser eignet sich dafür eine optische Aufbereitung. Das macht Abbildung 60. Sie stellt die Daten wieder als Punkte in einem Achsensystem dar, mit der Zirpfrequenz auf der Rechtsachse und dem zugehörigen Temperaturwert auf der Hochachse. Zudem zeigt sie uns die Regressionsgerade. Ihre Gleichung lautet:

$$\textit{Celsius-Temperatur} = -3{,}3 + 1{,}8 \cdot \textit{Zirpfrequenz}$$

Die Streuung der Datenpunkte um die Regressionsgerade ist nicht allzu groß. Die Gerade liefert eine recht gute Annäherung an die Datenpunkte, was, bei allen Vorbehalten aufgrund der geringen Zahl der Messwerte, auf einen engen Zusammenhang zwischen Temperatur und Frequenz hindeutet.

Abbildung 60: Regression von Temperatur versus Zirpfrequenz für eine Grillenart

Grillen sind offenbar genaue Temperaturfühler. Was liegt also näher, als sich eine possierliche Grille als Taschenthermometer zuzulegen?

Haben Sie ein gutes Ohr und ist Ihre Grille aktiv, können Sie mit der Gleichung der Regressionsgeraden aus der Tonhöhe auf die Temperatur schließen. Beträgt die Zirpfrequenz zum Beispiel 19 Hertz, speisen Sie diesen Wert in die Regressionsgleichung ein, und schon kann Ihnen mitgeteilt werden, dass die Außentemperatur momentan

$$-3{,}3 + 1{,}8 \cdot 19 = 30{,}9 \textit{ Grad Celsius}$$

beträgt.

Haben Sie kein gutes Gehör für Frequenzen in diesem Bereich, was Ihnen niemand verübeln sollte, so kann Ihnen auf andere Weise geholfen werden. Schon im Jahr 1897 hat der Biologe und Physiker Amos Dolbear nämlich festgestellt, dass auch die *Anzahl* der Zirpgeräusche einer Grille temperaturabhängig ist. Die Häufigkeit des Zirpens nimmt mit zunehmender Temperatur linear zu. Wird es wärmer, zirpt die Grille öfter.

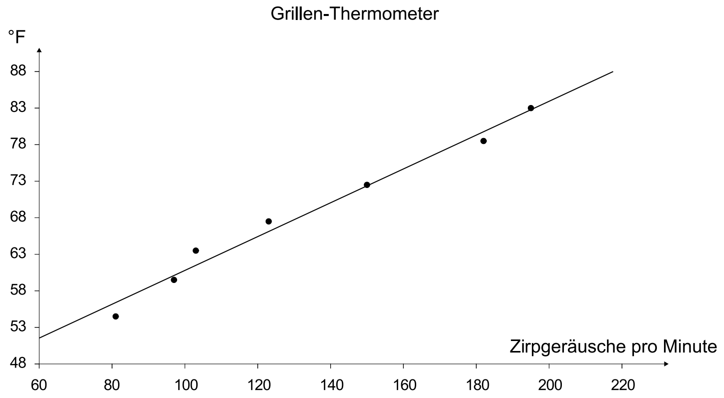

Abbildung 61: Zirprate und Temperatur für eine Grillenart

So lässt sich aus der Anzahl der Zirpgeräusche pro Minute ebenfalls auf die Außentemperatur zurückschließen. Die Zirprate streut sogar noch weniger als die Zirpfrequenz und hängt zudem so gut wie gar nicht vom Individuum oder dessen Alter ab.

Zirpgeräusche[12] pro Minute	81	97	103	123	150	182	195
Temperatur in Fahrenheit	54,5	59,5	63,5	67,5	72,0	78,5	83,0

Erst im Achsensystem wird die sehr geringe Streuung der Messwerte um die Regressionsgerade deutlich. Optik schlägt Tabellistik.

Die Regressionsgerade gehorcht der Gleichung:

$$\textit{Fahrenheit-Temperatur} = 37{,}7 + 0{,}23 \cdot \textit{Zirpgeräusche pro Minute}$$

Das ist übrigens eine sehr gute mathematische Bestätigung einer alten amerikanischen Bauernregel: Sie schließt von der Anzahl der minütlichen Zirpgeräusche auf die Temperatur, indem die Anzahl durch 4 geteilt wird und diesem Wert 40 hinzugezählt werden.

Bedenkt man ferner, dass Fahrenheit- und Celsius-Temperaturen in einem einfachen Zusammenhang stehen, so lässt sich die amerikanische Bauernregel leicht für europäische Bauern umrechnen. Da die Celsius-Temperatur sich ergibt, indem man von der Fahrenheit-Temperatur zunächst 32 abzieht, dann diesen Wert mit 5 multipliziert und anschließend durch 9 dividiert, so geht die obige Regressionsgleichung über in die entsprechende Formel:

$$\textit{Celsius-Temperatur} = 3{,}2 + 0{,}127 \cdot \textit{Zirpgeräusche pro Minute}$$

Diese Formel ist ausgesprochen genau. Und als Celsius-Variante der Bauernregel bietet sich an:

Zirpgeräusche pro Minute geteilt durch 8 plus 3 ergibt Celsius-Temperatur!

Interessanterweise sind benachbarte Grillen sehr stark synchronisiert. Sie zirpen im Einklang. Amos Dolbear hat gemeint,[13] Grillen in einem Feld seien synchronisiert bis hin zu dem Punkt, dass sie so genau dieselben Zeitintervalle von Geräusch und Stille einhalten, als folgten sie dem Stab eines Dirigenten.

Geimpft

Unser nächstes Beispiel verlässt die Arena des Spielerischen. Vielmehr handelt es sich um ein ernstes Thema: Es geht um die Kindersterblichkeit in ausgewählten Ländern und den möglichen Zusammenhang mit dem Impfschutz für Säuglinge. Die Datenanalyse führt zu einem wichtigen Ergebnis.

Die mit «Sterblichkeit von 1000» überschriebene Spalte in unten stehender Tabelle[14] enthält die Zahl der bis zum Alter von 5 Jahren gestorbenen Kinder unter je 1000 Lebendgeborenen. Die Spalte «Anzahl geimpft von 100» gibt die Anzahl je 100 Säuglinge an, die im ersten Lebensjahr eine Schutzimpfung gegen Diphtherie, Keuchhusten und Tuberkulose erhalten haben (DKT-Schnellschutzimpfung).

Land	**Anzahl geimpft von 100**	**Sterblichkeit von 1000**
Ägypten	89	55
Äthiopien	13	208
Bolivien	77	118
Brasilien	69	65
China	94	43
Deutschland	91	4
Finnland	95	7
Frankreich	95	9
Griechenland	54	9
Großbritannien	90	9
Indien	89	124
Italien	95	10
Japan	87	6
Kambodscha	32	185
Kanada	85	8
Mexiko	91	33
Polen	98	16
Russland	73	32
Senegal	47	145
Tschechische Republik	99	12
Türkei	76	87

Das zugehörige Streudiagramm mit Regressionslinie sieht so aus:

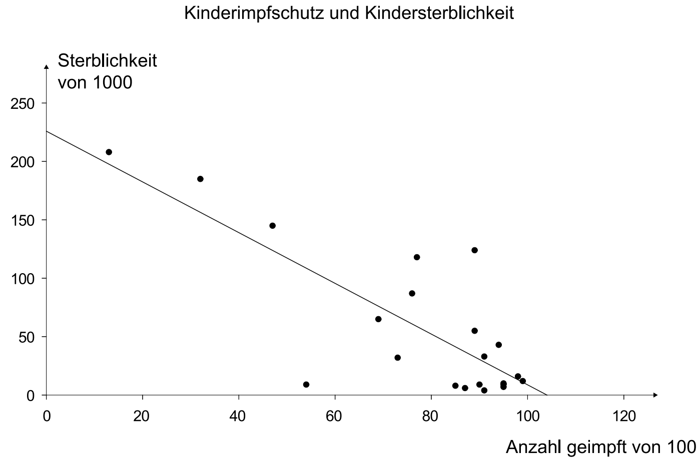

Abbildung 62: Kinderimpfschutz und Kindersterblichkeit in 21 Ländern

Die grafische Darstellung und die Regressionsgerade geben einen Eindruck von der Wirkung der DKT-Schnellschutzimpfung für Säuglinge. Das Diagramm zeigt über jeden Zweifel: Impfungen sind wichtig, und Impfen rettet Leben.

Zahlreiche wissenschaftlich-seriöse Studien haben belegt, dass sinnvolle und zeitgerechte Impfungen eine wirkungsvolle Vorsorgemaßnahme gegen Kindersterblichkeit sind.

[…]
