Tücken von Korrelationen: die Korrelationskoeffizienten von Pearson und Spearman

Ulrike Held
Horten-Zentrum, UniversitätsSpital, Zürich


Ulrike Held

Die Autorin erklärt, dass sie keine Interessenkonflikte im Zusammenhang mit diesem Beitrag hat.

In der klinischen Forschung werden häufig mehrere Parameter, wie z.B. das Gewicht und der systolische Blutdruck, an den Patienten gemessen. Zusätzlich zu der separaten Beschreibung jeder einzelnen Datenreihe, auch Variable genannt, ist es oft interessant zu untersuchen, ob es zwischen zwei Variablen Zusammenhänge gibt. Es kann z.B. interessieren, wie die eine Variable sich verhält, sobald die andere fällt oder steigt, wie die Art des Zusammenhangs ist und ob man den Zusammenhang quantifizieren kann, d.h., ob man Stärke und Richtung des Zusammenhangs angeben kann. Je nachdem, ob es sich um *metrische* (z.B. systolischer Blutdruck) oder *ordinale* (z.B. die Befindlichkeit gemessen auf einer Skala von 1 bis 7) Variablen handelt, gibt es unterschiedliche Zusammenhangs- oder Assoziationsmasse, die verwendet werden können. Anders als bei der linearen Regression ist die Festlegung einer Variablen als Ziel- und der anderen als Einflussgrösse nicht notwendig.

Wir betrachten hierzu wieder das fiktive Beispiel, welches wir bereits in Artikel «Wissenschaftliche Fragestellungen in der Medizin brauchen statistische Modelle»[1] herangezogen hatten: An 20 Patienten wurden der systolische Blutdruck und das Gewicht gemessen, und die Beobachtungen sind in Tabelle 1 dargestellt.

Es ist zu erkennen, dass bei dem Patienten mit der Nummer 6 ein fehlender Wert in der Variablen Gewicht vorliegt und bei Patient 8 ein fehlender Wert im systolischen Blutdruck.

Bevor man mit der Berechnung von Zusammenhangsmassen, wie z.B. dem Korrelationskoeffizienten für zwei metrische Variablen, beginnt, sollte man die vorliegenden zwei Messreihen graphisch in einem sogenannten *Streudiagramm* (engl. Scatterplot) darstellen. Für die graphische Darstellung der zwei Messreihen sowie auch für die Berechnung des Korrelationskoeffizienten gilt, dass nur vollständige Beobachtungspaare verwendet werden können. Für unseren Beispieldatensatz bedeutet dies also, dass nur 18 Beobachtungspaare eingehen und die einzelnen Beobachtungen einer Person, für die der zweite Wert fehlt, aus der Analyse ausgeschlossen werden müssen (hier Patienten Nummer 6 und 8). Man sollte sicherstellen, dass das Fehlen von Werten unabhängig von ihrem Wert selbst ist. Es darf z.B. nicht sein, dass häufig Patienten mit sehr niedrigen oder hohen Blutdruckmessungen fehlen, denn das würde die Ergebnisse verfälschen. Es gibt in der Praxis keine universell gültige Strategie, wie mit fehlenden Werten umgegangen werden soll, stattdessen sollte man in jedem Fall individuell abwägen. In Abbildung 1 sind die 18 vollständigen Datenpaare gegeneinander aufgetragen worden.

Die Stärke und Richtung des linearen Zusammenhangs zwischen zwei metrischen Messreihen können mit dem Pearson-Korrelationskoeffizienten beurteilt werden. Die Korrelation misst, wie stark der lineare Zusammenhang zwischen den beiden Datenreihen ist. Der Wert des Korrelationskoeffizienten liegt zwischen –1 und 1. Für den Fall, dass kein linearer Zusammenhang zwischen den beiden Messreihen vorliegt, ist der zugehörige Korrelationskoeffizient ungefähr gleich null, und man spricht davon, dass die beiden Variablen unkorreliert sind. Es kann in diesem Fall trotzdem ein Zusammenhang zwischen den Variablen vorliegen, der aber nicht linear ist.

Das Vorzeichen des Korrelationskoeffizienten gibt die Richtung des Zusammenhangs an: Ist es positiv, gehen grosse Werte der einen Variablen mit grossen Werten der anderen Variablen einher; ist das Vorzeichen negativ, so gehen grosse Werte der einen Variablen mit kleinen Werten der anderen einher. In Abhängigkeit vom untersuchten Zusammenhang gibt es Richtwerte zur Interpretation der Korrelation. Ab einem Wert grösser 0,3 (kleiner als –0,3) spricht man typischerweise von einem schwachen positiven (negativen) Zusammenhang, ab 0,5 bzw. –0,5 von einem moderaten Zusammenhang, und ab 0,8 bzw. –0,8

Tabelle 1. Messungen von Gewicht und systolischem Blutdruck bei 20 Männern.

Patient	Gewicht (kg)	Systolischer Blutdruck (mm Hg)
1	92,8	153,6
2	82,9	116,2
3	101,6	157,8
4	97,7	153,5
5	111,3	153,2
6	fehlend	123,3
7	73,3	128,4
8	87,2	fehlend
9	114,7	126,1
10	113,1	167,9
11	97,4	130,5
12	90,2	141,2
13	95,0	109,9
14	89,4	89,8
15	90,4	137,6
16	92,2	114,1
17	105,1	130,6
18	89,4	138,1
19	88,7	109,4
20	86,3	103,5

1 U. Held. «Wissenschaftliche Fragestellungen in der Medizin brauchen statistische Modelle». Schweiz Med Forum. 2010;10(32):528–30.

von einer starken linearen Assoziation zwischen den beiden Variablen.
Für unseren Beispieldatensatz ergibt sich eine Pearson-Korrelation von 0,48.
Der Pearson-Korrelationskoeffizient kann durch extreme Werte oder Ausreisser stark beeinflusst werden. In diesem Fall, oder wenn die vorliegenden Messreihen nicht metrisch, sondern ordinal skaliert sind, ist die Verwendung des Rangkorrelationskoeffizienten nach Spearman angezeigt. Beim Rangkorrelationskoeffizienten werden anstelle der Messwerte selbst nur noch deren *Ränge* verwendet. Das bedeutet, man sortiert die Beobachtungswerte ihrer Grösse nach und ersetzt dann jeden Messwert durch seinen Rang. Der Rangkorrelationskoeffizient liegt wiederum

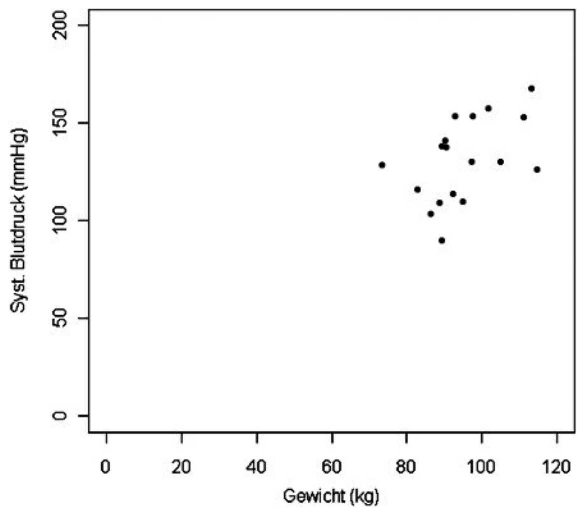

Abbildung 1
Streudiagramm der 18 Beobachtungspaare.

Tabelle 2. Messwerte und ihre zugehörigen Ränge, sortiert nach aufsteigender Grösse des Gewichts.

Patient	Gewicht (kg)	Rang Gewicht	Systolischer Blutdruck (mm Hg)	Rang Blutdruck
7	73,3	1	128,4	9
2	82,9	2	116,2	6
20	86,3	3	103,5	2
8	87,2	4		
19	88,7	5	109,4	3
14	89,4	6,5	89,8	1
18	89,4	6,5	138,1	13
12	90,2	8	141,2	14
15	90,4	9	137,6	12
16	92,2	10	114,1	5
1	92,8	11	153,6	17
13	95,0	12	109,9	4
11	97,4	13	130,5	10
4	97,7	14	153,5	16
3	101,6	15	157,8	18
17	105,1	16	130,6	11
5	111,3	17	153,2	15
10	113,1	18	167,9	19
9	114,7	19	126,1	8
6			123,3	7

im Intervall von –1 bis 1, und es gelten dieselben Interpretationsregeln wie für den Pearson-Koeffizienten. Für die Daten aus unserem obigen Beispiel haben wir die Beobachtungen und deren Ränge in Tabelle 2 dargestellt. Die Reihenfolge der Patienten ist sortiert nach zunehmendem Gewicht, somit ist die Rangfolge des Gewichts auch aufsteigend.
Hier fällt auf, dass in der Variablen Gewicht der Wert von 89,4 genau zweimal vorkommt. In diesem Fall wird zweimal der Durchschnittsrang 6,5 vergeben, anstelle der Ränge 6 und 7. Man spricht hier vom Vorliegen einer Bindung.
Für unser obiges Beispiel ist der geschätzte Rangkorrelationskoeffizient zwischen Gewicht und systolischem Blutdruck gleich 0,54. Aufgrund der Tatsache, dass hier nur 18 Beobachtungspaare vorliegen, sollte man den Rangkorrelationskoeffizienten dem Pearson-Koeffizienten vorziehen und von einem moderaten Zusammenhang zwischen Gewicht und systolischem Blutdruck ausgehen.

Glossar

Metrisch
Sind die Beobachtungen auf einer kontinuierlichen Skala gemessen worden, so spricht man von metrischen Grössen. Beispiele: Blutdruck, Grösse in cm.

Ordinal
Sind die Beobachtungen auf einer Rangskala gemessen worden, bei der eine natürliche Reihenfolge besteht, spricht man von ordinal skalierten Grössen. Beispiele: Zufriedenheit mit einem Produkt (sehr zufrieden, zufrieden, weniger zufrieden), schulische Leistung (sehr gut, gut, befriedigend usw.).

Rang
Die Beobachtungen einer Variablen oder Messreihe werden der Grösse nach geordnet und durchnummeriert: die kleinste Beobachtung erhält den Rang 1, die zweitkleinste Beobachtung den Rang 2 usw.

Streudiagramm
In einem Streudiagramm werden Wertepaare zweier Variablen gegeneinander abgetragen, wodurch eine Punktwolke entsteht. Englisch: Scatterplot.

Korrespondenz:
Dr. Ulrike Held
Horten-Zentrum
UniversitätsSpital Zürich
Postfach Nord
CH-8091 Zürich
ulrike.held@usz.ch

Empfohlene Literatur
– Bland JM, Altman DG. Statistics notes: correlation, regression, and repeated data. BMJ. 1994;308:896.
– Held L, Rufibach C, Seifert B. Einführung in die Biostatistik. 4. Auflage. Zürich: Abteilung Biostatistik, Institut für Sozial- und Präventivmedizin der Universität Zürich; Juli 2009. http://www.biostat.uzh.ch.
– Hüsler J, Zimmermann H. Statistische Prinzipien für medizinische Projekte. 4. Auflage. Bern: Huber-Verlag; 2006.
– Kreienbrock L, Schach S. Epidemiologische Methoden. 4. Auflage. München: Elsevier-Verlag; 2005.
– R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2008. ISBN 3-900051-07-0. URL http://www.R-project.org.