

Wolfram Fischer

Streuungsfächerkarte zur Visualisierung einer zweifachen Abhängigkeit

Zentrum für Informatik und wirtschaftliche Medizin, Wolfertswil SG (Schweiz)

Mai 2010

Einleitung

Zur Berechnung der Abhängigkeit eines Messwertes von einer oder mehreren Variablen werden Regressionsrechnungen eingesetzt. Eine Kontrollgrafik zur Beurteilung der Streuung der Messwerte in Abhängigkeit einer einzigen unabhängigen Variablen kann durch Aneinanderreihung von mehreren Boxplots erstellt werden. Wenn jedoch die Streuungen einer Messvariablen in Abhängigkeit von zwei unabhängigen Variablen dargestellt werden sollen, ist eine Grafik fixer Ausdehnung wesentlich praktischer und erlaubt zudem eine kompaktere Darstellung.¹

Beispiel

Im folgenden Beispiel soll gezeigt werden, ob die Resultate von Mathematikprüfungen nur vom sozioökonomischen Status der Prüflinge oder auch vom durchschnittlichen sozioökonomischen Status der MitstudentInnen an der gleichen Schule beeinflusst werden.

Methode

► Tafel 1

Zur Darstellung der Streuungen wurden «Streuungs-fächerkarten»² verwendet. Damit lassen sich in kompakter Weise die gleichen Streuungskennzahlen wie mit Boxplots abbilden, insbesondere also Median und Quartile. Dank der kreisförmigen Darstellung sind direkte Vergleiche gleichermaßen sowohl in waagrechter wie auch in senkrechter Richtung möglich.

Als Skala wird eine Kreislinie benutzt. Sie beginnt ganz links, z. B. mit Null. Die weiteren Werte werden im Uhrzeigersinn aufgetragen. Die weisse Mittellinie zeigt auf den Wert des Medians. Der schwarze Fächer zeigt die Streuung der mittleren 50% der Messwerte an; er reicht also vom ersten bis zum dritten Quartil. Mit den weissen Federn wird der Streubereich der mittleren 90% der Messwerte dargestellt. Die Länge der weissen Mittellinie korrespondiert mit der Anzahl Messwerte.

Daten

Das für das folgende Beispiel verwendete Datenset stammt aus dem R-Paket «nlme».³ Es beinhaltet Resultate zu 7185 Mathematikprüfungen. Die StudentInnen sind kategorisiert nach Geschlecht und Zugehörigkeit zu einer ethnischen Minorität.⁴

Grafik

► Tafel 2

In der Grafik wurden die Prüfungsergebnisse in Abhängigkeit vom sozioökonomischen Status der StudentInnen (x-Achse) und vom durchschnittlichen sozioökonomischen Status der MitstudentInnen an ihrer Schule (y-Achse) abgebildet. In den vier Hauptgrafikfeldern werden StudentInnen nach ihrem Geschlecht und nach der Zugehörigkeit zu einer Minorität unterschieden.

Resultate

Als Erstes fallen die weit geöffneten Fächer auf, d. h. die grossen Streuungen der Prüfungsergebnisse in praktisch allen Untergruppen. Der Vergleich der vier Hauptgrafikfelder zeigt

¹ Das folgende Beispiel stammt aus Fischer [Neue Grafiken I, 2010]: 38 f.

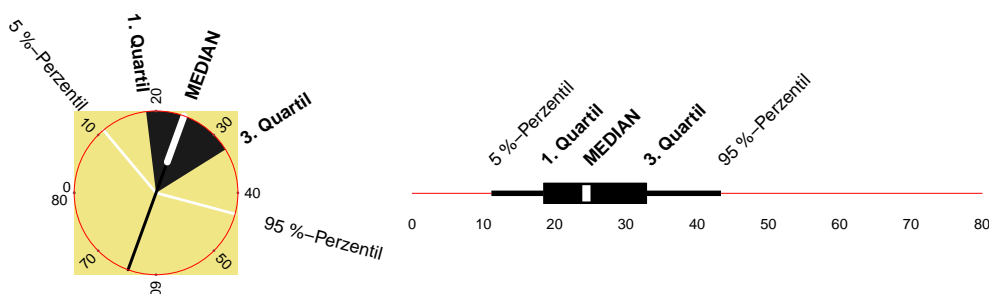
² Fischer [Neue Grafiken I, 2010]: 27 ff.

³ <http://www.r-project.org/>: «Linear and Nonlinear Mixed Effects» von Pinheiro J et al.

⁴ Im amerikanischen Original: «member of a minority racial group».

Tafel 1:

Legende zum Streuungs-fächer und Vergleich mit Boxplot



grundsätzlich bessere Prüfungsergebnisse bei StudentInnen, die nicht einer Minorität angehören, sowie leicht bessere Resultate der Männer im Vergleich zu den Frauen.

Innerhalb der Felder ist festzustellen, dass sowohl der sozioökonomische Status der StudentInnen als auch – etwas geringer – der durchschnittliche sozioökonomische Status der Schule positiv mit den Prüfungsergebnissen korrelieren.

Diese Grafik zeigt anschaulich, wie der mittlere Wert z.T. einer grossen Haupttendenz folgt, obwohl die Werte in den einzelnen Untergruppen stark streuen und man deshalb über eine mögliche Korrelation ins Zweifeln geraten könnte. (Besonders gut ist dies in den unteren beiden Feldern sichtbar.)

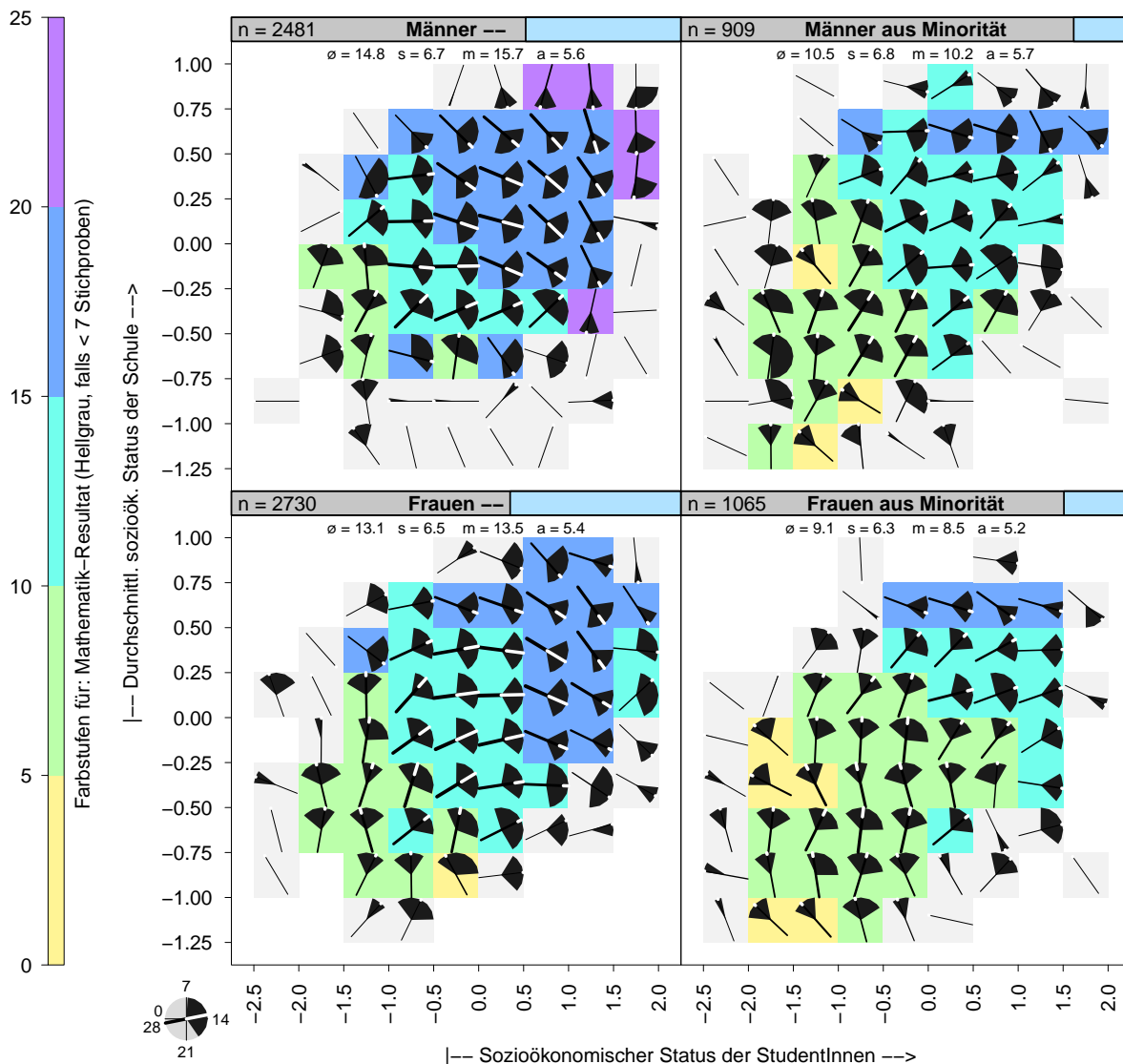
Konklusion

Literaturhinweis

Fischer (2010) Neue Grafiken I

Fischer W. *Neue Grafiken zur Datenvisualisierung*. Band 1: Speichengrafiken, Streuungsfächerkarten, Differenz-, Sequenz- und Wechseldiagramme. Wolfertswil (ZIM) 2010: 107 S. Internet: <http://www.fischer-zim.ch/studien/Neue-Grafiken-I-1003-Info.htm>.

Tafel 2: 7185 Mathematikprüfungen: Resultate nach Geschlecht und Zugehörigkeit zu einer Minorität



Die schwarzen Fächer zeigen die Werte der Variablen "Mathematik-Resultat" vom ersten bis zum dritten Quartil auf einer Kreisskala von 0 bis 28.
 Kennzahlen: ø = Durchschnitt, s = Standardabweichung, m = Median, a = ø abs. Abw.

Datenquelle:
R-Paket «nlme»

n = 7185

Z I M
[example.fanchart.mathachieve-1034]