

Multivariate Tests mit optimierten Scores

Zaino M¹, Kropf S²

¹Institut für Medizinische Informatik, Statistik und Epidemiologie, Universität Leipzig, Deutschland

²Institut für Biometrie und Medizinische Informatik, Otto-von-Guericke-Universität Magdeburg, Deutschland

mohammad.zaino@imise.uni-leipzig.de

Einleitung und Fragestellung In den neunziger Jahren entwickelten Läter und Mitarbeiter eine Familie von exakten multivariaten parametrischen Tests [1-4]. Diese Tests transformieren die hochdimensionalen Daten mittels datenabhängiger Gewichte in niedrigdimensionale Scores und führen dann mit diesen Scores klassische multivariate Tests durch. Solange die Gewichte dabei nach bestimmten Bedingungen gebildet werden, halten die Tests trotz der Vorverarbeitung der Daten zu Scores das Fehlerniveau exakt ein. Auf diese Weise sind multivariate Tests selbst in dem Fall durchführbar, dass die Merkmalszahl den Gesamtstichprobenumfang übersteigt.

Die Grundlage für diese Tests bildet die Theorie der links-sphärischen Matrizen, welche die klassische Annahme unabhängiger normalverteilter Merkmalsvektoren in den traditionellen univariaten und multivariaten Tests ersetzen kann [5]. In [1] bis [4] werden Bedingungen genannt, unter denen die aus den multivariat normalverteilten Beobachtungsvektoren abgeleiteten Scores bzw. Scorevektoren in ihrer Gesamtheit eine solche links-sphärische verteilte Matrix bilden. Die in einem Schritt abgeleiteten Scores können auch Ausgangspunkt für weitere Vorverarbeitungsschritte sein ([3]).

Die so gebildeten Scores können als rein technischer Zwischenschritt im multivariaten Test betrachtet werden. Die nähere Betrachtung der Scores bietet aber auch die Chance, die Ergebnisse des globalen multivariaten Tests für den Anwender wesentlich anschaulicher zu interpretieren. Dieser Aspekt spielte in den ursprünglichen Arbeiten nur eine untergeordnete Rolle. Daher wurde jetzt der Versuch unternommen, die Scores durch Selektions- und insbesondere Rotationsmethoden zu optimieren. Das primäre Ziel war dabei, die Interpretierbarkeit der Scores zu erhöhen. Als Nebeneffekt könnte sich dabei auch die Güte für Tests mit einzelnen Scores verbessern, da latente Variable besser in den Scores repräsentiert werden.

Material und Methoden Es wird hier das Problem des Vergleichs von K unabhängigen Stichproben p -dimensional normalverteilter Daten mit gleicher, aber unbekannter Kovarianzmatrix betrachtet, wobei die Gleichheit der Erwartungswertvektoren zu prüfen ist.

In den oben genannten zitierten Arbeiten [2] bis [4] werden verschiedene Versionen von Tests erarbeitet, welche die hochdimensionalen Ausgangsvektoren \mathbf{x}_{kj} ($k = 1, \dots, K; j = 1, \dots, n_k$) zunächst in niedrigdimensionale Scorevektoren \mathbf{z}_{kj} transformieren. Alle gehen von der totalen Produktsammenmatrix \mathbf{W} aus, welche die Summe der quadratischen Abweichungen aller Stichprobenelemente vom Gesamtmittelwert aus allen Stichproben beschreibt. Aus dieser Matrix \mathbf{W} werden dann Gewichtsmatrizen \mathbf{D} vom Format $p \times q$ berechnet, mit deren Hilfe die Scores anhand der linearen Transformation $\mathbf{z}_{kj} = \mathbf{D}'\mathbf{x}_{kj}$ bestimmt werden.

Die Auswahl der Matrix \mathbf{D} unterscheidet die hier betrachteten Testvarianten. Während beim PC_q -Test ([2,3]) die Gewichtsmatrix \mathbf{D} aus den ersten q Eigenvektoren des Hauptkomponenten-Eigenwertproblems $\mathbf{W}\mathbf{D} = \text{Diag}(\mathbf{W})\mathbf{D}\mathbf{A}$ gebildet wird, erhält man sie beim PC_{est} -Test ebenfalls aus den ersten Eigenvektoren des mehr faktoranalytisch orientierten Ansatzes $\mathbf{W}\mathbf{D} = \mathbf{V}\mathbf{D}\mathbf{A}$, wobei in [4] zwei Varianten für die den Maßstab der Variablen definierende Diagonalmatrix \mathbf{V} vorgeschlagen werden: der „einfache Schätzer“ $\mathbf{V} = 1/(n-1) [\text{Diag}(\mathbf{W}^{-1}[\text{Diag}(\mathbf{W}^{-1})]^{-1} \mathbf{W}^{-1})]^{-1}$, sowie der Schätzer $\mathbf{V} = 1/(n-1) [\text{Diag}(\mathbf{W}^L)]^2$ mit der so genannten linkssymmetrischen Wurzel \mathbf{W}^L , welche über die Bedingung $\mathbf{W}^L\mathbf{W}^L = \mathbf{W}$ bei maximalem Produkt der Diagonalelemente definiert ist. Ein Algorithmus zur Berechnung von \mathbf{W}^L ist in [4] gegeben.

In [6] wird gezeigt, dass die Bestimmung der Scoredimension ebenfalls aus den Daten erfolgen kann. Dabei können bewährte Methoden aus der Faktoranalyse wie das Kaiserkriterium (q als Anzahl der Eigenwerte in den obigen Eigenwertproblemen, welche den Mittelwert aller Eigenwerte überschreiten) genutzt werden, ohne das Signifikanzniveau des Tests zu verfälschen. Konkret wurde hier neben dem Kaiser-Kriterium ein Vorschlag von Jolliffe [7] aufgegriffen, in dem er den Grenzwert des Kaiserkriteriums um 70 % absenkt und damit eine höhere Scoreanzahl ermittelt. Analog wurde eine zweite Abwandlung betrachtet, welche den Grenzwert um den gleichen Faktor erhöht („invers Jolliffe“).

Ausgehend von den so bestimmten Gewichtsmatrizen \mathbf{D} wurde jetzt in gewisser Analogie zur Faktoranalyse nach q -dimensionalen Rotationsmatrizen \mathbf{A} gesucht, welche die wie oben ermittelten Gewichtsmatrizen modifizieren, $\mathbf{D}^* = \mathbf{D}\mathbf{A}$, und dabei eine bessere Interpretation der Scores zulassen. Diese Matrizen werden wie auch die Gewichtsmatrizen \mathbf{D} nur als Funktion der Produktsammenmatrix \mathbf{W} gebildet, so dass die Theoreme ihre Gültigkeit behalten und damit auch einzelne der q Scores in einem exakten Test untersucht werden können. Dabei ist allerdings zu beachten, dass alle so konstruierten Tests die globale multivariate Nullhypothese der Gleichheit aller K Erwartungswertvektoren prüfen. Die Spezialisierung bezieht sich lediglich auf die Power. Für die Optimierung der Rotationsmatrix \mathbf{A} wird als Zielfunktion das Varimax-Kriterium verwendet. Dabei wurden verschiedene Nebenbedingungen erprobt.

Das Vorgehen kann auf andere Fragestellungen als das hier betrachtete Mehrstichprobenproblem erweitert werden, wobei jeweils die dem Problem angepasste Produktsammenmatrix \mathbf{W} zu bestimmen ist. Das weitere Vorgehen ist dann analog.

In Beispieldatensätzen und Simulationsuntersuchungen wurden die drei obigen Testvarianten mit verschiedenen Kriterien für die Bestimmung einer geeigneten Scoreanzahl q und verschiedenen Rotationsmethoden erprobt. In all diesen Rechnungen wurden Situationen betrachtet, in denen die Merkmalsanzahl fast den Gesamtstichprobenumfang erreicht. Mit dem PC_q -Test wären auch höhere Merkmalszahlen realisierbar gewesen, nicht jedoch mit dem PC_{est} -Test.

Ergebnisse In den Rechnungen zeigte sich zunächst, dass auch mit diesen hochdimensionalen Situationen Tests mit hoher Güte erreichbar sind. In den betrachteten Situationen erwiesen sich der PC_q -Test und der PC_{est} -Test mit der linkssymmetrischen Wurzel in Kombination mit dem inversen Jolliffe-Kriterium als günstige Testvarianten. In den Simulationsrechnungen wurde die vorgegebene Anzahl von latenten Variablen am besten reproduziert. Im Beispiel erzielte man die am besten interpretierbare Korrelationsstruktur zwischen den Scores und den Originalvariablen (Ladungsmatrizen). Das Kaiser-Kriterium und erst recht das Jolliffe-Kriterium lieferten tendenziell zu hohe Faktorzahlen. Die Tests mit einzelnen Scores zeigten nach der Rotation klarere Aussagen als vor der Rotation, insbesondere wenn die Nebenbedingungen so gewählt wurden, dass die resultierenden Scores unkorreliert sind.

Diskussion Eine Rotation im Faktorraum kann die Interpretierbarkeit von Testergebnissen von scorebasierten multivariaten Tests wesentlich erhöhen und liefert dabei sogar Möglichkeiten, die Güte des Tests zu erhöhen. Bei der Vielfalt der Testmöglichkeiten und aller Modifikationen ist es jedoch wichtig, eine genaue Festlegung der geplanten Teststrategie vorab vorzunehmen oder aber zu fixieren, wie sie anhand der Ladungsmatrizen (jedoch nicht der Testergebnisse!) noch im Zuge der Analyse festgelegt wird.

Literatur

- [1] Läter J. Exact t and F Tests for Analyzing Studies with Multiple Endpoints. *Biometrics* 1996; 52: 964-970.
- [2] Läter J, Glimm E, Kropf S. New Multivariate Tests for Data with an Inherent Structure. *Biometrical Journal* 1996; 38: 5-23.
- [3] Läter J, Glimm E, Kropf S. Multivariate Tests Based on Left-Spherically Distributed Linear Scores. *Annals of Statistics* 1998; 26: 1972-1988.
- [4] Läter J, Kropf S, Glimm E. Exact Stable Multivariate Tests for Applications in Clinical Research. In: 1998 Proceedings of the Biopharmaceutical Section, Annual Meeting of the American Statistical Association, Dallas, Texas, August 9-13, 1998. American Statistical Association; 1999: 46-55.
- [5] Fang KT, Zhang YT. Generalized Multivariate Analysis. Berlin Heidelberg: Springer; 1990.
- [6] Kropf S. Hochdimensionale multivariate Verfahren in der medizinischen Statistik. Aachen: Shaker Verlag; 2000.
- [7] Jolliffe IT. Discarding Variables in a Principal Component Analysis. I: Artificial Data. *Applied Statistics* 1992; 21: 160-173.