

planung & analyse

Zeitschrift für Marktforschung und Marketing www.planung-analyse.de
Eine Marke der dfv Mediengruppe

6/2015 D11700F

Datenschutz

EU-Verordnung
ante portas

Special

Forschung für
B2B-Märkte

Schwerpunkt Kommunikation

Um jeden
Preis
auffallen



Diskriminanzanalyse

Ausgehend von einer Gruppierung von Objekten beschäftigt sich eine Diskriminanzanalyse mit

- ▶ der Identifikation der Variablen bezüglich derer sich diese Gruppen voneinander trennen lassen,
- ▶ der Zuordnung von „neuen“ Objekten zu den Gruppen auf Basis ihrer Variablenausprägungen (Klassifizieren).

Mit der ersten Aufgabe befasst sich dieser Beitrag, mit der zweiten der kommende Beitrag dieser Reihe.

Ableitung der Diskriminanzfunktionen

Abbildung 1 zeigt zwölf Objekte, die anhand von zwei Eigenschaften charakterisiert und jeweils einer Gruppe (rot, blau und grün) zugeordnet sind. Mit Hineinlegen einer ersten Geraden gelingt es, die blaue und die grüne Gruppe vollständig voneinander zu trennen. Eine zweite Gerade trennt zudem eindeutig zwischen der blauen und der roten Gruppe. Das Hineinlegen von Geraden in das durch die Eigenschaften aufgespannte Koordinatensystem korrespondiert mit der Bestimmung von Diskriminanzfunktionen, im Fall mit zwei Eigenschaften also von $y = b_0 + b_1 \cdot x_1 + b_2 \cdot x_2$

wobei x_1 und x_2 die Eigenschaftsausprägungen, b_0 das Absolutglied sowie b_1 und b_2 die Diskriminanzkoeffizienten darstellen, in denen sich die diskriminatorische Bedeutung der Variablen widerspiegelt. Für die erste Trennung zwischen blau und grün ist insbesondere Eigenschaft 1 verantwortlich. Das heißt für die entsprechende Diskriminanzfunktion sollte der Betrag von b_1 größer als der von b_2 sein. Für die zweite Trennung zwischen blau und rot ist vor allem Eigenschaft 2 maßgeblich. Insofern sollte in einer zweiten Diskriminanzfunktion der Betrag von b_2 größer als der von b_1 sein. Durch Einsetzen der Eigenschaftsausprägungen in eine Diskriminanzfunktion lässt sich für jedes Objekt sein Diskriminanzwert y berechnen.

Ein Umstellen der Diskriminanzfunktion nach x_2 führt zu:

$$x_2 = (y - b_0) / b_2 - b_1 / b_2 \cdot x_1$$

und beschreibt allgemein eine mögliche Trenngerade. Die Gerade, die durch den Ursprung des Koordinatensystems und orthogonal zu der nach x_2 umgestellten Diskriminanzfunktion verläuft, ist die Diskriminanzachse: $x_2 = b_2 / b_1 \cdot x_1$.

Auf dieser können die Diskriminanzwerte dargestellt werden, indem von jedem Objekt das Lot auf die Diskriminanzachse gefällt wird, in der Abbildung beispielhaft dargestellt für ein Objekt der roten Gruppe. Ziel ist es, die Steigung b_2 / b_1 der Diskriminanzachse so zu bestimmen, dass

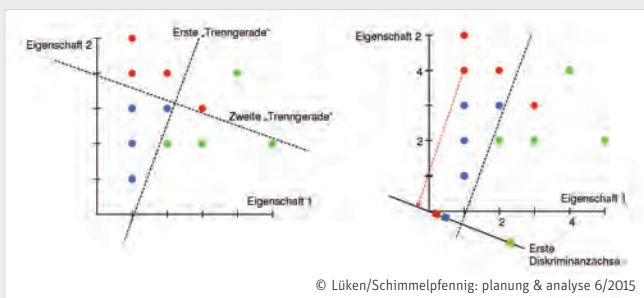


Abbildung 1: Bestimmung einer Diskriminanzachse.

sich das arithmetische Mittel der Diskriminanzwerte einer Gruppe möglichst stark von denen der anderen Gruppen unterscheidet – wie in der Abbildung das der grünen ■ von dem der roten ■ und blauen ■. Die eingezeichnete erste Diskriminanzachse erfüllt dieses Ziel. Ihre Steigung determiniert das Verhältnis der Koeffizienten der ersten Diskriminanzfunktion, aber nicht deren absolute Höhe. Mathematisch sind die Koeffizienten so zu bestimmen, dass die Streuung der Diskriminanzwerte zwischen den Gruppen – die durch eine Diskriminanzfunktion erklärte Streuung – in Relation zur nicht erklärten Streuung der Diskriminanzwerte innerhalb der Gruppen möglichst groß wird. Um eindeutige Koeffizienten zu erhalten, werden diese schließlich so normiert, dass unter anderem der Mittelwert aller Diskriminanzwerte einer Diskriminanzfunktion gleich null ist. Somit ergeben sich für das Datenbeispiel die Diskriminanzfunktionen:

$$y_1 = -1,07 + 0,96 \cdot x_1 - 0,35 \cdot x_2 \text{ und } y_2 = -3,76 + 0,36 \cdot x_1 + 1,02 \cdot x_2$$

Gruppe	Arithmetisches Mittel der Diskriminanzwerte		Erklärte Streuung der Diskriminanzwerte	
	Erste Diskriminanzfunktion	Zweite Diskriminanzfunktion	Erste Diskriminanzfunktion	Zweite Diskriminanzfunktion
blau	-0,66	-1,02	$4 \cdot (-0,66 - 0)^2 = 1,74$	$4 \cdot (-1,02 - 0)^2 = 4,16$
rot	-0,79	0,95	$4 \cdot (-0,79 - 0)^2 = 2,50$	$4 \cdot (0,95 - 0)^2 = 3,61$
grün	1,42	0,05	$4 \cdot (1,42 - 0)^2 = 8,07$	$4 \cdot (0,05 - 0)^2 = 0,01$
Summe	-	-	12,31	7,78

© Lüken/Schimmelpfennig: planung & analyse 6/2015

Abbildung 2: Die durch die Diskriminanzfunktionen erklärte Streuung.

Die relative Wichtigkeit einer Diskriminanzfunktion für die Trennung zwischen den Gruppen wird durch den Anteil der durch sie erklärten Streuung an der gesamten erklärten Streuung gemessen (siehe Abbildung 2). Insofern besitzt die erste Diskriminanzfunktion eine relative Wichtigkeit von 61 Prozent und die zweite von 39 Prozent.

(Diskriminatorische) Bedeutung der Variablen

Um Vergleichbarkeit zu gewährleisten, sind die Koeffizienten zu standardisieren. Dazu werden sie multipliziert mit der Standardabweichung der entsprechenden Variable innerhalb der Gruppen. Diese beträgt für die erste Eigenschaft 0,97 und für die zweite 0,93. Somit ist beispielsweise der standardisierte Koeffizient für x_1 in der ersten Diskriminanzfunktion $0,96 \cdot 0,97 = 0,93$.

Zur Beurteilung der diskriminatorischen Bedeutung der Variablen werden alle Diskriminanzfunktionen herangezogen, um einen mittleren Diskriminanzkoeffizienten zu berechnen. Dieser ist gleich dem mit der relativen Wichtigkeit der Diskriminanzfunktionen gewichteten Mittelwert des Betrags der standardisierten Diskriminanzkoeffizienten:

$$\text{Für Eigenschaft 1: } 0,61 \cdot 0,93 + 0,39 \cdot 0,34 = 0,70$$

$$\text{Für Eigenschaft 2: } 0,61 \cdot 0,33 + 0,39 \cdot 0,95 = 0,57$$

Insofern trennt in diesem Datenbeispiel Eigenschaft 1 zwischen den drei Gruppen etwas stärker als Eigenschaft 2.

Johannes Lüken und **Prof. Dr. Heiko Schimmelpfennig**, Experten für Multivariate Analysen bei IfaD, Institut für angewandte Datenanalyse.

In Ausgabe 1/2016: *Klassifizieren mittels Diskriminanzanalyse*

▶ Literatur

Backhaus, K.; Erichson, B.; Plinke, W.; Weiber, R.: Diskriminanzanalyse. In: Multivariate Analysemethoden, 13. Auflage, Berlin, 2011, S. 187-248.

Decker, R.; Temme, T.: Diskriminanzanalyse. In: Herrmann, A.; Homburg, C.: Marktforschung, 2. Auflage, Wiesbaden, 2000, S. 295-336.