DISCUSSION PAPERS IN STATISTICS AND ECONOMETRICS

SEMINAR OF ECONOMIC AND SOCIAL STATISTICS
UNIVERSITY OF COLOGNE

No. 6/07

Anmerkungen zur Aggregation von Intelligenzquotienten

by

Gabriel Frahm
Gert Mittring

4th version
July 31, 2007

DISKUSSIONSBEITRÄGE ZUR STATISTIK UND ÖKONOMETRIE

SEMINAR FÜR WIRTSCHAFTS- UND SOZIALSTATISTIK
UNIVERSITÄT ZU KÖLN

Albertus-Magnus-Platz, D-50923 Köln, Deutschland

This page intentionally left blank.

DISCUSSION PAPERS IN STATISTICS AND ECONOMETRICS

SEMINAR OF ECONOMIC AND SOCIAL STATISTICS
UNIVERSITY OF COLOGNE

No. 6/07

Anmerkungen zur Aggregation von Intelligenzquotienten

by

Gabriel Frahm[1]
Gert Mittring[2]

4th version
July 31, 2007

Zusammenfassung

Bei der Messung der Intelligenz eines bestimmten Probanden liegen typischerweise unterschiedliche Testergebnisse vor und der untersuchende Psychologe möchte die vorliegenden Messwerte im Kontext der Ergebnisse anderer Probanden bewerten. Dabei will er das Potenzial seines Probanden weder überschätzen noch unterschätzen. Es stellt sich jedoch heraus, dass der arithmetische Mittelwert aller Messwerte des Probanden im Allgemeinen zu einer *verzerrten* Einschätzung seines Potenzials führt. In der vorliegenden Arbeit wird eine Methode entwickelt, die eine konsistente Aggregation unterschiedlicher Intelligenzquotienten zu einem „Meta-IQ“ ermöglicht. Darüber hinaus werden einfache Aggregationsformeln unter Berücksichtigung bestimmter Abhängigkeitsstrukturen zwischen den einzelnen Testergebnissen eines Probanden vorgestellt.

Schlüsselworte: Aggregation, Faktormodell, Hauptkomponentenmodell, Intelligenzmessung, Korrelation von Testergebnissen, „Meta-IQ“, statistische Verfahren.

[1]Universität zu Köln, Lehrstuhl für Statistik und Ökonometrie, Meister-Ekkehart-Str. 9, 50937 Köln.

[2]Intertel Research Officer, Mensa Österreich, Von-Sandt-Str. 32, 53225 Bonn.

ANMERKUNGEN ZUR AGGREGATION VON INTELLIGENZQUOTIENTEN

GABRIEL FRAHM†
UNIVERSITÄT ZU KÖLN
LEHRSTUHL FÜR STATISTIK UND ÖKONOMETRIE
MEISTER-EKKEHART-STR. 9
50937 KÖLN

GERT MITTRING‡
INTERTEL RESEARCH OFFICER
MENSA ÖSTERREICH
VON-SANDT-STR. 32
53225 BONN

Zusammenfassung. Bei der Messung der Intelligenz eines bestimmten Probanden liegen typischerweise unterschiedliche Testergebnisse vor und der untersuchende Psychologe möchte die vorliegenden Messwerte im Kontext der Ergebnisse anderer Probanden bewerten. Dabei will er das Potenzial seines Probanden weder überschätzen noch unterschätzen. Es stellt sich jedoch heraus, dass der arithmetische Mittelwert aller Messwerte des Probanden im Allgemeinen zu einer *verzerrten* Einschätzung seines Potenzials führt. In der vorliegenden Arbeit wird eine Methode entwickelt, die eine konsistente Aggregation unterschiedlicher Intelligenzquotienten zu einem „Meta-IQ" ermöglicht. Darüber hinaus werden einfache Aggregationsformeln unter Berücksichtigung bestimmter Abhängigkeitsstrukturen zwischen den einzelnen Testergebnissen eines Probanden vorgestellt.

Notes on the Aggregation of Intelligence Quotients

Abstract. The brainpower of some person is typically evaluated by applying several intelligence tests. Thereupon the psychologist has to interpret the results relative to the outcomes of other test persons. He aims at a fair judgement without overestimating or underestimating the capabilities of the test person. Unfortunately, it turns out that the mean of all measurements of the person generally leads to a *biased* estimate. Thus we present a method for aggregating different intelligence quotients which leads to an unbiased estimate, the so-called 'Meta-IQ'. Furthermore, we derive simple formulas by taking specific dependence structures of the outcomes into account.

Schlüsselworte: Aggregation, Faktormodell, Hochbegabtendiagnostik, Intelligenzmessung, Korrelation, „Meta-IQ", statistische Verfahren.

Key words: aggregation, correlation, factor model, gifted people, intelligence, 'Meta-IQ', statistical methods.

Ausgangslage

Bei der Untersuchung des intellektuellen Potenzials eines Probanden werden häufig mehrere Intelligenztests eingesetzt. Dabei kann es zu recht unterschiedlichen Messwerten kommen, und der untersuchende Psychologe wird vor die Frage gestellt, welcher Intelligenzquotient (IQ) nun „richtig" sei bzw. wie das Untersuchungsergebnis insgesamt zu bewerten sei, ohne das Potenzial des Probanden zu überschätzen oder es zu unterschätzen. Insbesondere, wenn sprachgebundene und sprachfreie Verfahren kombiniert zum Einsatz kommen, wird das Problem akut. Eine einigermaßen korrekte Antwort ist dann umso wichtiger, wenn von dem Ergebnis einschneidende

Maßnahmen für den Probanden abhängen, wie beispielsweise die Wahl der weiterführenden Schulgattung, die Vergabe eines Stipendiums oder eine berufliche Zusatzqualifikation.

Der vorliegende Aufsatz beruht auf der Erkenntnis, dass der arithmetische Mittelwert aller Messwerte eines Probanden im Allgemeinen zu einer *verzerrten* Einschätzung seines Potenzials führt. Im Folgenden wird eine Methode entwickelt, um eine konsistente Aggregation von unterschiedlichen Intelligenzquotienten (IQs) quasi zu einem „Meta-IQ", wie die beiden Autoren ihn nennen wollen, zu ermöglichen. Aus einer allgemein gültigen Formel werden danach einfache Formeln unter Berücksichtigung bestimmter Abhängigkeitsstrukturen zwischen den einzelnen Testergebnissen eines Probanden abgeleitet.

1. Grundlegende Annahmen und Notation

Der *Score* S (d.h. die Punktzahl) eines beliebigen Intelligenztests besitze die stetige Verteilungsfunktion F. Dann kann S mittels der Transformation

$$Z := \Phi^{-1}\big\{F(S)\big\}$$

zunächst in die standardnormalverteilte Zufallsvariable Z überführt werden. Das Symbol Φ^{-1} steht hierbei für die Quantilfunktion der Standardnormalverteilung. Durch eine weitere Transformation

$$X := 100 + 15Z$$

erhält man einen *normierten* Intelligenztest, d.h. für den IQ X einer beliebigen Testperson gilt $X \sim \mathcal{N}(100, 15^2)$. In praxi wird die Verteilungsfunktion F gemeinhin durch die empirische Verteilungsfunktion von S geschätzt. Auf diesen Punkt werden wir nicht näher eingehen, möchten allerdings hervorheben, dass die folgenden Überlegungen nicht von diesem Schätzproblem betroffen sind.

Nun betrachte man zwei normierte Intelligenztests X_1 und X_2, d.h.

$$X_1, X_2 \sim \mathcal{N}\big(100, 15^2\big).$$

Wir treffen nun die zusätzliche Annahme, dass X_1 und X_2 *gemeinsam* normalverteilt sind, d.h.

$$\big(X_1, X_2\big) \sim \mathcal{N}\left(\begin{bmatrix}100\\100\end{bmatrix}, \begin{bmatrix}15^2 & 15^2\rho_{12}\\ 15^2\rho_{21} & 15^2\end{bmatrix}\right),$$

wobei $-1 < \rho_{12} < 1$ der Korrelationskoeffizient von X_1 und X_2 ist. Dieser quantifiziert das Ausmaß des linearen Zusammenhangs zwischen X_1 und X_2 .

Die Annahme einer gemeinsamen Normalverteilung ist nicht ganz unproblematisch. Durch eine geeignete Transformation kann die Randverteilung eines Scores prinzipiell in jede andere Verteilung überführt werden. Allerdings ist die gemeinsame Abhängigkeitsstruktur, also die sogenannte *Copula* (Nelsen, 2006, S. 18) eines Zufallsvektors, invariant gegenüber streng monoton steigenden Transformationen seiner Komponenten. Somit müssen wir uns zwecks folgender Überlegungen auf eine bestimmte Copula beschränken und der Einfachheit halber gehen wir von der *Gauss-Copula*, d.h. der gemeinsamen Abhängigkeitsstruktur eines normalverteilten Zufallsvektors, aus.

Im Folgenden werden wir auch allgemeiner einen normalverteilten Zufallsvektor $X = (X_1, \ldots, X_m)$ betrachten, d.h.

$$X \sim \mathcal{N}\left(\begin{bmatrix}100\\100\\ \vdots\\100\end{bmatrix}, \begin{bmatrix}15^2 & 15^2\rho_{12} & \ldots & 15^2\rho_{1m}\\ 15^2\rho_{21} & \ddots & & \vdots\\ \vdots & & \ddots & \vdots\\ 15^2\rho_{m1} & \ldots & \ldots & 15^2\end{bmatrix}\right)$$

stellt dann einen multivariaten Intelligenztest bestehend aus m Komponenten dar. Mit einer etwas kompakteren Notation erhalten wir

$$X \sim \mathcal{N}\big(100 \cdot 1, 15^2\rho\big), \tag{1.0.1}$$

wobei $1 := (1, 1, \ldots, 1)$ ein Einsvektor sowie

$$\rho := \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1m} \\ \rho_{21} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \rho_{m1} & \cdots & \cdots & 1 \end{bmatrix}$$

eine positiv definite Korrelationsmatrix ist.

2. Aggregation von Intelligenzquotienten

2.1. **Bivariater Fall.** Ein Meta-IQ $Y = f(X_1, X_2)$ muss nun ebenfalls die Eigenschaft

$$Y = f\big(X_1, X_2\big) \sim \mathcal{N}\big(100, 15^2\big) \tag{2.1.1}$$

erfüllen. Der arithmetische Mittelwert erfüllt diese Eigenschaft im Allgemeinen jedoch nicht, denn

$$\overline{X} := \frac{X_1 + X_2}{2} \sim \mathcal{N}\bigg(100, \frac{1+\rho}{2} \cdot 15^2\bigg).$$

Lediglich für den realitätsfremden Grenzfall $\rho = 1$ (d.h. die beiden Intelligenztests sind praktisch identisch) wäre die gewünschte Eigenschaft (2.1.1) erfüllt. Für $-1 < \rho < 1$ kommt es jedoch bezüglich aller überdurchschnittlich intelligenten Probanden zu einer chronischen *Unterschätzung* ihrer Potenziale und umgekehrt. Falls die einzelnen Tests nämlich nur schwach korreliert sind, ist es relativ unwahrscheinlich, dass ein Proband in beiden Tests überdurchschnittlich gut abschneidet.

Dieser unerwünschte Effekt kann leicht behoben werden. Statt des arithmetischen Mittelwerts erfüllt nämlich die aggregierte Größe

$$Y := 100 + \frac{\overline{X} - 100}{\sqrt{(1+\rho)/2}} \sim \mathcal{N}\big(100, 15^2\big) \tag{2.1.2}$$

die gewünschte Eigenschaft. Die richtige Aggregation hängt also maßgeblich von der Korrelation der Ergebnisse beider Intelligenztests ab.

2.2. **Multivariater Fall.** Nun seien m Intelligenztests und die dazugehörigen IQs $X_1, \ldots, X_m$ gegeben. Hierbei können wir nun auf die kompakte Schreibweise aus Gl. 1.0.1 zurückgreifen. Bei der Aggregation der m Testergebnisse mittels einer Aggregationsfunktion f sollte nun grundsätzlich die Eigenschaft

$$Y = f\big(X_1, \ldots, X_m\big) \sim \mathcal{N}\big(100, 15^2\big) \tag{2.2.1}$$

erfüllt sein. Für den arithmetischen Mittelwert gilt in diesem Fall

$$\overline{X} = \frac{1}{m} \cdot \sum_{t=1}^{m} X_t \sim \mathcal{N}\bigg(100, \frac{1'\rho 1}{m^2} \cdot 15^2\bigg)$$

und die Forderung (2.2.1) ist wiederum verletzt. Hat man also m Intelligenztests ausgewertet, so ist der arithmetische Mittelwert aller IQs für die Beurteilung eines Probanden ungeeignet. Allerdings kann man auch hier auf die aggregierte Größe

$$Y = 100 + m \cdot \frac{\overline{X} - 100}{\sqrt{1'\rho 1}} \sim \mathcal{N}\big(100, 15^2\big) \tag{2.2.2}$$

ausweichen, wobei $1'\rho 1 = \sum_{i,j=1}^{m} \rho_{ij}$. Man beachte, dass der Meta-IQ in Gleichung 2.1.2 gerade aus Gl. 2.2.2 im Spezialfall $m = 2$ resultiert.

Angenommen, man möchte nun die einzelnen Testergebnisse unterschiedlich gewichten. Das bedeutet, es existiert ein Wägungsschema $w = (w_1, \ldots, w_m)$ mit der Eigenschaft $1'w = 1$ und man möchte aus allen Testergebnissen einen Meta-IQ berechnen. Man erhält dann

$$Y = 100 + \frac{\sum_{t=1}^{m} w_t X_t - 100}{\sqrt{w'\rho w}} \sim \mathcal{N}\big(100, 15^2\big) \tag{2.2.3}$$

mit $w'\rho w = \sum_{i,j=1}^{m} w_i w_j \rho_{ij}$. Im Übrigen erhält man im Fall einer Gleichgewichtung, d.h. $w = (1/m, \ldots, 1/m)$, die Formel aus Gl. 2.2.2. Der Ausdruck in (2.2.3) kann somit als allgemeine Aggregationsformel verwendet werden.

3. Spezielle Aggregationsformeln

3.1. **Bestimmung der Korrelationskoeffizienten.** Grundsätzlich stellt sich die Frage, wie man die Korrelationsmatrix ρ ermitteln soll. Mittels einer empirischen Studie könnte man ρ_{ij} $(i, j = 1, \ldots, m)$ beispielsweise anhand des Pearsonschen Korrelationskoeffizienten

$$\hat{\rho}_{ij} = \frac{\sum_{k=1}^{n} (x_{ik} - 100)(x_{jk} - 100)}{\sqrt{\sum_{k=1}^{n} (x_{ik} - 100)^2} \sqrt{\sum_{k=1}^{n} (x_{jk} - 100)^2}}$$

schätzen und die Einträge in ρ durch die entsprechenden Schätzwerte ersetzen. Hierbei spiegeln $x_{i1}, \ldots, x_{in}$ und $x_{j1}, \ldots, x_{jn}$ die empirisch ermittelten IQs von n Probanden bezüglich der Tests i und j wider. In vielen Fällen wird eine geeignete empirische Studie allerdings nicht vorhanden bzw. eine entsprechende Auswertung aufgrund fehlender Daten unmöglich sein.

3.2. **Allgemeine Intelligenz.** Man geht im Allgemeinen davon aus, dass ein positiver linearer Zusammenhang zwischen den Ergebnissen einzelner Intelligenztests besteht. Ohne Kenntnis von ρ kann man der Einfachheit halber eine *Äquikorrelationsstruktur* unterstellen, d.h. $\rho_{ij} = r$ für $i, j = 1, \ldots, m$ $(i \neq j)$ mit $0 < r < 1$. Etwas allgemeiner kann man auch von einem *allgemeinen Faktor der Intelligenz* ausgehen und gewisse Annahmen hinsichtlich der *Faktorladungen* einzelner Tests treffen. Beide Ansätze führen zu recht einfachen Aggregationsformeln.

3.2.1. *Hauptkomponentenmodell.* Im Falle der Äquikorrelation ergibt sich für die Korrelationsmatrix gerade $\rho = r11' + (1 - r)I_m$, wobei I_m die m-dimensionale Einheitsmatrix darstellt. Die Matrix ρ besitzt dann den maximalen Eigenwert $\lambda = rm + (1 - r)$. Das bedeutet nach Anwendung einer *Hauptkomponentenanalyse* erhalten wir für die größte Hauptkomponente ein Anteil von

$$\gamma := \frac{\lambda}{m} = r + \frac{1 - r}{m} \geq \frac{1}{m}$$

an der Gesamtstreuung aller IQs. Der Parameter γ lässt sich in aller Regel subjektiv besser abschätzen, als der Korrelationskoeffizient r. Wenn man z.B. davon ausgeht, dass $\gamma = 50\%$ der Gesamtstreuung aller IQs durch die größte Hauptkomponente erklärt werden und man verwendet $m = 4$ Intelligenztests, so erhält man

$$r = \frac{\gamma m - 1}{m - 1} = \frac{0.5 \cdot 4 - 1}{4 - 1} = 0.\bar{3}$$

als entsprechenden Korrelationskoeffizienten. Für die Berechnung des Meta-IQ gemäß Gl. 2.2.2 benötigt man sogar lediglich den Anteil γ, d.h. sowohl der Korrelationskoeffizient r als auch die Anzahl m der Intelligenztests werden nicht explizit benötigt. Es gilt nämlich

$$1'\rho 1 = m\left\{rm + (1 - r)\right\} = m\lambda = m^2\gamma\,.$$

Damit erhält man für den Meta-IQ aus Gl. 2.2.2 die einfache Darstellung

$$Y = 100 + \frac{\overline{X} - 100}{\sqrt{\gamma}} \sim \mathcal{N}\big(100, 15^2\big)$$

und man beachte, dass $1/m \leq \gamma < 1$. Wie wir nun sehen können, hängt der Meta-IQ eines Probanden maßgeblich vom Einfluss der größten Hauptkomponente ab. Es sei darauf hingewiesen, dass beim Hauptkomponentenmodell *nicht* angenommen wird, dass die Ergebnisse der einzelnen Intelligenztests nach Abzug der größten Hauptkomponente unkorreliert sind.

3.2.2. *Faktormodell.* Das auf Spearman (1904) zurückgehende Faktormodell beinhaltet die Korrelationsstruktur $\rho = gg' + \Delta$, wobei der Vektor g den Einfluss des *allgemeinen Faktors* bezüglich jedes einzelnen Tests quantifiziert ($g_1, \ldots, g_m$ sind also die sogenannten *Faktorladungen*) und Δ eine Diagonalmatrix ist, welche die Residualstreuung enthält. Wenn man z.B. davon ausgeht, dass das Ergebnis des ersten Intelligenztests zu 80%, des zweiten zu 60%, u.s.w. durch den allgemeinen Faktor bestimmt wird, erhält man für die Faktorladungen die Werte $g_1 = \sqrt{0.8}$, $g_2 = \sqrt{0.6}$, etc. Beim Faktormodell wird nun explizit angenommen, dass die Ergebnisse einzelner Intelligenztests nach Abzug des allgemeinen Faktors *unkorreliert* sind, d.h. testspezifische mentale Fähigkeiten zum Ausdruck bringen. Man erhält nun $1'\rho 1 = (1'g)^2 + 1'\delta$, wobei $\delta := \operatorname{diag}(\Delta)$ und es gilt $1'\delta = m - g'g$. Daraus ergibt sich

$$1'\rho 1 = (1'g)^2 + m - g'g$$

und bezüglich Gl. 2.2.2 erhalten wir

$$Y = 100 + \frac{\overline{X} - 100}{\sqrt{\bar{g}^2 + (1/m - \|g/m\|^2)}} \sim \mathcal{N}\big(100, 15^2\big)\,, \tag{3.2.1}$$

wobei $\bar{g} := 1'g/m$ und $\|g/m\| := \sqrt{(g/m)'(g/m)}$. Falls die Testergebnisse vollkommen unkorreliert sind (d.h. $g = 0$), erhalten wir für den Nenner in Gl. 3.2.1 gerade den Wert $\sqrt{1/m}$. Im Grenzfall perfekter positiver Korrelation (d.h. $g = 1$) ergibt sich stattdessen der Wert 1 und der Meta-IQ entspricht dem arithmetischen Mittelwert aller Testergebnisse eines Probanden. Im speziellen Fall der Äquikorrelation (d.h. $g = \sqrt{r}\,1$) gilt

$$1'\rho 1 = rm^2 + (1-r)m$$

und Gl. 3.2.1 vereinfacht sich zu

$$Y = 100 + \frac{\overline{X} - 100}{\sqrt{r + (1-r)/m}} \sim \mathcal{N}\big(100, 15^2\big)\,. \tag{3.2.2}$$

Wenn man also z.B. glaubt, dass der allgemeine Faktor zu 50% in *jedes* der m Testergebnisse einfließt (und die Tests außer dem allgemeinen Faktor lediglich *spezifische* mentale Fähigkeiten messen), setzt man $r = 0.5$ in Gl. 3.2.2 ein. Diese Aggregationsformel verdeutlicht, dass eine Korrektur des arithmetischen Mittelwerts der Testergebnisse eines Probanden insbesondere dann notwendig ist, wenn

(1) die einzelnen Tests zu einem großen Anteil spezifische mentale Fähigkeiten messen und/oder
(2) der zu untersuchende Proband hochbegabt ist (da sich die Korrektur umso stärker auswirkt, je größer $\overline{X} - 100$ ist).

4. Fazit

Der arithmetische Mittelwert aller IQs eines Probanden führt zu einer verzerrten Einschätzung seiner mentalen Fähigkeiten. Unter der Annahme einer gemeinsamen Normalverteilung der IQs können jedoch allgemeine und konsistente Aggregationsformeln hergeleitet werden. Wird eine bestimmte Korrelationsstruktur hinsichtlich der einzelnen Testergebnisse vorausgesetzt, können daraus besonders einfache Rechenregeln zwecks einer unverzerrten Beurteilung des Meta-IQ eines Probanden abgeleitet werden. Für die Anwendung dieser Rechenregeln genügt im Allgemeinen eine grobe Abschätzung der wahren Korrelationsstruktur. Eine konsistente Aggregation unterschiedlicher Testergebnisse wird insbesondere in der Hochbegabtendiagnostik empfohlen.

Danksagung

Die Autoren möchten sich herzlich bei Frau Dr. Ida Fleiß für ihre wertvollen Anregungen und konstruktive Kritik als auch bei Carsten Körner für seinen Hinweis auf einen Tippfehler bedanken.

Literatur

Nelsen, R. (2006): An Introduction to Copulas. Springer, zweite Aufl.

Spearman, C. (1904): 'General intelligence,' objectively determined and measured. *The American Journal of Psychology* **15**, S. 201–292.

†Universität zu Köln, Lehrstuhl für Statistik und Ökonometrie, Meister-Ekkehart-Str. 9, 50937 Köln. E-Mail: frahm@statistik.uni-koeln.de, Tel.: 0221 470-4267.

‡Intertel Research Officer, Mensa Österreich, Von-Sandt-Str. 32, 53225 Bonn. E-Mail: mittring.fleiss@tiscali.de, Tel.: 0228 460500.