methevalreport
4(2) 2002

Christof Nachtigall & Ute Suhl

Der Regressionseffekt
Mythos und Wirklichkeit

metheval
psychologische methodenlehre
und evaluationsforschung jena

Impressum

methevalreport
erscheint seit 1999
in unregelmäßigen Abständen
als „graue“ Schriftenreihe des Lehrstuhls für
Psychologische Methodenlehre und Evaluationsforschung
am Institut für Psychologie der Friedrich-Schiller-Universität Jena

Herausgeber:
Prof. Dr. Rolf Steyer
Sekr.: +49 (3641) 945 230
Durchwahl: +49 (3641) 945 231
Fax: +49 (3641) 945 232

rolf.steyer@uni-jena.de

Redaktion:
Dipl.Psych. Friedrich Funke
sff@uni-jena.de

Typographie:
cand.psych. Silke Zachariae
zachariae@web.de

Standort:
Thüringer Universitäts- und Landesbibliothek
Lesesaal Zweigstelle Psychologie

Internet
http://www.uni-jena.de/svw/metheval/report/

Bestellungen:
Methodenlehre und Evaluationsforschung
Institut für Psychologie
Steiger 3 Haus 1
D-07743 Jena
Deutschland

Copyright:
Bei unveröffentlichten Arbeiten verbleibt das Urheberrecht bei der Autorin oder beim Autor.
Das Copyright für Texte, die in anderen Publikationsorganen erschienen sind, liegt bei diesen Organen.

Der Regressionseffekt

Mythos und Wirklichkeit

Christof Nachtigall und Ute Suhl

Der 'Regressionseffekt', auch 'Regression zur Mitte' genannt, ist zunächst eine schwer verständliche Sache. Man kann diesen Effekt in verkürzter Form so beschreiben: Messwerte von Personen tendieren dazu, bei einer weiteren Messung näher am Mittelwert zu liegen. Wie kann das geschehen? Sind hier statistische Geister im Spiel?

Es ist einleuchtend, dass ein solcher Effekt große Bedeutung für die Interpretation von Effekten hat Wenn sich z. B. in klinischen Wirksamkeitsstudien die Mittelwerte von untersuchten Personen zwischen zwei Messzeitpunkten verändern, möchte man diese Effekte auf die Behandlung zurückführen und nicht auf unbegreifliche statistische Seltsamkeiten. Im Rahmen der Entwicklungspsychologie wird der Regressionseffekt als allgegenwärtig beschrieben (Furby, 1973, S. 172), und Lord sieht darin einen Hauptgrund dafür, dass "studies of growth may become confusing or confused" (Lord, 1963, S. 24).

Viel wird über den Regressionseffekt geschrieben, doch wird er nur selten explizit auf den Punkt gebracht. Viele Darstellungen des Regressionseffektes sind auch in einschlägigen Lehrbüchern verkürzt und liefern kein umfassendes Verständnis des Phänomens. Manchmal werden Regressionseffekte als statistische Artefakte beschrieben, die durch mangelnde Reliabilität des Messinstrumentes verursacht werden (Bortz & Döring, 1995, S. 517). In anderen Lehrbüchern erscheint der Regressionseffekt primär als ein Effekt der Verwendung von Regressionsgeraden (z. B. Mortensen, 1995, S.109). Die Sache ist mithin für den Nicht-Experten schwer zu durchschauen. Ziel dieses Beitrags ist es, eine verständliche und klärende Einführung zu diesem Thema zu geben.

1. Was ist Regression zur Mitte?

Der Ausdruck 'Regression zur Mitte' geht auf Galton[1] zurück, der festgestellt hatte, dass die durchschnittliche Abweichung der Nachkommen vom Mittelwert geringer ist als die durchschnittliche Abweichung der Eltern vom Mittelwert (siehe Gigerenzer et al., 1989, S.165). Galton bezog diese Aussage z. B. auf die menschliche Körpergröße. Diese Aussage ist zunächst leicht misszuverstehen. Bedeutet Regression zur Mitte, dass alle Menschen in ferner Zukunft eine einheitliche (nämliche mittlere) Körpergröße haben werden? Wenn Kinder von Generation zu Generation immer näher am Mittelwert liegen, dann würde genau dies geschehen. Der Blick auf entsprechende Daten zeigt ein anderes Bild: Bei einer Untersuchung der Körpergröße von insgesamt 488 weiblichen Psychologiestudierenden der Universität Jena erweist sich das Merkmal 'Körpergröße' als annähernd normalverteilt mit einer Streuung von s=5.85 cm (vgl. Abbildung 1)[2]. Bei der ebenfalls erfragten ‚Größe der Mütter' der Befragten liegt die Streuung bei s=5.72 cm. Offensichtlich ist kein Trend zur Einheitsgröße erkennbar. Zwar nimmt die durchschnittliche Größe innerhalb einer Generation von 165cm auf 169 cm um 4 cm zu, die Streuung bleibt jedoch so gut wie konstant.

---

[1] Sir Francis Galton (1822-1911), ein Vetter Charles Darwins, beschäftigte sich vornehmlich mit Fragen der Vererbungslehre, wobei er statistische Konzepte wie Korrelation, Regression und Normalverteilung verwendete, um die Unterschiede und Ähnlichkeiten zwischen Eltern und Nachkommen einer Spezies zu beschreiben.

[2] Die Daten können mit der Datei *http://www.uni-jena.de/svw/metheval/materialien/reports/daten/koerpergroesse.sav zum* Zwecke eigener Analysen und zum Nachvollziehen der hier beschriebenen Auswertungen heruntergeladen werden.

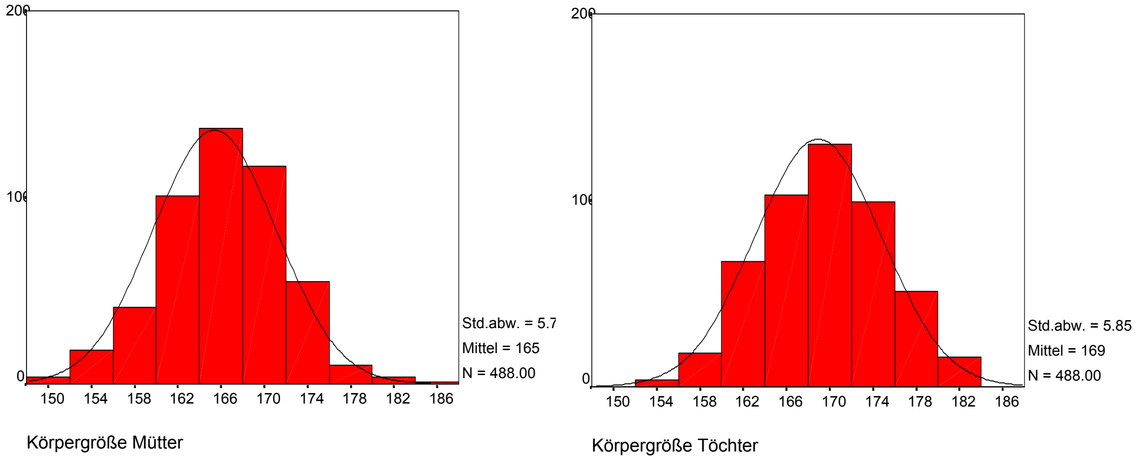

Abbildung 1: Verteilung der Körpergröße bei Müttern und Töchtern. Die durchgezogene Linie entspricht einer Normalverteilung.

Dennoch findet sich hier Regression zur Mitte. Untersucht man nämlich die Körpergröße von großen Müttern und deren Töchtern, so ergibt sich folgendes: Wir betrachten diejenigen Mütter als ‚groß', deren Körpergröße mindestens eine Standardabweichung über dem Mittelwert liegt. Um Werte leichter vergleichen zu können, verwenden wir z-standardisierte Variablen (*X=Körpergröße der Mutter, Y=Körpergröße der Tochter*). Zur Erinnerung: z-transformierte Werte geben die Abweichungen vom Mittelwert, gemessen in Standardabweichungen, wieder. Ein *X*-Wert von genau 1 bedeutet also eine Mutter mit einer Körpergröße von $165\,\text{cm} + 1 \cdot 5.72\,\text{cm} = 170.72$ cm. Im Datensatz finden sich n=70 Mütter, die mindestens eine Standardabweichung über dem Durchschnitt liegen. Der Mittelwert in dieser Gruppe der großen Mütter beträgt 1.6.

Betrachten wir nun die Töchter dieser großen Mütter. Tabelle 1 zeigt Kennwerte der Körpergröße für diese Teilstichprobe. Obwohl *alle* Mütter mehr als eine Standardabweichung über dem Durchschnitt liegen, sind deren Töchter im Schnitt nur 0.82 Standardabweichungen über dem Durchschnitt. Hier zeigt sich der Regressionseffekt: Die Töchter großer Mütter sind im Durchschnitt weniger groß. Zwar liegt der Wert immer noch über dem Gesamtdurchschnitt (z-Wert 0), doch ist dieser Wert weniger extrem als die Ausgangswerte. Entsprechende Ergebnisse zeigen die Analysen der Daten kleiner Mütter und deren Töchter: Die Töchter von kleinen Müttern sind im Durchschnitt weniger klein.

Tabelle 1: Kennwerte des z-transformierten Merkmals Körpergröße der Töchter. Die Daten stammen aus der Teilstichprobe der ‚großen' Mütter (d.h. $z_{\text{Größe Mutter}}>1$, N=70).

	N	*Minimum*	*Maximum*	***Mittelwert***	*Standardabweichung*
z-Wert(Körpergröße Töchter)	*70*	*-1.52*	*2.24*	***.82***	*.84*

Dieses Phänomene ist keineswegs auf das Beispiel Körpergröße von Eltern und Kindern beschränkt. Betrachten wir aus derselben Untersuchung die Merkmale Körpergröße und Gewicht der befragten Studentinnen. Auch hier findet sich Regression zur Mitte: In der Stichprobe gibt es 95 ‚große' Frauen (z-Wert > 1). Wie Tabelle 2 zeigt, liegt das durchschnittliche Gewicht dieser Frauen ebenfalls näher am Mittelwert. Der z-Wert beträgt nur 0.85.

Tabelle 2 Kennwerte des Merkmals ‚Gewicht' (in kg) und der z-transformierten Gewichtswerte. Die Daten stammen aus der Teilstichprobe der ‚großen' Frauen (d.h. $z_{Größe}>1$, N=90).

	N	*Minimum*	*Maximum*	***Mittelwert***	*Standardabweichung*
Gewicht (in kg)	*90*	*51*	*87*	***65.10***	*7.92*
Z-Wert (Gewicht)	*90*	*-.96*	*3.66*	***.85***	*1.02*

Dieser Effekt der Regression zur Mitte läßt sich an beliebigen Merkmalen beobachten. Notwendig sind verbundene Messungen, d.h. an derselben Person werden zwei Merkmale oder wiederholt das gleiche Merkmal erhoben. Eine andere Möglichkeit besteht, wenn - wie im Mütter-Töchter-Beispiel - die Personen paarweise einander zugeordnet werden können.

2. Definition des Regressionseffektes

Formal bedeutet Regression zur Mitte, dass bei paarweise verbundenen Messungen von X und Y für eine gegebene Abeichung eines Wertes x vom Mittelwert $\bar{x}$ kleinere Abweichungen der zugehörigen y-Werte von deren Mittelwert $\bar{y}$ zu erwarten sind. Der Ausdrucke ‚kleiner' bezieht sich dabei auf die z-Metrik, d. h. wir bewerten die Abweichungen vom Mittelwert wie in den obigen Beispielen geschehen in Standardabweichungen des jeweiligen Merkmals.

Im Rahmen von Zufallsexperimenten betrachten wir die Erwartungswerte $E(X)$ und $E(Y)$ statt der Stichprobenmittelwerte $\bar{x}$ und $\bar{y}$. Damit kann der Regressionseffekt formal definiert werden: Für alle $x \neq E(X)$ gilt

$$\frac{|E(Y|X=x)-E(Y)|}{Std(Y)} < \frac{|x-E(X)|}{Std(X)} \tag{1}$$

Der Ausdruck $E(Y|X=x)$ beschreibt den theoretischen Mittelwert der Variable Y in der Subpopulation $X=x$, z.B. die durchschnittliche Größe von Töchtern deren Mütter die Größe x haben. Betrachten wir z-standardisierte Variablen, dann vereinfacht sich Gleichung (1) zu

$$|E(Y|X=x)|<|x| \text{ für } x \neq 0. \tag{2}$$

Soll der Regressionseffektes in einer konkreten Situation nachgewiesen werden, ist die Gültigkeit von Gleichung (1) bzw. (2) zu zeigen.

3. Erklärung des Regressionseffektes

Wie kommt der Regressionseffekt zustande? Betrachten wir zwei Merkmale X und Y, die sich nicht vollständig gegenseitig determinieren. Die Ausprägung von Y bei einer Person wird zum Teil durch die Ausprägung von X sowie durch weitere Faktoren bestimmt. Machen wir uns das am Beispiel der Körpergröße klar: Die Körpergröße der Mütter wird zum einen durch die gemeinsame genetische Komponente von Müttern und Töchtern bestimmt, zum andern auch durch die individuelle Kombination der von den unterschiedlichen Vorfahren stammenden Wachstumsgene der Mutter. Zusätzlich spielen noch weitere Faktoren wie z.B. Ernährung und Umweltbedingungen eine gewisse, wenn auch untergeordnete Rolle. Unter großen Müttern sind vornehmlich solche zu finden, bei denen viele dieser Einflussfaktoren einen positiven Beitrag zur Körpergröße geleistet haben. Hinsichtlich der Größe der Töchter dieser großen Mütter wirkt sich die gemeinsame genetische Komponente sowie zusätzlich die individuelle genetische Konstellation der Töchter auf deren Körpergröße aus. Die gemeinsamen Komponenten stimmen bei den Töchtern großer Mütter mit denen ihrer Mütter überein und sind in der Regel positiv. Deshalb sind diese Töchter im Mittel immer noch überdurchschnittlich groß. Aber die weiteren Einflussfaktoren können sowohl in positive wie in negative Richtung wirken. Hier überwiegt im Gegensatz zu den Müttern keinerlei Richtung, es gibt in dieser Teilpopulation keinerlei Bevorzugung von positiv oder negativ wirkenden Einflussfaktoren. Bei den Müttern war dies anders, wurde doch die Teilgruppe

© 2002 metheval - jena

‚große Mütter' gerade dadurch ausgewählt, dass die Mutter groß sind und demzufolge die das Wachstum steuernden Einflussfaktoren im wesentlichen in die positive Richtung gewiesen haben.

Verdeutlichen wir diesen Mechanismus durch ein vereinfachtes Gedankenexperiment: Angenommen, die Körpergröße von Müttern und Töchtern läge nur in den fünf Stufen von sehr klein (150 cm), klein (160 cm), mittel (170 cm), groß (180 cm) bis sehr groß (190 cm) vor. Die von Müttern und Töchtern gemeinsam geteilten Wachstumsgene würden festlegen, ob ein Mensch klein (160 cm) oder groß (180) wird. Die davon unabhängige individuelle Disposition würde dann zusätzlich die Körpergröße um 10 cm steigern oder verringern. Untersucht man nun sehr große Mütter von 190 cm Körpergröße, so sind die gemeinsamen *und* die individuellen Anteile positiv. Für die Töchter dieser Mütter weist zwar die gemeinsame Komponente auf ‚groß' (180), die individuelle Komponente kann aber zur Größe von 190 cm oder 170 cm führen. Der Durchschnitt liegt unter 190 cm, die Töchter sehr großer Mütter sind daher weniger groß als ihre Mütter.

Verallgemeinern wir diese Argumentation auf beliebige Merkmale X und Y, dann läßt sich die Situation so formulieren: X und Y haben etwas ‚gemeinsames', d.h., sie lassen sich darstellen als[3]

$$X=E(X|Y)+\varepsilon_x \qquad \text{und} \qquad Y=E(Y|X) + \varepsilon_y\,.$$

Die Regressionen $E(X|Y)$ bzw. $E(Y|X)$ spiegeln das ‚Gemeinsame' beider Variablen wieder, die Residuen ε_x und ε_y die davon unabhängigen weiteren Einflussfaktoren, die X und Y bestimmen. Betrachtet man eine Teilpopulation von Personen mit einer überdurchschnittlichen Merkmalsausprägung x (im Beispiel die ‚großen' Mütter), dann ist die gemeinsame Komponente $E(X|Y)$ positiv und es gilt in dieser Teilpopulation $E(\varepsilon_x)>0$. Der Erwartungswert für Y in dieser Teilpopulation ist

$$E(Y) = E(Y|X=x) + E(\varepsilon_y).$$

Hier gilt jedoch $E(\varepsilon_y)=0$, da $E(Y)$ in dieser Teilpopulation gerade $E(Y|X=x)$ ist. In Analogie zum inhaltlichen Beispiel bedeutet dies, das in dieser Subpopulation der Erwartungswert von Y nur durch das ‚Gemeinsame', nicht durch die weiteren Faktoren bestimmt wird. Mit dieser formalen Betrachtung haben wir nun alle Hilfsmittel zur Hand um zu bestimmen, wann der Regressionseffekt auftritt und wann nicht.

4. Wann tritt der Regressionseffekt auf?

Betrachten wir zunächst den Fall linearer Zusammenhänge. Dazu orientieren wir uns wieder am Körpergrößenbeispiel. Dort ist die gemeinsame Verteilung der Mutter- und Tochtergröße gut durch eine bivariate Normalverteilung beschreibbar. In solchen Fällen ist bekannt, dass die Regression $E(Y|X)$ linear ist (z.B. Johnson & Kotz, 1972). Da wir z-standardisierte Variablen betrachten, ist die Steigung der Geraden gleich der Korrelation von X und Y, der Achsenabschnitt ist 0 (vgl. Wirtz & Nachtigall, 2002, S. 114). In der Stichprobe findet sich eine Korrelation zwischen X und Y von r=0.46. Wir erhalten damit den (geschätzten) lineare Zusammenhang

$$Y=0.46\cdot X +\varepsilon$$

und es ist

$$E(Y|\mathrm{X})= 0.46\cdot X.$$

Es zeigt sich der Regressionseffekt: Für Mütter mit einer (z-standardisierten) Größe von x=1 beträgt die zu erwartende Größe 0.46. Der Wert liegt näher am Mittelwert als der Ausgangswert. Genauso verhält es sich für alle anderen x-Werte außer der Null. Es liegt also Regression zur Mitte vor.

Allgemein wird deutlich, dass im Falle einer linearen Regression mit nicht perfekter Korrelation die vorhergesagten Werte *immer* näher am Mittelwert liegen. In diesem Fall ist der Regressionseffekt immer gegeben. Formal ausgedrückt. Ist $|Kor(X,Y)|<1$ und die Regression linear, dann gilt für z-transformierte Variablen

$$E(Y|X)=Kor(X,Y)\cdot X$$

[3] Um Zusammenhänge durch Regressionen zu beschreiben, müssen die Variablen X und Y numerisch sein (vgl. Steyer, 2002).

und Gleichung (2) ist erfüllt. Da bei den für die Psychologie relevanten Variablen perfekte Zusammenhänge praktisch nie vorkommen, anderseits lineare Zusammenhänge häufig angenommen werden, sind Regressionseffekte in der psychologischen Forschung als Normalfall zu betrachten.

Anders verhält es sich jedoch bei nichtlinearen Zusammenhängen. Hier kann es zu Situationen kommen, bei denen die zu erwartenden Werte von Y weiter vom Mittelwert weg liegen als die Ausgangswerte. Abbildung 2 zeigt einen solchen Zusammenhang zweier Variablen X und Y anhand von Stichprobendaten im Streudiagramm. In diesem Beispiel ist die Variable X gleichverteilt mit Werten 1, 2 und 3. Die Variable Y ist kontinuierlich und für jede Stufe von X normalverteilt mit Streuung 1. Es wurde die Regression durch die Werte $E(Y|X{=}1){=}0$, $E(Y|X{=}2){=}0$ und $E(Y|X{=}3){=}9$ vorgegeben. Zu Illustrationszwecken wurde eine Stichprobe vom Umfang N=300 simuliert[4]. Es ist $E(X){=}2$ und $E(Y){=}3$. Betrachtet man $E(Y|X{=}3){=}9$, dann ist dieser Wert ‚weiter' vom Erwartungswert entfernt als der Ausgangswert. Zum Nachweis über Gleichung (1) sind die Streuungen zu berechnen. Mit etwas Rechnerei läßt sich $\text{Std}(X){=}0.816$ und $\text{Std}(Y){=}4.359$ bestimmen. Demnach ist für x=3 der Abstand zu $E(X)$ gleich (3-2)/0.816=1.225 Standardabweichungen, der Abstand $E(Y|X{=}3)$ zu $E(Y)$ jedoch (9-3)/4.359=1.376 Standardabweichungen. Da 1.376 > 1.225, ist Gleichung (1) verletzt. Regression zur Mitte liegt nicht vor.

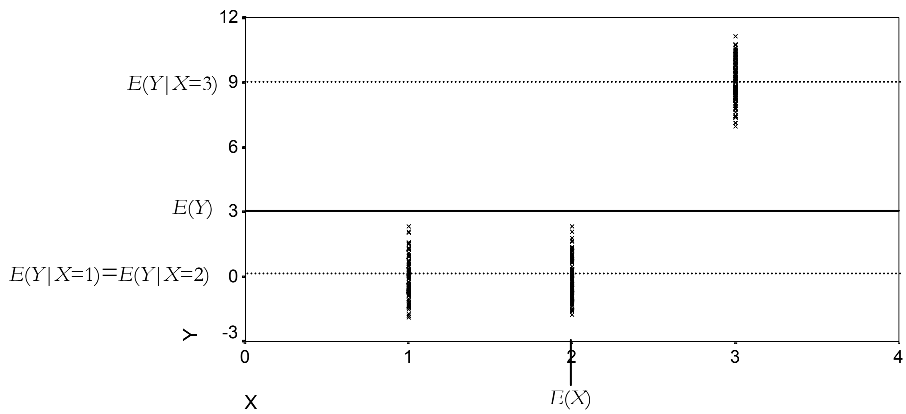

Abbildung 2: Ein nichtlinearer Zusammenhang ohne Regression zur Mitte. Für x=3 liefert die Regression den Wert 9, welcher ‚weiter' (ausgedrückt in Standardabweichungen) vom Mittelwert entfernt ist als der Wert 3 selbst.

Eine Tendenz ‚weg von der Mitte' findet sich in diesem Beispiel auch, wenn wir den Punkt x=2 betrachten. x=2 ist der Erwartungswert von X. Bei linearer Regressionen geht die Regressionsgerade immer durch die Erwartungswerte. Ist ein Merkmal durchschnittlich, so sind für das andere Merkmal ebenfalls durchschnittliche Werte zu erwarten. Im hiesigen Beispiel sieht das anders aus: Es ist $E(Y|X{=}2){=}0{<}3{=}\ E(Y)$. Für durchschnittliche Werte von X sind keineswegs durchschnittliche Werte von Y zu erwarten.

5. Mythos und Wirklichkeit

Bis heute existieren eine Reihe von Missverständnissen, Irrtümern und Mythen über den Regressionseffekt, welche wir nun im Einzelnen betrachten werden.

- ***Mythos 1: Der Regressionseffekt führt zu einer Schrumpfung der Streuung***

  Wenn Töchter großer Mütter weniger groß werden, die Töchter kleiner Mütter weniger klein werden, müßte sich beim Merkmal Körpergröße' doch mit der Zeit die Streuung immer kleiner werden, sich im Laufe der Generationen vielleicht sogar eine Einheitsgröße einstellen? Die Daten des Beispiels zeigen das Gegen-

[4] Die Daten dieses Beispiels befinden sich in unter http://www.uni-jena.de/svw/metheval/materialien/reports/daten/gegenbeispiel.sav.

© 2002 metheval - jena

teil. Die Streuung der Körpergröße von Müttern und Töchtern ist nahezu konstant bzw. wächst leicht an (vgl. Abbildung 1). Wie ist das mit dem Regressionseffekt vereinbar? Dazu ist zu sagen, dass es eine Art Gegenbewegung zum Regressionseffekt gibt: Mütter mit mittlerer Größe haben in aller Regel Töchter, die aufgrund der individuellen genetischen Komponente weiter nach oben oder unten vom Mittelwert abweichen. Dies gleicht den ‚Schwund an Streuung' wieder aus. Die Aufhellung des nächsten Mythos wird den Irrtum hinter diesem Mythos weiter verdeutlichen.

- ***Mythos 2: Regression zur Mitte ist ein gerichteter Effekt***
  Wir haben bisher ausgiebig analysiert, dass große Mütter weniger große Töchter haben und uns die Gründe dafür klargemacht. Wie ist es umgekehrt? Betrachten wir große Töchter, dann sind mit der gleichen Begründung wie vorher weniger große Mütter zu erwarten. Ein Blick in die Daten bestätigt dies: Definiert man große Töchter durch eine Mindestgröße von einer Standardabweichung über dem Mittelwert, so erfüllen 95 Personen der Stichprobe dieses Kriterium. Die Kennwerte der z-transformierten Körpergröße in dieser Gruppe gibt Tabelle 3 wieder.

Tabelle 3 Kennwerte des z-transformierten Merkmals ‚Körpergröße der Mütter'. Die Daten stammen aus der Teilstichprobe der ‚großen' Frauen (d.h. $z_{Größe}>1$, N=95).

	N	*Minimum*	*Maximum*	*Mittelwert*	*Standardabweichung*
z-Wert(Körpergröße Mütter)	*95*	*-2.16*	*3.60*	*.63*	*.99*

Tatsächlich sind die Mütter großer Töchter im Durchschnitt weniger groß. Der Mittelwert in der Stichprobe ist 0.63, obwohl alle Töchter eine Größe von über 1 hatten. Entsprechende Ergebnisse zeigen die Analysen der Daten kleiner Töchter und deren Müttern.

Dies macht deutlich, dass Regression zur Mitte kein gerichtetes Phänomen ist. Sowohl von X auf Y als auch umgekehrt gibt es Regression zur Mitte. Wird dieselbe Personeneigenschaft zweimal an den gleichen Personen gemessen (Messung X und Messung Y), so werden Personen mit überdurchschnittlichen Werten in Messung X in der anderen Messung tendenziell näher beim Durchschnitt liegen. Dies kann z.B. auch bei Messungen zum gleichen Zeitpunkt mit parallelen Messinstrumenten beobachtet werden[5]. Der Regressionseffekt hat nichts mit zeitlicher Reihenfolge zu tun. Im Allgemeinen ist der Regressionseffekt ungerichtet (vgl. auch Furby, 1973). Es gibt jedoch eine wichtige Ausnahme, die wir im Zusammenhang mit dem nächsten Mythos kennenlernen werden.

- ***Mythos 3: Prä-Post-Differenzen sind durch Regression zur Mitte verfälscht***
  Für besonders viel Irritation sorgt der Regressionseffekt, wenn es um zeitliche Veränderung von Merkmalen im Rahmen von Messwiederholungen geht. Um Veränderung über die Zeit zu erfassen, bietet es sich an, bei einem vor- und nach einer Intervention erhobenen Merkmal (Prä-Post-Design) die Differenz der Messwerte zu betrachten. Die Verwendung solcher Differenzenscores wird jedoch unter anderem mit dem Argument kritisiert, dass aufgrund des Regressionseffektes Veränderung ohnehin zu erwarten seien. Es werden Alternativen vorgeschlagen, bei denen der Posttestwert mit dem vom Regressionseffekt ‚bereinigten' Posttestwert verglichen wird (z.B. Steyer et al., 1997). Dabei wird jedoch nicht berücksichtigt, dass der Regressionseffekt nicht nur vom Prä- zum Posttest, sondern auch vom Post- zum Prätest auftritt (siehe Mythos 2). Differenzenscores sind daher keineswegs automatisch vom Regressionseffekt verfälscht, sonder stellen vielmehr im Allgemeinen einen erwartungstreuen Schätzer für die wahre Veränderung dar

[5] Wer dies selber nachprüfen möchte, findet im Datensatz http://www.uni-jena.de/svw/metheval/materialien/reports/daten/wt-daten.sav die Ergebnisse von N=706 Personen, deren Kenntnisse in Wahrscheinlichkeitsrechnung mit dem Fragebogen zur Wahrscheinlichkeitstheorie (FWT) (Nachtigall und Wolf, 2001) untersucht wurden. Die Variablen test1 und test2 sind parallele Testhälften des FWT.

(vgl. Rogosa, 1995; Nachtigall & Suhl, 2002), wohingegen die einseitige ‚Bereinigung' zu verfälschten Schätzern der wahren Veränderung führt.

Allerdings gibt es eine wichtige Ausnahme, bei der es aufgrund des Regressionseffektes zu gänzlich falschen Ergebnissen bei Prä-Post-Differenzen kommen kann. Machen wir uns das an einem Beispiel klar: Bei einer Untersuchung zur Wirksamkeit einer neuen Behandlungsmethode wird die Befindlichkeit der Patienten in einem Vortest erhoben. Daraufhin wird die Gruppe der Patienten, denen es am schlechtesten geht, mit der neuen Therapie behandelt. Nach der Behandlung wird erneut die Befindlichkeit erhoben. Selbst wenn die Behandlung vollkommen wirkungslos ist, wird aufgrund des Regressionseffektes der Mittelwert in der behandelten Gruppe bei der Post-Messung zum Gesamtmittelwert tendieren und einen 'falschen' Behandlungseffekt vorgaukeln. Umgekehrt würde ein tatsächlicher Behandlungseffekt dadurch verschleiert, wenn nur die andere Extremgruppe der Patienten mit überdurchschnittlich guter Befindlichkeit behandelt würde. Generell ist mit Verfälschungen durch Regressionseffekte immer dann zu rechnen, wenn Extremgruppen untersucht werden. Damit ist z.B. dann zu rechnen, wenn sich Patienten aufgrund ihres momentanen Leidensdruckes selbst zu einer therapeutischen Behandlung anmelden. Da die Symptomschwere zum Zeitpunkt der Anmeldung in der Regel höher ist der individuelle Durchschnittswert, ist zu anderen Messzeitpunkten mit niedrigeren Messwerten zu rechnen. In einer solchen Situation sind Prä-Post-Vergleiche kein geeignetes Mittel zur Abschätzung der therapeutischen Wirkung. Statt dessen sollten Vergleiche mit Kontrollgruppen herangezogen werden.

- ***Mythos 4: Der Regressionseffekt ist ein statistischer Effekt***

In der statistischen Lehrbüchern wird der Regressionseffekt manchmal als eine Eigenschaft von Regressionsgeraden dargestellt (z. B. Mortensen, 1995, S.109). Für z-transformierte Variablen läßt sich die Regressionsgerade von Y auf X schreiben als

$$\hat{Y} = r \cdot X,$$

wobei r die Pearson-Korrelation ist (Wirtz & Nachtigall, 2002, S. 114). Ist $|r|$ kleiner als 1, dann sind die vorhergesagten Werte immer näher am Mittelwert als die Ausgangswerte.

Diese Argumentation ist keineswegs falsch, nur weckt sie den Eindruck, der Regressionseffekt sei eine Eigenschaft von Regressionsgeraden. Statt dessen ist Regression zur Mitte ein allgemeineres Phänomen, welches nicht-determinstische lineare Zusammenhänge immer betrifft, aber auch bei nicht-linearen Zusammenhänge vorkommen kann. Im Falle eines linearen Zusammenhanges kann dieser durch eine Gerade beschrieben werden, mit Hilfe derer der Regressionseffekt nachweisbar ist.

Generell gründet der Regressionseffekt neben dem in Abschnitt 3 beschrieben ‚Funktionsmechanismus' auch auf grundlegende Eigenschaften von Mittel- bzw. Erwartungswerten. Es geht um zu erwartende Durchschnittswerte, für Einzelfälle können die Werte einer Variable Y auch im linearen Fall weiter vom Mittelwert entfernt liegen als der Ausgangswert der Variablen X. Konkret: Große Mütter können noch größere Töchter haben. In diesem Sinne handelt es sich beim Regressionseffekt durchaus um einen 'statistischen' Effekt, allerdings dürften viele Menschen mittlere Werte im Rahmen ihres intuitiven Denkens ganz selbstverständlich verwenden, ohne sich immer darüber klar zu sein, dass sie mit einem 'statistischen Modell' arbeiten.

- ***Mythos 5: Regression zur Mitte beruht auf mangelnder Reliabilität des Messinstrumentes***

Häufig wird Regression zur Mitte lediglich im Zusammenhang mit mangelnden Reliabilität eines Messinstrumentes diskutiert. Bortz & Döring (1995, S. 517) sprechen in ihrem Methodenlehrbuch von Regressionseffekten als ...“statistischen Artefakten (...), die durch mangelnde Reliabilität der Messinstrumente verursacht werden.“ Das Beispiel der Körpergröße verdeutlicht jedoch, dass selbst bei einer perfekt reliablen Messung der Regressionseffekt auftreten würde. Entscheidend ist, dass die untersuchten Merkmale aus einer gemeinsame determinierten Komponente und einer davon unabhängigen Restkomponente bestehen (vgl. Abschnitt 3). Diese Restkomponente kann weitere Einflußfaktoren und / oder Messfehler repräsentieren.

© 2002 metheval - jena

Mythos 6: Regression zur Mitte ist unvermeidlich

Diese Auffassung findet sich in der Literatur noch verbreitet. So bezeichnet Rogosa (1995, S.19) Gleichung (1) im Falle nicht perfekter Korrelation als allgemein gültig. Dies trifft jedoch nicht zu, wie wir am Beispiel eines nicht-linearen Zusammenhanges gesehen haben. Bei linearen Zusammenhängen kann allerdings immer von Regression zur Mitte ausgegangen werden. Welche Konsequenzen dass für den praktischen Forscher hat, soll im Folgenden nochmals bilanziert werden.

6. Fazit

Der Regressionseffekt ist ein wichtiger und zu berücksichtigender Effekt, wenn es um Versuchsplanung und Interpretation von Effekten insbesondere im Rahmen von Längsschnittuntersuchungen geht. Speziell besteht bei der Untersuchung von Extremgruppen die große Gefahr, dass der Regressionseffekt mögliche Behandlungseffekte verfälscht. Er ist kein statistisches Artefakt in dem Sinne, dass er nur auf mangelnde Reliabilität zurückzuführen ist oder dass nur die durch ein Regressionsgleichung vorhergesagten Werte näher am Mittelwert liegen. Die Datenbeispiele zeigen, dass der Effekt 'wirklich' existiert, z.B. dann, wenn der Zusammenhang der Merkmale tatsächlich linear ist. Allerdings stellt der Regressionseffekt keine Verfälschung in dem Sinne dar, dass bei Längsschnittuntersuchungen die Messwerte mit der Zeit alle zur Mitte wandern. Im Gegenteil, der Effekt ist unabhängig von zeitlicher Abfolge. Regression zur Mitte findet sich bei den Töchtern großer Mütter und bei den Müttern großer Töchter. Die in der Literatur vorgebrachten Einwände gegen die Verwendung von Differenzenscores aufgrund des Regressionseffektes (z.B. Steyer, 1997) sind daher so nicht haltbar (vgl. hierzu auch Rogosa, 1995, sowie Nachtigall & Suhl, 2002, Suhl & Nachtigall, 2002).

Zu welchen Konsequenzen eine falsche Interpretation des Regressionseffektes führen kann, zeigt auf eindrucksvolle Weise das von Furby (1973) präsentierte Beispiel von ‚Intelligenz‘ und ‚sozio-ökonomischem Status‘ (SÖS). Beide Merkmaler sind positiv korreliert. Aufgrund des Regressionseffektes findet sich bei Kindern von hochintelligenten Eltern im Mittel eine geringere Intelligenz. Aufgrund der positiven Korrelation von Intelligenz und SÖS sind auch die Kinder von Eltern mit hohem SÖS weniger Intelligent als ihre Eltern. Eysenck (1971, zitiert nach Furby, 1973) interpretierte diesen Regressionseffekt als negativen Zusammenhang von Intelligenz und SÖS-bestimmten Entwicklungsfaktoren. Macht demnach hoher Status die Kinder dumm? Beachten wir den Regressionseffekt in umgekehrter Richtung, so sind die Eltern hochintelligenter Kinder weniger intelligent als ihre Kinder. Da eine positive Korrelation zwischen Intelligenz und SÖS besteht, kommen überdurchschnittlich intelligente Kinder vermehrt aus einem Elternhaus mit überdurchschnittlichem SÖS. Analog Eysencks Interpretation ließe sich folgern, dass überdurchschnittlicher SÖS die Kinder noch intelligenter werden läßt als ihre Eltern, mithin SÖS ein wichtiger *positiver* Faktor für die Intelligenzentwicklung der Kinder ist. Macht hoher Status demnach die Kinder klug?

Das Beispiel zeigt deutlich, zu welch gravierenden Fehlinterpretationen ein falsches Verständnis des Regressionseffektes führen kann.

Literatur

Bortz, J. & Döring, N. (1995): *Forschungsmethoden und Evaluation.* (2. Aufl.), Berlin: Springer.

Furby, L. (1973): Interpreting Regression toward the Mean in Developmental Research. *Developmental Psychology, 8, 2,* 172-179.

Gigerenzer, G., Swijtink, Z., Porter, Th., Daston, L., Beatty, J. & Krüger, L. (1999): *Das Reich des Zufalls. Wissen zwischen Wahrscheinlichkeit, Häufigkeit und Unschärfen.* Berlin: Spektrum

Johnson, N. L. & Kotz, S. (1972): *Distributions in Statistics: Continous Multivariate Distributions.* New York: Wiley.

Lord, F. M. (1963): Elementary Models for measuring change. In C. W. Harris (Ed.), *Problems in measuring change.* Madison: University of Wisconsin Press.

Mortensen, U. (1995): *Mathematische und statistische Methoden I. Deskriptive Verfahren.* Münster: Lit.

Nachtigall, C. & Suhl, U. (2002). Warum kompliziert, wenn es einfach geht. Teil 1: Zur Analyse intraindividueller Veränderung. *metheval report 4* (3), www.uni-jena.de/svw/metheval/report/

Nachtigall, C. & Wolf, A. (2001). Fragebogen zur Wahrscheinlichkeitstheorie (FWT). Handanweisung. *metheval report 3* (2), www.uni-jena.de/svw/metheval/report/

Rogosa, D. (1995). Myths and Methods: Myths about longitudinal research. In J. M. Gottman (Ed.). *The analysis of change* (p. 3-66), Mahwah: Erlbaum.

Steyer, R. (2002): *Wahrscheinlichkeit und Regression.* Berlin: Springer.

Steyer, R., Hannöver, W., Telser, Ch. & Kriebel, R. (1997). Zur Evaluation intraindividueller Veränderung. *Zeitschrift für Klinische Psychologie*, 26, 291-299.

Suhl, U. & Nachtigall, C. (2002). Warum kompliziert, wenn es einfach geht. Teil 2: Ergebnisse einer Simulationsstudie. *metheval report 4* (4), www.uni-jena.de/svw/metheval/report/

Wirtz, M. & Nachtigall, C. (2002): *Deskriptive Statistik. Statistische Methoden für Psychologen. Teil 1.* (2. Aufl.), Weinheim: Juventa.

© 2002 metheval - jena