# Vergleichende Analyse zielgerichteter und kontextfreier Verdachtsgenerierungsverfahren

Pitlik L., 2011. XI.

## Einführung

Es gab zwei unterschiedlichen Annäherungen zum Problem „Verdachtsgenerierung“:

* <http://miau.gau.hu/miau/156/bericht_frageboegen.docx>
* <http://miau.gau.hu/miau/158/la158.docx>

Die erste Version nahm an, dass Fragebögen in einer Firma immer mit konkreten Zielformulierungen verbunden sind. Daher scheint es gerecht zu sein, Verdacht kontextabhängig abzuleiten. In diesem Fall sollte man für jede Frage die Ideale Antwort festlegen können und jede weitere Antwortvariante als eine Art Rangzahl (an der idealen Antwort gemessen) ausdrücken können. Demnach entsteht der Verdacht dann: wenn die unerklärbaren Abweichungen pro Person und Frage überdurchschnittlich sind, wobei die Unerklärbarkeit mit inversen Modellen und/oder Abweichungen zwischen Antidiskriminierungsmodellen und deren Zerlegungsberechnungen begründet wird.

Die zweite Version wurde für solche Fragebögen ausgearbeitet, in denen keine Konzeptionen darüber vorliegen, ob es überhaupt eine Art ideale Antwort pro Frage geben darf. In diesem Fall werden pro Frage explorative Modelle gebaut und zu deren Kontrolle konsolidierte Treppenfunktionen abgeleitet. De Verdacht entsteht dann, wenn überdurchschnittlich hohe Abweichungen zwischen den tatsächlichen und den normierten (erwarteten) Antworten existieren (in den Modellen, die stabil genug sind). Instabile Modelle werden durch Antidiskriminierungsberechnungen aus Modellkorrelationen, -abweichungen und Konsolidierungseffekten (wie Streuungen, MAX- und MIN-Abweichungen, Parallelitätsverhältnissen von Teilergebnissen) identifiziert.

## Ergebnisse

Bei vergleichenden Analysen stellen sich immer folgende Fragen:

* welche Unterschiede gibt es bezüglich der verdächtigen Objekte?
* welche Verdachtsgenerierung ist plausibler für menschliche Intuitionen?
* welche Vorgehensweise brauchte weniger Rechenschnitte, Rechenzeit?

Abbildung 1 und 2 weisen darauf hin, dass

* die letzte Frage (X33) immer verdächtig ist, was hauptsächlich dadurch zu erklären ist, dass die letzte Frage nicht mehr ernst genug genommen wird.
* die ersten (5) Fragen verdächtiger sind, als eine andere, gleich große, zufällig gewählte Fragengruppen, weil hier noch eine Art gedanklichen Synchronisationseffekte zu vermuten sind.
* Verdacht und Stabilität laufen relativ überzeigend parallel.

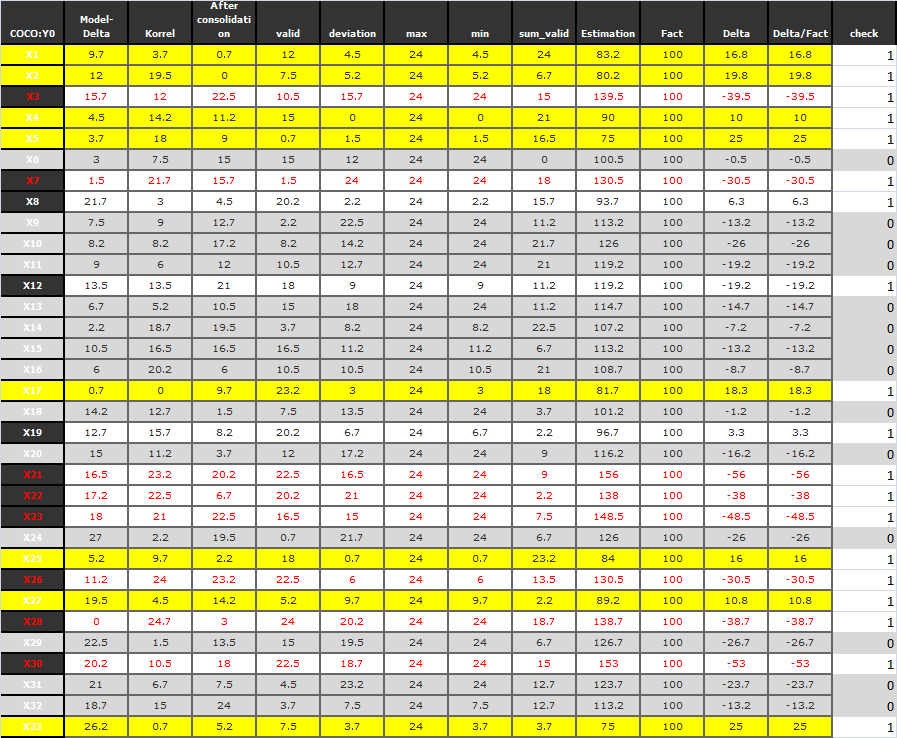


Abbildung 1: Modellierbarkeit der Fragen - kontextfrei (Quelle: eigene Berechnungen)

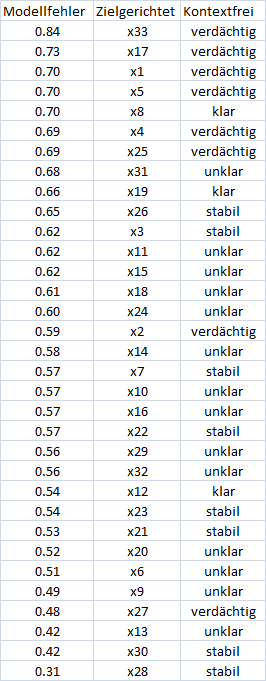
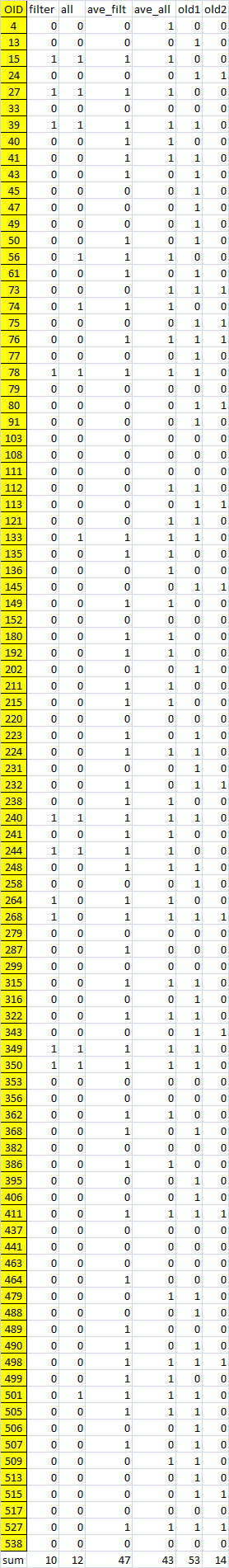


Abbildung 2: Zielgerichtete und kontextfrei Modellschwächen (Quelle: eigene Berechnungen)

Abbildung 3: Verdachtsmomente (Quelle: eigene Berechnungen)

Wie Abbildung 3 zeigt, es gibt mehrere Möglichkeiten Verdacht zu generieren. Die gefilterte Variante wurde ohne solche Fragen abgeleitet, die als Modell instabil sind und die durchschnittlichen Abweichungen über der Trendlinie liegen. Die ALL-Variante hat alle Fragen zur Berechnung mit einbezogen (und die durchschnittlichen Abweichungen über der Trendlinie liegen). Die AVERAGE-Varianten haben sowohl eine gefilterte und eine volle Version und hier wurde Verdacht für jede Person mit überdurchschnittlicher Abweichung vom Normwert pro Frage und Person deklariert. Die Spalte OLD1 zeigt die Verdachtsmomente laut Unerklärbarkeitseffekte und OLD2 steht hier für unstabile Teilmodelle (vgl. automatische SWOT-Analyse) anhand der Antidiskriminierungsberechnungen.

Wie man sieht: es gibt Personen, die aus vielen Perspektiven Verdächtig sind (d.h. beinah überall steht der Wert 1). Bis auf die zielgerichteten, SWOT-basierten Berechnungen kann man behaupten, dass Verdacht eine Art universelles Phänomen sei. Die verdächtigen Gruppen stimmen etwa 50% miteinander überein (trotz vollkommen unterschiedlicher Annäherungen).

Die Wirkung der Verdachtsmomente pro Person lässt sich selbstverständlich in den Durchschnittswerten pro Frage entdecken. Anhängig davon wie breit der Begriff Verdacht verstanden werden soll, d.h. wie viele (und welche) Verdächtige eine selbständige Gruppe bilden, die Fragendurchschnitte können sich gravierend ändern – und das abzuleiten, war eigentlich das Hauptziel der ganzen Berechnungsserie!

Die Abbildung 3 ist also eine Art Leitfaden. Die unterschiedlichen Verdachtspalten können als Filterimpulse verwendet werden. Wenn die weiteren Spalten die tatsächlichen Antworten sind, dann lassen sich Teilmengendurchschnitte pro Frage mit wenigen Mausklicken festlegen.

## Zusammenfassung

Das Ziel wurde also erreicht: es gibt fundierte und automatisierbare Vorgehensweisen kontextfrei und/oder kontextabhängig Verdachtsmomente für Personen und Fragen abzuleiten. Somit sollte man leider die bisherige Praxis, alle Antworten ohne weiteres in eine Durchschnittsberechnung mit einfließen zu lassen, zunächst als unnötiges Risiko beurteilen. Die gesamte Markt der Befragungen ließe sich also revolutionieren, falls die standardisierte Auswertungssoftware mit Inkonsistenz/Verdacht-Kontrollmechanismen ergänzt wären. Die Berechnungsprozesse stehen dafür zur Verfügung. Was noch fehlt, ist das plattformfreie Abschließen der Automatisierungsschritte…

## Anhang

Unterschiede der Durchschnitte pro Frage für ausgewählte Teilmenge (Gruppen von Befragten):…