

Kow KURODA, Hikaru YOKONO, Keiga ABE, Tomoyuki
TUTIYA, Yoshihiki ASAOKA, Yuichiro KOBAYASHI, Toshiyuki
KANAMARU, & Takumi TAGAWA
Annual Meeting of NLP
2019-03-12, Nagoya Univ

Insights from a large scale web survey for Acceptability Rating Data for Japanese (ARDJ)

本日の話題提供の概要

- 日本語の容認度評定データ
(Acceptability Rating Data of Japanese: ARDJ) プロジェクトの簡単な紹介
 - 調査の概要
 - 得られた容認度評定の概要と分類
 - 容認度評定のモデル化 (2D)
- 共同研究者: 阿部 慶賀
- 研究協力者: 浅尾 仁彦, 金丸敏幸, 小林 雄一郎, 田川 拓海, 土屋 智行, 横野 光
 - 考察と議論
 - おまけ
 - 標準化なしのHCPC
- 外部協力者: 粟津 俊二, 寺井 あすか, 山泉 実
 - 容認度評定のモデル化 (3D)
- Survey 2, Phase 2 (大規模 web 調査)

ARDJ s2u データ公開サイト

- 282種類の刺激文に対する述べ1,882名による容認度の4値評定
- ファイルは無償で入手可能ですが
- 中身へのアクセスにはパスワードが必要です
 - 利用をご希望の方は私にお問い合わせを

The screenshot shows a GitHub repository page for 'kow-k / Acceptability-Rating-Data-of-Japanese'. The repository has 0 issues, 0 pull requests, 0 projects, and 0 wiki pages. The 'Code' tab is selected. The 'data' folder is expanded, showing a list of files and their status:

File	Status
outdated	moved files
README.md	edited a file
data-mapped-classified-unified.xlsx.zip	renamed files
responses-for-set1-grouped.xlsx.zip	renamed files
responses-for-set1-ungrouped.xlsx.zip	renamed files
s1.grN.refined1.tsv.zip	added a file
s2u.grN.refined1b.collapsed.placed.tsv.zip	updated files
s2u.grN.refined1b.placed.tsv.zip	updated files

Below the file list, there is a section titled 'README.md' containing the following text:

Data from Acceptability Rating Database for Japanese
(ARDJ)

なぜ（わざわざ）ARDJ なのか？



ARDJ プロジェクトの経緯 1/2

- 科研費 (2016-2018年度)
 - 研究課題名: 言語研究者の容認度評定力の認証システムの試作: 容認度評定データベースを基礎にして(挑戦的萌芽研究 16K13223)
 - <https://github.com/kow-k/Acceptability-Rating-Data-of-Japanese>
 - 2016年度
 - 色々と調査 [4]
 - 2017年度
 - 逸脱文候補を自動生成するスクリプトの開発
 - <https://github.com/kow-k/Japanese-sentence-mutators>
 - 予備実験1を実施
 - 2018年度
 - 本実験1の第1部
 - 大学生を対象にした小規模調査 (200名程度)
 - 本実験1の第2部
 - 一般の方々を対象にした大規模調査 (1600+名規模)
 - を実施

ARDJ プロジェクトの経緯 2/2

- [1] 黒田 航 (2016). 心理学的により現実的な容認度評定のモデルを求めて. JCLA17ワークショッピ発表.
- [2] 黒田 航, 阿部 慶賀, 横野 光, 田川 拓海, 小林 雄一郎, 金丸 敏幸, 土屋 智行, 浅尾 仁彦 (2016). (言語学者による) 容認度評定の認証システムを試作する構想: 入念に設計された日本語文の容認度評定データベースに基づいて. 日本認知科学会第33回大会発表論文集, pp. 557-562.
- [3] Kuroda, Kow (2016). In search of a psychologically realistic model of acceptability judgements: Towards a better understanding of how "context(ual) effects" work. 「思考と言語」研究会発表.
- [4] 黒田 航, 浅尾 仁彦, 金丸 敏幸, 小林 雄一郎, 田川 拓海, 横野 光, 土屋 智行, 阿部 慶賀 (2017). 言語学は事例をどう扱っているのか? 見本抽出から明らかになった扱い方の(意外な)片寄り. 言語処理学会23回大会発表論文集, pp. 458-461, 2017.
- [5] Kuroda, K., Hikaru, Y., Abe, K., Tsuchiya, T., Asao, Y., Kobayashi, Y., Kanamaru, T., and Tagawa, T. (2018). Development of Acceptability Rating Data for Japanese (ARDJ): An Initial report. In Proc. of the Annual Meeting of Natural Language Processing Society, pp. 65-68, 2018.
- [6] 黒田 航 (2018). 意味の社会性を意識した動詞の分類とその理論的含意. 日本認知科学会第35回大会発表論文集, pp. 65-68, 2018.

なぜ ARDJ か？ 1/3

- 私が ARDJ を始めた動機
- (理論) 言語学者は、容認性判断 **acceptability judgment** (あるいは容認度評定 **acceptability rating**) の内実が自明なものとして扱っているが、事実無根
- そもそも、理論から独立した測定が行なわれていない
 - 生成言語学は、容認可能な文と容認不可能な文の区別が自明で、“容認可能性=文法性”の同一視が自明だとする
 - 認知言語学は、容認可能な文と容認不可能な文の区別は程度の問題なので、二つの区別が説明に値しないとする

なぜ ARDJ か？ 2/3

- 何が問題か？
- 理論言語学の観測観/データ観は素朴過ぎる
- 観測には誤差やバイアスがつきものだって事が、まったく自覚されていない
- 理論言語学の研究では確証バイアス (*confirmation bias*) の悪影響が甚大
- ベースライン反応を確認せんかったら、専門家の判断があっても評価しようがない

なぜ ARDJ か？ 3/3

- 容認性判断の実態調査が必要
 - ただし容認性判断は次の意味で容認度評定の特殊な場合だと考える
 - 容認性判断: 容認できる \Leftrightarrow できないの二値反応
 - 容認度評定: 完全に容認できる $\Leftrightarrow \dots \Leftrightarrow$ 完全にできない間の多値反応

容認度評定課題の例

- 次に30種類の文をお見せします。個々の文を次の4つの基準で評定して下さい。
 - 0. 違和感がなく自然に理解できる文,
 - 1. 違和感を感じるが理解可能な文,
 - 2. 違和感を感じて理解困難な文,
 - 3. 不自然な理解不能な文
- 提示の順序は回答者ごとに変えてあります。文の先頭についている整理数字は無視して下さい。

注意

- ・一次元化してあるけれども、これが妥当な想定かは怪しい

Web Survey の 手順



刺激文の構築

- 要件
 - 理論的バイアスをなるべく回避して刺激文を作る
 - そのためには、自動生成に勝るものはない
- 手順
 - BCCWJ の LWP から頻度基盤に動詞を見本抽出し、属性を付与し、それを FCA で分類（黒田 2018）
 - 広い範囲の動詞タイプを30選定,
 - 産物
 - 282種類の文
 - 30文で構成される gr0–gr9 で、延べ300文
 - s001, s002, …, s280, s281.0, s281.1, …, s281.9, s282.0, s282.1, …, s282.9

NINJAL-LWP

- NINJAL-LWP で出
現動詞を頻度順に並
べたもの

NINJAL-LWP for BCCWJ

→ 2語比較検索

ここに調べたい語を入力 (読みも可) [続り込み] [元に戻す]

すべて [名詞] [動詞] [形容詞] [遠休詞] [副詞] [オノマトペ]

見出し	読み	ローマ字表記	頻度
いる-非自立	イル	iru	958,064
する	スル	suru	601,911
なる	ナル	naru	484,861
ある	アル	aru	479,785
言う	イウ	iu	285,579
思う	オモウ	omou	199,915
できる	デキル	dekiyu	194,228
くる-非自立	クル	kuru	171,836
見る	ミル	miru	141,739
いる	イル	iru	95,901
てる-非自立	テル	teru	94,962
いく-非自立	イク	iku	93,950
しまう-非自立	シマウ	shimau	90,868
考える	カンガエル	kangaeru	84,500
分かる	ワカル	wakaru	81,524
持つ	モツ	motsu	78,996
行う	オコナウ	okonau	78,612
出る	デル	deru	74,409
下さる-非自立	クダサル	kudasaru	66,220
ある-非自立	オル	oru	64,387
やる	ヤル	yaru	63,451

page / 248 | 100 | 24,719件中 1 - 100を表示

NINJAL-LWP for BCCWJ Copyright © 2012-2016 National Institute for Japanese Language and Linguistics, Lago Institute of Language. All rights reserved.

属性のコーディング

・見本の手動コーディング

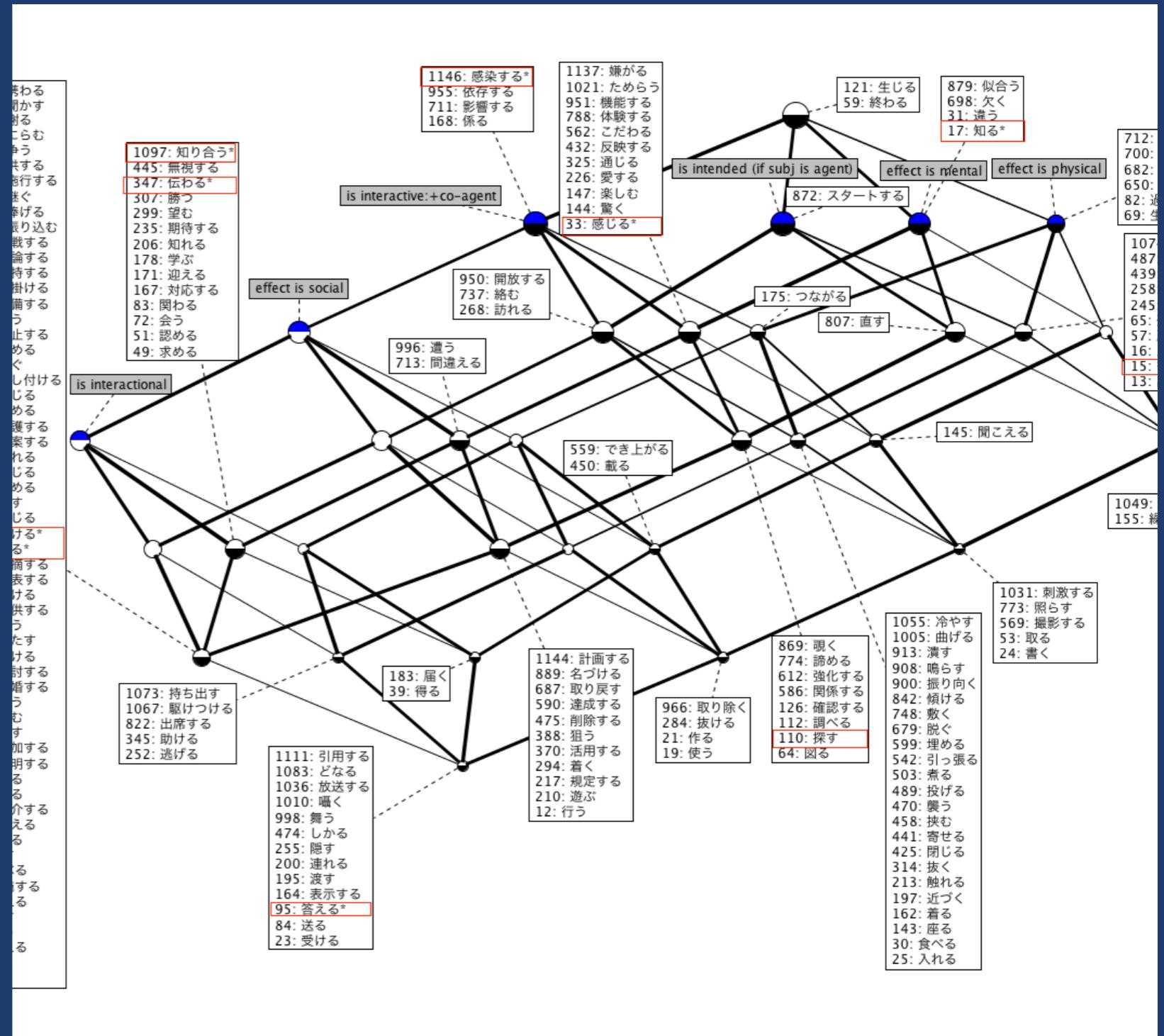
- F1: 物理的効果が伴うか: [effect is physical]]
- F2: 心理的効果が伴うか: [effect is mental]
- F3 社会的効果が伴うか: [effect is social]
- F4: 複数の意志ある主体の間の相互作用か: [event is interactive]
- F5: 複数の存在の相互作用か: [event is interactional]
- F6: 効果が (主語が意志ある主体である時に) 意図されたものか: [effect is intended (if subject is agent)]

Excel screenshot showing the 'v-freqs-for-BCCWJ.xlsx' spreadsheet. The table contains data for various verbs, categorized by their semantic features. The columns represent different attributes: lexeme, pronunciation, frequency, auxiliary, sampling, and six semantic features (F1 to F6). The rows list specific verbs like '出る*', '行く*', etc., along with their corresponding feature values.

lexeme	pronounce	freq	aux	sampled	effect is physical	effect is mental	effect is social	is interactive	is interactional	intended (if subject is agent)	check sum1	is_tivi	
出る*		74409	0	1	1	0	0	0.5	0	1	0	2.5	0
行く*		60396	0	1	1	0	0	0.5	0	1	0	2.5	0
入る*		57410	0	1	1	0	0	0.5	0	1	0	2.5	0
知る*		56618	0	1	0	1	0	0.5	0	0.5	0	2	
使う*		52857	0	1	1	1	1	1	0	1	0	5	0
作る*		50137	0	1	1	1	0.5	1	0	1	0	4.5	
参加する		8296	0	1	0.5	1	1	1	1	1	0	5.5	
表示する		7615	0	1	1	1	1	1	1	1	0	6	0
対応する		7473	0	1	0.5	1	1	1	1	1	0	5.5	
届く		6725	0	1	1	1	1	1	1	0	0	5	0
払う		6694	0	1	0.5	1	1	1	1	1	0	5.5	
結婚する		6513	0	1	0.5	1	1	1	1	1	0	5.5	0
検討する		6126	0	1	0.5	1	1	1	1	1	0	5.5	0
知れる		6123	0	1	0.5	1	1	1	1	1	0	4.5	
避ける		5962	0	1	0.5	1	1	1	1	1	0	5.5	0
果たす		5949	0	1	0.5	1	1	1	1	1	0	5.5	0
愛する		5919	0	1	0.5	1	0.5	1	0.5	0.5	0	4	0
しかる		0	1	1	1	1	1	1	1	1	0	6	
届く	きさやく	0	1	1	1	1	1	1	1	1	0	6	0
放送する		0	1	1	1	1	1	1	1	1	0	6	
持ち出す		0	1	1	0.5	1	1	1	1	1	0	5.5	0
どなる		0	1	1	1	1	1	1	1	1	0	6	0
引用する		0	1	1	1	1	1	1	1	1	0	6	0

FCA の結果

- 赤く着色してある動詞（のlemma）が刺激文に使う動詞の候補
 - 詳しくは
 - [6] 黒田 航 (2018). 意味の社会性を意識した動詞の分類とその理論的含意. 日本認知科学会第35会大会発表論文集, pp. 65-68, 2018.



動詞の内訳

v.index	v.form	count	ratio
v111	伝えた	17	0.057
v1147	知り合った	12	0.040
v116	答えた	1	0.003
v1197	感染した	1	0.003
v131	探した	1	0.003
v145	聞こえた	17	0.057
v155	繰り返した	9	0.030
v18	聞いた	23	0.077
v183	届いた	8	0.027
v210	遊んだ	17	0.057
v22	行った	1	0.003
v25	入れた	16	0.053
v26	話しかけた	1	0.003

v.index	v.form	count	ratio
v326	黙った	2	0.007
v338	負けた	1	0.003
v345	助けた	17	0.057
v377	表れた	1	0.003
v40	教えた	3	0.010
v44	感じた	10	0.033
v447	つないだ	8	0.027
v450	載った	18	0.060
v470	襲った	29	0.097
v713	間違った	9	0.030
v807	直した	23	0.077
v829	助かった	9	0.030
v831	届けた	17	0.057
v958	習った	29	0.097
sum		300	1.000

変異/編集タイプ の内訳

- o(original): 原文
- s(swapping): 分節 (NP/PP) 単位のランダムなかき混ぜ
- p(postposition): 後置詞のランダムな置換
- v(erb): 動詞のランダムな文脈類似語への置換
- n(nominal): 名詞類 (形容動詞の語幹を含む) のランダムな文脈類似語への置換

edit.type	count	ratio
o(original)	36	0.12
s(swapping)	70	0.23
p(postposition)	58	0.19
v(verb)	65	0.22
n(nominal)	71	0.24
sum	300	1.00

パターンの内訳

- P1: _が _で _に _と V した.
- P2: _が _で _に _を V した.
- P3: _が _で _を _に V した.
- P4: _が _で _から _を V し
た.
- P5: _が _で _と _を V した.

pattern	count	ratio
p1	64	0.21
p2	53	0.18
p3	67	0.22
p4	77	0.26
p5	39	0.13
sum	300	1

set2 の gr0 の 30 文

s.index	id	v.index	pattern	author	edit.type	gr	ver	gr.index	sentence
s10	10	v25	p4	3	n	gr0	A	1	担当者が携帯で出張もさから電話を入れた。
s20	20	v958	p3	1	p	gr0	A	2	熊がサークスで自転車をトレーナーを習った。
s30	30	v831	p3	1	n	gr0	A	3	伝書鳩が戦地で進攻を司令官に届けた。
s40	40	v807	p3	1	s	gr0	A	4	職員がまっすぐに絵画を美術館で直した。
s50	50	v831	p3	1	v	gr0	A	5	伝書鳩が戦地で戦況を司令官に送り届けた。
s60	60	v713	p5	3	s	gr0	A	6	子供が足し算とかけ算を宿題で間違えた。
s70	70	v345	p1	3	s	gr0	A	7	続編で宿敵がピンチに主人公と助けた。
s80	80	v18	p1	3	n	gr0	A	8	娘が病院で医者に軽症と聞いた。
s90	90	v807	p2	1	v	gr0	A	9	脚本家が話し合いで前日に台詞を考え直した。
s100	100	v470	p4	2	o	gr0	A	10	暴漢が鋭利な刃物で背後から人を襲った。
s110	110	v470	p5	2	s	gr0	A	11	仲間とコンビニを失業者が盗んだ自動車で襲った。
s120	120	v145	p3	1	v	gr0	A	12	ランナーが路上で悲鳴を夕暮れ時にかき消した。
s130	130	v807	p4	3	n	gr0	A	13	先生が赤ペンで一から詩を直した。
s140	140	v958	p5	3	v	gr0	A	14	弟が家で妹と料理を習わせた。
s150	150	v470	p5	2	v	gr0	A	15	失業者が盗んだ自動車で仲間とコンビニを襲った。
s160	160	v25	p3	3	n	gr0	A	16	男性が茶碗で角砂糖を紅茶に入れた。
s170	170	v145	p3	1	v	gr0	A	17	ランナーが路上で悲鳴を夕暮れ時に聞き取れた。
s180	180	v470	p5	2	p	gr0	A	18	失業者を盗んだ自動車で仲間とコンビニを襲った。
s190	190	v145	p3	1	o	gr0	A	19	ランナーが路上で悲鳴を夕暮れ時に聞こえた。
s200	200	v807	p2	1	n	gr0	A	20	脚本家が話し合いで翌日に台詞を直した。
s210	210	v345	p1	3	n	gr0	A	21	宿敵が続編で苦境に主人公と助けた。
s220	220	v111	p2	1	s	gr0	A	22	社長が社員に口頭で解雇を伝えた。
s230	230	v111	p2	1	v	gr0	A	23	社長が口頭で社員に解雇を読み取った。
s240	240	v210	p1	3	s	gr0	A	24	部下が温泉に北海道で同僚と遊んだ。
s250	250	v958	p1	1	s	gr0	A	25	医学生が解剖実習で看護師と医師に習った。
s260	260	v116	p3	2	s	gr0	A	26	見解をその芸能人が質問者に会見で答えた。
s270	270	v958	p5	3	s	gr0	A	27	家で弟が妹と料理を習った。
s280	280	v831	p1	1	p	gr0	A	28	新婚さんが役所が窓口に転居届と届けた。
s281	281	v1147	p1	1	p	gr0	A	29	夫が職場で真夜中に妻へ知り合った。
s282	282	v44	p4	1	n	gr0	A	30	学生が合格発表の場で足下から幸福を感じた。

set2 の gr1 の 30文

s.index	id	v.index	pattern	author	edit.type	gr	ver	gr.index	sentence
s1	1	v183	p2	1	n	gr1	A	1	幼児が肩車で茎に手を届いた。
s11	11	v210	p1	3	n	gr1	A	2	同僚が北海道で温泉に同僚と遊んだ。
s21	21	v958	p5	3	o	gr1	A	3	弟が家で妹と料理を習った。
s31	31	v25	p3	3	s	gr1	A	4	スプーンで男性が角砂糖を紅茶に入れた。
s41	41	v18	p2	3	n	gr1	A	5	祖母がキッチンで早朝にラジオを聞いた。
s51	51	v18	p1	3	o	gr1	A	6	娘が病院で医者に重症と聞いた。
s61	61	v713	p5	3	n	gr1	A	7	子供が宿題で足し算と分数を間違えた。
s71	71	v155	p2	3	s	gr1	A	8	抗議をデモで政府に大衆が繰り返した。
s81	81	v958	p1	1	n	gr1	A	9	医学生が解剖実習で医師に看護院と習った。
s91	91	v25	p3	3	p	gr1	A	10	男性がスプーンを角砂糖を紅茶に入れた。
s101	101	v829	p1	1	o	gr1	A	11	青年が震災で救出時に飼い犬と助かった。
s111	111	v210	p3	1	p	gr1	A	12	シロクマが動物園で氷と足元に遊んだ。
s121	121	v111	p4	1	p	gr1	A	13	司令官で無線で本部から命令を伝えた。
s131	131	v183	p2	1	s	gr1	A	14	肩車で幼児が枝に手を届いた。
s141	141	v345	p1	3	n	gr1	A	15	戦が続編でピンチに主人公と助けた。
s151	151	v470	p3	2	s	gr1	A	16	通り魔が休日の路上で不意に通行者を襲った。
s161	161	v18	p4	1	o	gr1	A	17	部下が給湯室で上司から噂を聞いた。
s171	171	v807	p3	1	s	gr1	A	18	絵画を美術館で職員がまっすぐに直した。
s181	181	v831	p3	1	s	gr1	A	19	司令官に戦地で戦況を伝書鳩が届けた。
s191	191	v713	p5	3	o	gr1	A	20	子供が宿題で足し算とかけ算を間違えた。
s201	201	v470	p4	2	v	gr1	A	21	暴漢が鋭利な刃物で背後から人を悩ました。
s211	211	v958	p4	1	v	gr1	A	22	後輩が現場で先輩から基礎を学んだ。
s221	221	v807	p3	1	n	gr1	A	23	職員がギメで絵画をまっすぐに直した。
s231	231	v470	p5	2	o	gr1	A	24	失業者が盗んだ自動車で仲間とコンビニを襲った。
s241	241	v470	p2	2	n	gr1	A	25	チーターが全速力で見事に井倉を襲った。
s251	251	v145	p5	1	p	gr1	A	26	市民が広場で銃声と叫び声で聞こえた。
s261	261	v111	p2	1	s	gr1	A	27	解雇を口頭で社員に社長が伝えた。
s271	271	v958	p4	1	p	gr1	A	28	後輩が現場で先輩から基礎で習った。
s281	281	v1147	p1	1	p	gr1	A	29	夫が職場で真夜中に妻へ知り合った。
s282	282	v44	p4	1	n	gr1	A	30	学生が合格発表の場で足下から幸福を感じた。

set2 の gr2 の 30 文

s.index	id	v.index	pattern	author	edit.type	gr	ver	gr.index	sentence
s2	2	v470	p3	2	p	gr2	A	1	通り魔が休日の路上が通行者を不意に襲った。
s12	12	v450	p4	1	v	gr2	A	2	高得点者が掲示で一位から十位を見れた。
s22	22	v345	p1	3	s	gr2	A	3	ピンチに続編で宿敵が主人公と助けた。
s32	32	v25	p4	3	s	gr2	A	4	携帯で担当者が出張先から電話を入れた。
s42	42	v40	p2	2	v	gr2	A	5	講師が受験対策で学生に良い解法を教わった。
s52	52	v450	p4	1	o	gr2	A	6	高得点者が掲示で一位から十位を載った。
s62	62	v1147	p1	1	v	gr2	A	7	夫が職場で真夜中に妻と会った。
s72	72	v958	p5	3	n	gr2	A	8	弟が家で妹とゴーヤーチャンプルーを習った。
s82	82	v807	p3	1	p	gr2	A	9	職員が美術館で絵画でまっすぐに直した。
s92	92	v450	p3	1	n	gr2	A	10	学的が論文賞で顔写真を広報誌に載った。
s102	102	v25	p3	3	s	gr2	A	11	紅茶にスプーンで角砂糖を男性が入れた。
s112	112	v338	p3	3	s	gr2	A	12	初戦を優勝候補がトーナメントで突然に負けた。
s122	122	v450	p3	1	p	gr2	A	13	学者が論文賞で顔写真から広報誌に載った。
s132	132	v155	p2	3	v	gr2	A	14	大衆がデモで政府に抗議を重ねた。
s142	142	v470	p3	2	p	gr2	A	15	通り魔が休日の路上を通行者を不意に襲った。
s152	152	v470	p5	2	p	gr2	A	16	失業者が盗んだ自動車が仲間とコンビニを襲った。
s162	162	v470	p5	2	n	gr2	A	17	雇用者が盗んだ自動車で仲間とコンビニを襲った。
s172	172	v829	p1	1	v	gr2	A	18	青年が震災で救出時に飼い犬と救った。
s182	182	v447	p5	1	p	gr2	A	19	医師が手術から血管と神経をつないだ。
s192	192	v210	p3	1	v	gr2	A	20	シロクマが動物園で氷を足元に楽しんだ。
s202	202	v713	p5	3	p	gr2	A	21	子供が宿題で足し算とかけ算から間違えた。
s212	212	v713	p5	3	v	gr2	A	22	子供が宿題で足し算とかけ算を忘れた。
s222	222	v958	p4	1	s	gr2	A	23	先輩から現場で後輩が基礎を習った。
s232	232	v831	p1	1	p	gr2	A	24	新婚さんが役所で窓口に転居届で届けた。
s242	242	v470	p5	2	v	gr2	A	25	失業者が盗んだ自動車で仲間とコンビニを悩まし
s252	252	v470	p3	2	s	gr2	A	26	休日の路上で通り魔が通行者を不意に襲った。
s262	262	v183	p2	1	v	gr2	A	27	幼児が肩車で枝に手を送った。
s272	272	v326	p4	3	s	gr2	A	28	捕虜が忠誠心から秘密を尋問で黙った。
s281	281	v1147	p1	1	p	gr2	A	29	夫が職場で真夜中に妻へ知り合った。
s282	282	v44	p4	1	n	gr2	A	30	学生が合格発表の場で足下から幸福を感じた。

set2 の gr3 の 30 文

s.index	id	v.index	pattern	author	edit.type	gr	ver	gr.index	sentence
s3	3	v25	p4	3	p	gr3	A	1	担当者が携帯で出張先を電話を入れた。
s13	13	v326	p4	3	p	gr3	A	2	捕虜が尋問に忠誠心から秘密を黙った。
s23	23	v470	p4	2	s	gr3	A	3	暴漢が鋭利な刃物で人を背後から襲った。
s33	33	v210	p3	1	s	gr3	A	4	氷を動物園でシロクマが足元に遊んだ。
s43	43	v470	p4	2	v	gr3	A	5	暴漢が鋭利な刃物で背後から人を襲いかかった。
s53	53	v18	p4	1	n	gr3	A	6	上司が給湯室で上司から噂を聞いた。
s63	63	v958	p4	1	v	gr3	A	7	後輩が現場で先輩から基礎を通った。
s73	73	v345	p4	3	p	gr3	A	8	人が河原で洪水が子猫を助けた。
s83	83	v450	p4	1	p	gr3	A	9	高得点者が掲示で一位から十位が載った。
s93	93	v18	p4	1	p	gr3	A	10	部下で給湯室で上司から噂を聞いた。
s103	103	v40	p4	2	o	gr3	A	11	字のうまい青年が外国で生活の必要から書道を教え
s113	113	v111	p4	1	v	gr3	A	12	司令官が無線で本部から命令を読みとった。
s123	123	v111	p4	1	p	gr3	A	13	司令官が無線で本部が命令を伝えた。
s133	133	v958	p3	1	s	gr3	A	14	熊がトレーナーに自転車をサークルで習った。
s143	143	v807	p3	1	v	gr3	A	15	職員が美術館で絵画をまっすぐに直せた。
s153	153	v958	p1	1	o	gr3	A	16	医学生が解剖実習で医師に看護師と習った。
s163	163	v450	p3	1	s	gr3	A	17	顔写真を論文賞で学者が広報誌に載った。
s173	173	v958	p4	1	n	gr3	A	18	親友が現場で先輩から基礎を習った。
s183	183	v183	p2	1	n	gr3	A	19	幼児が船虫で枝に手を届いた。
s193	193	v155	p2	3	v	gr3	A	20	大衆がデモで政府に抗議を繰返した。
s203	203	v18	p1	3	p	gr3	A	21	娘が病院を医者に重症と聞いた。
s213	213	v958	p1	1	n	gr3	A	22	医師に解剖実習で医学生が看護師と習った。
s223	223	v807	p4	3	n	gr3	A	23	君が赤ペンで一から文章を直した。
s233	233	v831	p3	1	v	gr3	A	24	伝書鳩が戦地で戦況を司令官に持ち帰った。
s243	243	v25	p3	3	v	gr3	A	25	男性がスプーンで角砂糖を紅茶に取り出した。
s253	253	v450	p4	1	n	gr3	A	26	高得点者が掲示で一位から十ワーストを載った。
s263	263	v958	p1	1	v	gr3	A	27	医学生が解剖実習で医師に看護師と教わった。
s273	273	v958	p4	1	n	gr3	A	28	後輩が現場で同僚から基礎を習った。
s281	281	v1147	p1	1	p	gr3	A	29	夫が職場で真夜中に妻へ知り合った。
s282	282	v44	p4	1	n	gr3	A	30	学生が合格発表の場で足下から幸福を感じた。

set2 の gr4 の 30 文

s.index	id	v.index	pattern	author	edit.type	gr	ver	gr.index	sentence
s4	4	v145	p5		p	gr4	A	1	市民が広場で銃声と叫び声が聞こえた。
s14	14	v450	p4		n	gr4	A	2	高得点者が掲示で一位から九位を載った。
s24	24	v958	p4		s	gr4	A	3	後輩が現場で基礎を先輩から習った。
s34	34	v807	p4	3	v	gr4	A	4	先生が赤ペンで一から文章を作り直した。
s44	44	v958	p3		v	gr4	A	5	熊がサークルで自転車をトレーナーに習わせた。
s54	54	v345	p1	3	o	gr4	A	6	宿敵が続編でピンチに主人公と助けた。
s64	64	v829	p1		v	gr4	A	7	青年が震災で救出時に飼い犬と救えた。
s74	74	v470	p4	2	p	gr4	A	8	暴漢で鋭利な刃物で背後から人を襲った。
s84	84	v111	p2		p	gr4	A	9	社長が口頭で社員に解雇に伝えた。
s94	94	v831	p3		n	gr4	A	10	伝書鳩が戦地で戦況を司令官で届けた。
s104	104	v470	p2	2	s	gr4	A	11	見事に全速力でチーターがガゼルを襲った。
s114	114	v450	p4		s	gr4	A	12	一位から掲示で高得点者が十位を載った。
s124	124	v447	p5		n	gr4	A	13	医師が手術で肺と神経をつないだ。
s134	134	v447	p5		v	gr4	A	14	医師が手術で血管と神経を組んだ。
s144	144	v958	p3		o	gr4	A	15	熊がサークルで自転車をトレーナーに習った。
s154	154	v450	p3		v	gr4	A	16	学者が論文賞で顔写真を広報誌に書いた。
s164	164	v807	p3		v	gr4	A	17	職員が美術館で絵画をまっすぐに直った。
s174	174	v25	p4	3	v	gr4	A	18	担当者が携帯で出張先から電話を投げ入れた。
s184	184	v345	p4	3	s	gr4	A	19	人が子猫を洪水から河原で助けた。
s194	194	v210	p1	3	p	gr4	A	20	部下が北海道で温泉を同僚と遊んだ。
s204	204	v210	p1	3	o	gr4	A	21	部下が北海道で温泉に同僚と遊んだ。
s214	214	v25	p4	3	s	gr4	A	22	担当者が電話を出張先から携帯で入れた。
s224	224	v831	p1		s	gr4	A	23	新婚さんが役所で転居届と窓口に届けた。
s234	234	v450	p4		s	gr4	A	24	十位を掲示で一位から高得点者が載った。
s244	244	v18	p1	3	n	gr4	A	25	娘が病院で医者に心筋梗塞と聞いた。
s254	254	v18	p1	3	s	gr4	A	26	娘が病院で重症と医者に聞いた。
s264	264	v145	p3		p	gr4	A	27	ランナーが路上で悲鳴から夕暮れ時に聞こえた。
s274	274	v18	p2	3	v	gr4	A	28	母がキッチンで早朝にラジオを聞けた。
s281	281	v1147	p1		p	gr4	A	29	夫が職場で真夜中に妻へ知り合った。
s282	282	v44	p4		n	gr4	A	30	学生が合格発表の場で足下から幸福を感じた。

set2 の gr5 の 30 文

s.index	id	v.index	pattern	author	edit.type	gr	ver	gr.index	sentence
s5	5	v25	p4	3	n	gr5	A	1	担当者が携帯で出張さきから電話を入れた。
s15	15	v450	p3	1	v	gr5	A	2	学者が論文賞で顔写真を広報誌に見れた。
s25	25	v958	p1	1	p	gr5	A	3	医学生で解剖実習で医師に看護師と習った。
s35	35	v183	p2	1	p	gr5	A	4	幼児が肩車で枝から手を届いた。
s45	45	v18	p4	1	s	gr5	A	5	部下が噂を上司から給湯室で聞いた。
s55	55	v111	p4	1	v	gr5	A	6	司令官が無線で本部から命令をくみ取った。
s65	65	v26	p3	2	v	gr5	A	7	彼女が手紙で真実を不意に話しかけた。
s75	75	v210	p3	1	s	gr5	A	8	シロクマが足元に氷を動物園で遊んだ。
s85	85	v111	p2	1	n	gr5	A	9	オーナーが口頭で社員に解雇を伝えた。
s95	95	v210	p1	3	p	gr5	A	10	部下が北海道で温泉が同僚と遊んだ。
s105	105	v807	p2	1	v	gr5	A	11	脚本家が話し合いで前日に台詞を戻した。
s115	115	v450	p3	1	n	gr5	A	12	学者が論文賞で顔図版を広報誌に載った。
s125	125	v713	p5	3	n	gr5	A	13	子どもが宿題で足し算とかけ算を間違えた。
s135	135	v345	p4	3	n	gr5	A	14	中国人が河原で洪水から子猫を助けた。
s145	145	v155	p2	3	n	gr5	A	15	大衆が反戦で政府に抗議を繰り返した。
s155	155	v829	p1	1	p	gr5	A	16	青年が震災を救出時に飼い犬と助かった。
s165	165	v345	p4	3	v	gr5	A	17	人が河原で洪水から子猫を力づけた。
s175	175	v131	p2	2	n	gr5	A	18	親戚連れが潮干狩りで海に貝を探した。
s185	185	v831	p1	1	v	gr5	A	19	新婚さんが役所で窓口に転居届と送った。
s195	195	v145	p5	1	v	gr5	A	20	市民が広場で銃声と叫び声をかき消した。
s205	205	v1147	p3	4	o	gr5	A	21	関係者がインターネットで個人情報を相互に知り
s215	215	v831	p1	1	v	gr5	A	22	新婚さんが役所で窓口に転居届と受け取った。
s225	225	v345	p4	3	s	gr5	A	23	人が洪水から河原で子猫を助けた。
s235	235	v713	p5	3	v	gr5	A	24	子供が宿題で足し算とかけ算を分かった。
s245	245	v807	p3	1	o	gr5	A	25	職員が美術館で絵画をまっすぐに直した。
s255	255	v807	p4	3	p	gr5	A	26	先生が赤ペンで一から文章で直した。
s265	265	v831	p3	1	n	gr5	A	27	伝書鳩が戦地で戦況を司令分掌に届けた。
s275	275	v1197	p1	3	o	gr5	A	28	私が遊園地でインフルエンザに家族と感染した。
s281	281	v1147	p1	1	p	gr5	A	29	夫が職場で真夜中に妻へ知り合った。
s282	282	v44	p4	1	n	gr5	A	30	学生が合格発表の場で足下から幸福を感じた。

set2 の gr6 の 30文

s.index	id	v.index	pattern	author	edit.type	gr	ver	gr.index	sentence
s6	6	v829	p1	1	p	gr6	A	1	青年が震災が救出時に飼い犬と助かった。
s16	16	v18	p2	3	v	gr6	A	2	母がキッチンで早朝にラジオを尋ねた。
s26	26	v447	p5	1	v	gr6	A	3	医師が手術で血管と神経を結んだ。
s36	36	v18	p1	3	v	gr6	A	4	娘が病院で医者に重症と聴いた。
s46	46	v831	p3	1	s	gr6	A	5	戦況を戦地で伝書鳩が司令官に届けた。
s56	56	v183	p2	1	v	gr6	A	6	幼児が肩車で枝に手を出せた。
s66	66	v18	p4	1	n	gr6	A	7	上官が給湯室で上司から噂を聞いた。
s76	76	v713	p5	3	s	gr6	A	8	足し算とかけ算を宿題で子供が間違えた。
s86	86	v155	p2	3	p	gr6	A	9	大衆までデモで政府に抗議を繰り返した。
s96	96	v807	p2	1	o	gr6	A	10	脚本家が話し合いで前日に台詞を直した。
s106	106	v829	p1	1	s	gr6	A	11	青年が救出時に震災で飼い犬と助かった。
s116	116	v831	p1	1	n	gr6	A	12	新婚さんが役所で窓口に転居届けと届けた。
s126	126	v18	p4	1	p	gr6	A	13	部下に給湯室で上司から噂を聞いた。
s136	136	v470	p4	2	n	gr6	A	14	乱闘が鋭利な刃物で背後から人を襲った。
s146	146	v111	p4	1	s	gr6	A	15	無線で司令官が本部から命令を伝えた。
s156	156	v155	p2	3	p	gr6	A	16	大衆がデモに政府に抗議を繰り返した。
s166	166	v807	p2	1	p	gr6	A	17	脚本家まで話し合いで前日に台詞を直した。
s176	176	v145	p3	1	n	gr6	A	18	マラソンが路上で悲鳴を夕暮れ時に聞こえた。
s186	186	v18	p1	3	s	gr6	A	19	医者に病院で娘が重症と聞いた。
s196	196	v470	p3	2	o	gr6	A	20	通り魔が休日の路上で通行者を不意に襲った。
s206	206	v345	p4	3	o	gr6	A	21	人が河原で洪水から子猫を助けた。
s216	216	v447	p5	1	s	gr6	A	22	医師が血管と神経を手術でつないだ。
s226	226	v958	p1	1	v	gr6	A	23	医学生が解剖実習で医師に看護師と習わせた。
s236	236	v450	p4	1	v	gr6	A	24	高得点者が掲示で一位から十位を書けた。
s246	246	v807	p4	3	o	gr6	A	25	先生が赤ペンで一から文章を直した。
s256	256	v345	p1	3	p	gr6	A	26	宿敵が続編からピンチに主人公と助けた。
s266	266	v807	p4	3	v	gr6	A	27	先生が赤ペンで一から文章を戻した。
s276	276	v807	p2	1	n	gr6	A	28	脚本家が話し合いで前日にフレーズを直した。
s281	281	v1147	p1	1	p	gr6	A	29	夫が職場で真夜中に妻へ知り合った。
s282	282	v44	p4	1	n	gr6	A	30	学生が合格発表の場で足下から幸福を感じた。

set2 の gr7 の 30 文

s.index	id	v.index	pattern	author	edit.type	gr	ver	gr.index	sentence
s7	7	v210	p1	3	s	gr7	A	1	同僚と北海道で温泉に部下が遊んだ。
s17	17	v447	p5	1	s	gr7	A	2	血管と神経を手術で医師がつないだ。
s27	27	v111	p2	1	o	gr7	A	3	社長が口頭で社員に解雇を伝えた。
s37	37	v145	p5	1	n	gr7	A	4	コミュニティが広場で銃声と叫び声を聞こえた。
s47	47	v450	p3	1	o	gr7	A	5	学者が論文賞で顔写真を広報誌に載った。
s57	57	v145	p5	1	n	gr7	A	6	市民が広場で足音と叫び声を聞こえた。
s67	67	v18	p1	3	v	gr7	A	7	娘が病院で医者に重症と訊いた。
s77	77	v829	p1	1	n	gr7	A	8	青年が震災で救出正午に飼い犬と助かった。
s87	87	v25	p3	3	n	gr7	A	9	男性がスプーンで角砂糖をコーヒーに入れた。
s97	97	v958	p1	1	p	gr7	A	10	医学生が解剖実習で医師に看護師で習った。
s107	107	v807	p2	1	s	gr7	A	11	前日に話し合いで脚本家が台詞を直した。
s117	117	v470	p2	2	p	gr7	A	12	チーターから全速力で見事にガゼルを襲った。
s127	127	v450	p4	1	p	gr7	A	13	高得点者が掲示を一位から十位を載った。
s137	137	v145	p5	1	v	gr7	A	14	市民が広場で銃声と叫び声を聴こえた。
s147	147	v807	p2	1	p	gr7	A	15	脚本家が話し合いへ前日に台詞を直した。
s157	157	v210	p3	1	o	gr7	A	16	シロクマが動物園で氷を足元に遊んだ。
s167	167	v958	p5	3	p	gr7	A	17	弟が家で妹と料理まで習った。
s177	177	v18	p2	3	p	gr7	A	18	母がキッチンが早朝にラジオを聞いた。
s187	187	v470	p2	2	v	gr7	A	19	チーターが全速力で見事にガゼルを襲いかかった。
s197	197	v155	p2	3	s	gr7	A	20	大衆が抗議を政府にデモで繰り返した。
s207	207	v958	p5	3	p	gr7	A	21	弟が家で妹から料理を習った。
s217	217	v155	p2	3	n	gr7	A	22	大衆がデモで政府に弾劾を繰り返した。
s227	227	v145	p3	1	n	gr7	A	23	ランナーが軒下で悲鳴を夕暮れ時に聞こえた。
s237	237	v345	p4	3	n	gr7	A	24	人が河原で土砂から子猫を助けた。
s247	247	v210	p3	1	n	gr7	A	25	シロクマがほ乳類園で氷を足元に遊んだ。
s257	257	v450	p3	1	s	gr7	A	26	学者が顔写真を論文賞で広報誌に載った。
s267	267	v447	p5	1	o	gr7	A	27	医師が手術で血管と神経をつないだ。
s277	277	v829	p1	1	n	gr7	A	28	青年が震災で救出時に飼い主と助かった。
s281	281	v1147	p1	1	p	gr7	A	29	夫が職場で真夜中に妻へ知り合った。
s282	282	v44	p4	1	n	gr7	A	30	学生が合格発表の場で足下から幸福を感じた。

set2 の gr8 の 30文

s.index	id	v.index	pattern	author	edit.type	gr	ver	gr.index	sentence
s8	8	v807	p4	3	s	gr8	A	1	先生が一から赤ペンで文章を直した。
s18	18	v210	p3	1	n	gr8	A	2	オコジョが動物園で氷を足元に遊んだ。
s28	28	v958	p1	1	n	gr8	A	3	医学生が解剖実習で医師に看護居士と習った。
s38	38	v345	p4	3	p	gr8	A	4	人が河原で洪水から子猫が助けた。
s48	48	v210	p1	3	v	gr8	A	5	部下が北海道で温泉に同僚と楽しんだ。
s58	58	v25	p4	3	p	gr8	A	6	担当者が携帯で出張先から電話で入れた。
s68	68	v345	p1	3	v	gr8	A	7	宿敵が続編でピンチに主人公と助け出した。
s78	78	v111	p2	1	n	gr8	A	8	社長が口頭で社内に解雇を伝えた。
s88	88	v345	p4	3	v	gr8	A	9	人が河原で洪水から子猫を助け出した。
s98	98	v210	p3	1	p	gr8	A	10	シロクマに動物園で氷を足元に遊んだ。
s108	108	v210	p1	3	n	gr8	A	11	隊長が北海道で温泉に同僚と遊んだ。
s118	118	v829	p1	1	s	gr8	A	12	飼い犬と震災で救出時に青年が助かった。
s128	128	v447	p5	1	p	gr8	A	13	医師が手術を血管と神経をつないだ。
s138	138	v18	p4	1	v	gr8	A	14	部下が給湯室で上司から噂を尋ねた。
s148	148	v470	p2	2	p	gr8	A	15	チーターを全速力で見事にガゼルを襲った。
s158	158	v470	p2	2	o	gr8	A	16	チーターが全速力で見事にガゼルを襲った。
s168	168	v111	p4	1	o	gr8	A	17	司令官が無線で本部から命令を伝えた。
s178	178	v183	p2	1	o	gr8	A	18	幼児が肩車で枝に手を届いた。
s188	188	v345	p1	3	p	gr8	A	19	宿敵へ続編でピンチに主人公と助けた。
s198	198	v18	p2	3	n	gr8	A	20	母が居間で早朝にラジオを聞いた。
s208	208	v18	p2	3	p	gr8	A	21	母がキッチンから早朝にラジオを聞いた。
s218	218	v40	p2	2	n	gr8	A	22	講座が受験対策で女学生に良い固有値を教えた。
s228	228	v145	p3	1	s	gr8	A	23	路上でランナーが悲鳴を夕暮れ時に聞こえた。
s238	238	v111	p2	1	p	gr8	A	24	社長が口頭で社員に解雇で伝えた。
s248	248	v145	p5	1	s	gr8	A	25	広場で市民が銃声と叫び声を聞こえた。
s258	258	v831	p1	1	o	gr8	A	26	新婚さんが役所で窓口に転居届と届けた。
s268	268	v470	p3	2	v	gr8	A	27	通り魔が休日の路上で通行者を不意に苛(さいな)ん
s278	278	v111	p2	1	v	gr8	A	28	社長が口頭で社員に解雇を伝わった。
s281	281	v1147	p1	1	p	gr8	A	29	夫が職場で真夜中に妻へ知り合った。
s282	282	v44	p4	1	n	gr8	A	30	学生が合格発表の場で足下から幸福を感じた。

set2 の gr9 の 30 文

s.index	id	v.index	pattern	author	edit.type	gr	ver	gr.index	sentence
s9	9	v470	p3	2	v	gr9	A	1	通り魔が休日の路上で通行者を不意に追い払った。
s19	19	v807	p4	3	s	gr9	A	2	先生が赤ペンで文章を一から直した。
s29	29	v210	p1	3	v	gr9	A	3	部下が北海道で温泉に同僚と遊べだ。
s39	39	v470	p2	2	n	gr9	A	4	温血動物が全速力で見事にガゼルを襲った。
s49	49	v145	p5	1	s	gr9	A	5	市民が銃声と叫び声を広場で聞こえた。
s59	59	v145	p3	1	s	gr9	A	6	ランナーが悲鳴を路上で夕暮れ時に聞こえた。
s69	69	v111	p4	1	n	gr9	A	7	司令官が受信で本部から命令を伝えた。
s79	79	v155	p2	3	o	gr9	A	8	大衆がデモで政府に抗議を繰り返した。
s89	89	v713	p5	3	p	gr9	A	9	子供が宿題で足し算とかけ算に間違えた。
s99	99	v22	p3	5	o	gr9	A	10	船が遠回りで海路を安全に行った。
s109	109	v25	p3	3	o	gr9	A	11	男性がスプーンで角砂糖を紅茶に入れた。
s119	119	v470	p2	2	s	gr9	A	12	チーターが全速力でガゼルを見事に襲った。
s129	129	v377	p4	3	v	gr9	A	13	波紋が投石で落下点から水面を表れた。
s139	139	v25	p3	3	v	gr9	A	14	男性がスプーンで角砂糖を紅茶に放り込んだ。
s149	149	v111	p4	1	s	gr9	A	15	命令を無線で本部から司令官が伝えた。
s159	159	v183	p2	1	p	gr9	A	16	幼児が肩車が枝に手を届いた。
s169	169	v831	p3	1	o	gr9	A	17	伝書鳩が戦地で戦況を司令官に届けた。
s179	179	v958	p4	1	o	gr9	A	18	後輩が現場で先輩から基礎を習った。
s189	189	v18	p1	3	p	gr9	A	19	娘で病院で医者に重症と聞いた。
s199	199	v25	p4	3	o	gr9	A	20	担当者が携帯で出張先から電話を入れた。
s209	209	v831	p3	1	p	gr9	A	21	伝書鳩を戦地で戦況を司令官に届けた。
s219	219	v958	p3	1	n	gr9	A	22	熊が四重奏で自転車をトレーナーに習った。
s229	229	v958	p5	3	v	gr9	A	23	弟が家で妹と料理を教わった。
s239	239	v470	p2	2	v	gr9	A	24	チーターが全速力で見事にガゼルを見舞った。
s249	249	v450	p3	1	p	gr9	A	25	学者が論文賞に顔写真を広報誌に載った。
s259	259	v145	p3	1	p	gr9	A	26	ランナーが路上で悲鳴を夕暮れ時に聞こえた。
s269	269	v831	p1	1	n	gr9	A	27	新婚さんが役所で窓口に転居書類と届けた。
s279	279	v18	p2	3	o	gr9	A	28	母がキッチンで早朝にラジオを聞いた。
s281	281	v1147	p1	1	p	gr9	A	29	夫が職場で真夜中に妻へ知り合った。
s282	282	v44	p4	1	n	gr9	A	30	学生が合格発表の場で足下から幸福を感じた。

gr0–g9 の特性

- 30文のうち、次の 2文 (s281, s283) が全 gr に共通
- s281:v1147:p1:1:p:gr9:29
 - 夫が 職場で 真夜中に 妻へ 知り合った。
- s282: v44:p4:1:n:gr9:30
 - 学生が 合格発表の場で 足下から 幸福を 感じた。
- 他の28文は gr に特有

評定課題の実施



なぜ属性を取得するのか

- 容認度評定は、個人の様々な特性を反映する
 - そうでないはずがない
 - とは言え、反映のすべてが有意味とは限らない
- 容認度評定値を左右しそうな特性を評定課題と一緒に取得しておく事は、調査としては本質的
 - これをやっておけば後で属性ごとに層別解析ができるし、回帰もできる

容認度評定の前に11種類の質問 1/2

- Q1. あなたは現在、何才ですか？
- Q2. あなたの性別はどれですか？
- Q3. あなたの母語は日本語ですか？
- Q4. あなたは過去にどの地域（県単位）で暮らしたことがありますか？該当するものすべてを指定して下さい。
- Q5. 過去に日本語が話されない地域で1年以上の生活をしたことありますか？
- Q6. 学んだ事のある外国語の種類の数を概数で答えて下さい（期間の長さに関係なく、10個以上の場合は区別なしで）.

容認度評定の前に11種類の質問 2/2

- Q7. 英語を含めた異国語を学んだ期間の今までの合計を、およその年数で答えて下さい。
- Q8. 外国語/異国語を話す人と日常的に接しますか？
- Q9. 一ヶ月当たりの読書量を、およその冊数で答えて下さい
(画像中心の雑誌は含めません)。
- Q10. 小学校より後の教育機関でどれ位学んだか、およその年数を2桁の数字で答えて下さい。
例えば普通大学の学部卒業の場合、10年 (=中学で3年 + 高等学校で3年 + 大学で4年) です。
- Q11. あなたは自分の考え方や感じ方が大きく分けて文系だと思いますか？理系だと思いますか？
(自分の主観評価で構いません)

Q1-Q11

質問 回答 178

4セクション中8個目のセクション

あなたの特性について

これから10種類の特性をお聞きします。質問の提示の順序は回答者ごとに変更されています。質問の先頭に付いている整理番号は無視して下さい。

1. あなたは現在、何才ですか？*

1. 13才以下
2. 14才
3. 15才
4. 16才
5. 17才
6. 18才
7. 19才
8. 20才
9. 21才
10. 22才
11. 23才



質問 回答 178

10. 小学校より後の教育機関でどれ位学んだか、およよその年数を2桁の数字で答えて下さい。例えば普通大学の学部卒業の場合、10年（=中学で3年+高等学校で3年+大学で4年）です。

1. (6年+) 3年以下
2. (6年+) 4年から6年
3. (6年+) 7年から10年
4. (6年+) 11年から12年
5. (6年+) 13年から15年
6. (6年+) 16年以上

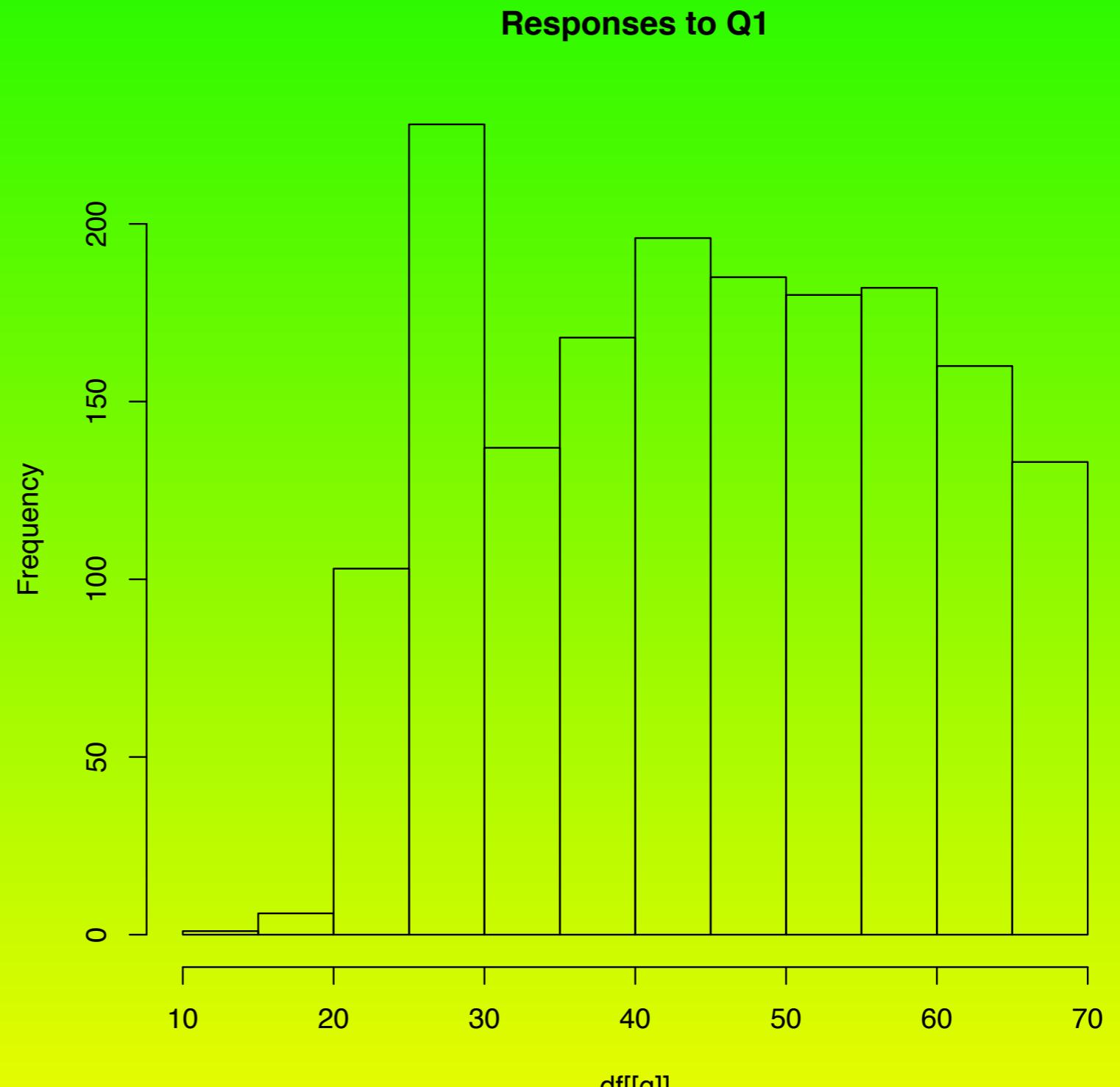
11. あなたは自分の考え方や感じ方が大きく分けて文系だと思いますか？理系だと思いますか？（自分の主観評価で構いません）

- とっても文系だと思う
- 文系っぽいだと思う
- どちらとも言えない
- 理系っぽいと思う
- とっても理系だと思う

特性に関する質問は終わりです。次に30種類の文に関する印象をお聞きします。

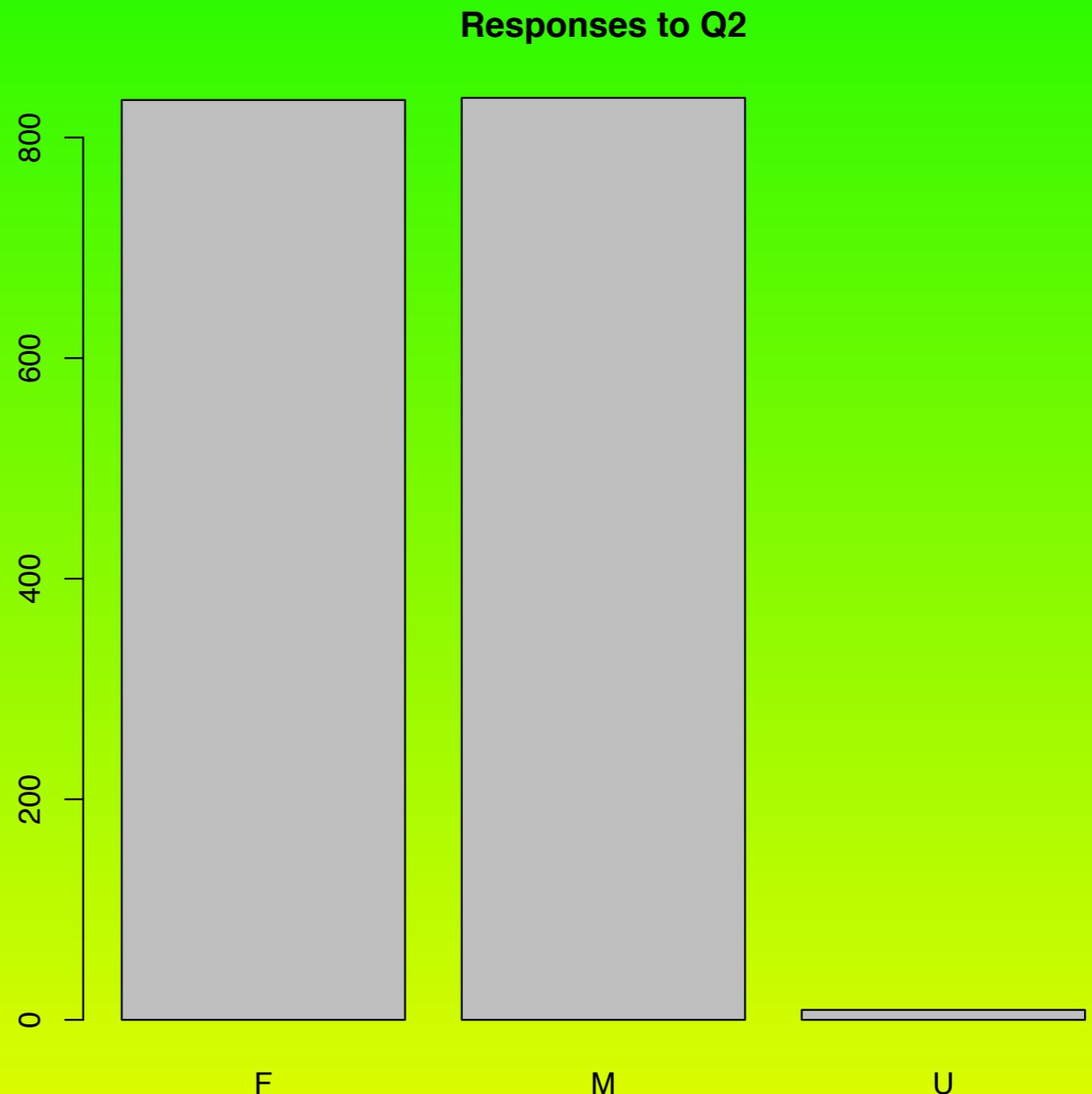
Q1への回答

- Q1. あなたは現在,
何才ですか？
-



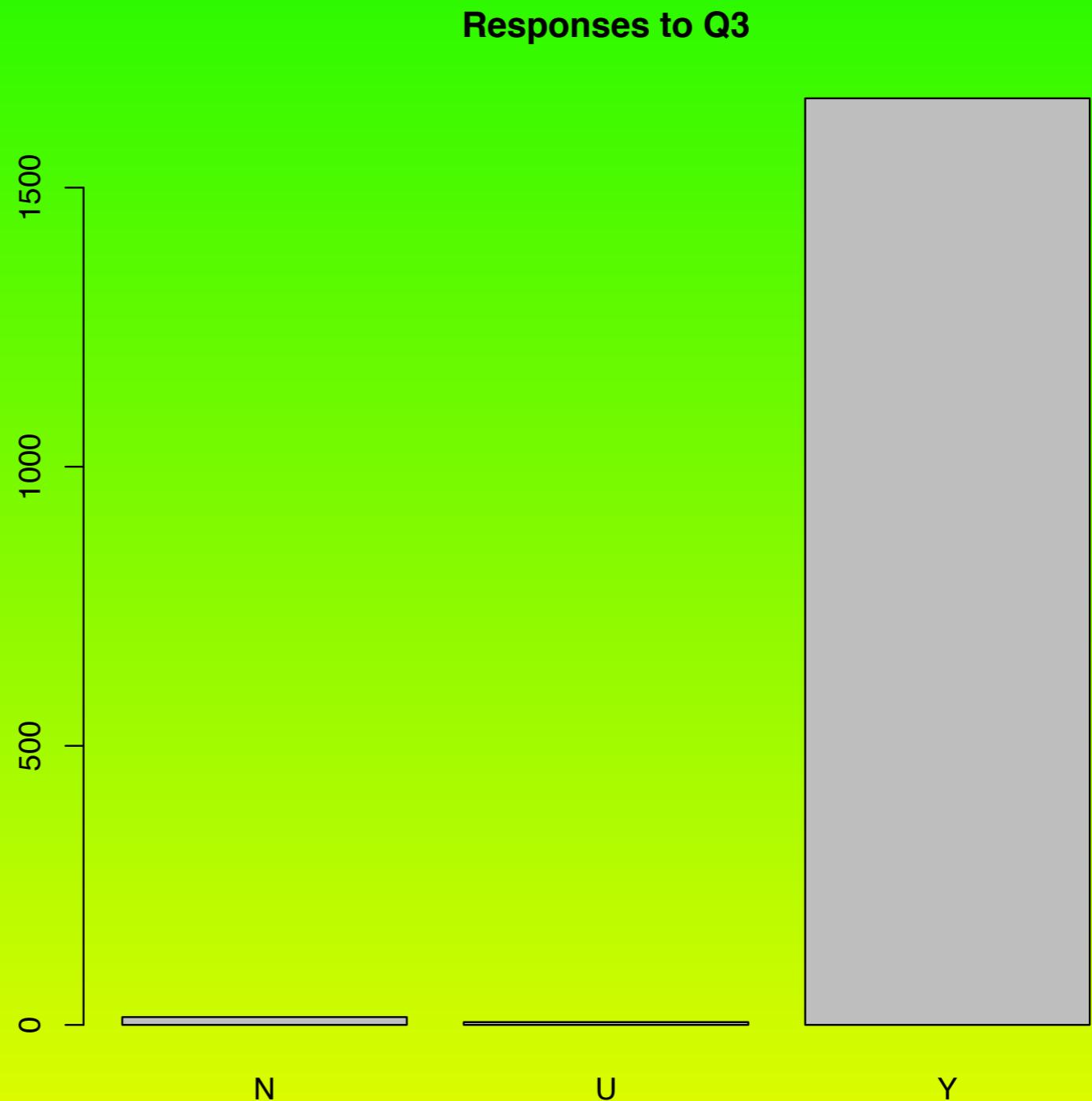
Q2への回答

- Q2. あなたの性別
はどれですか？



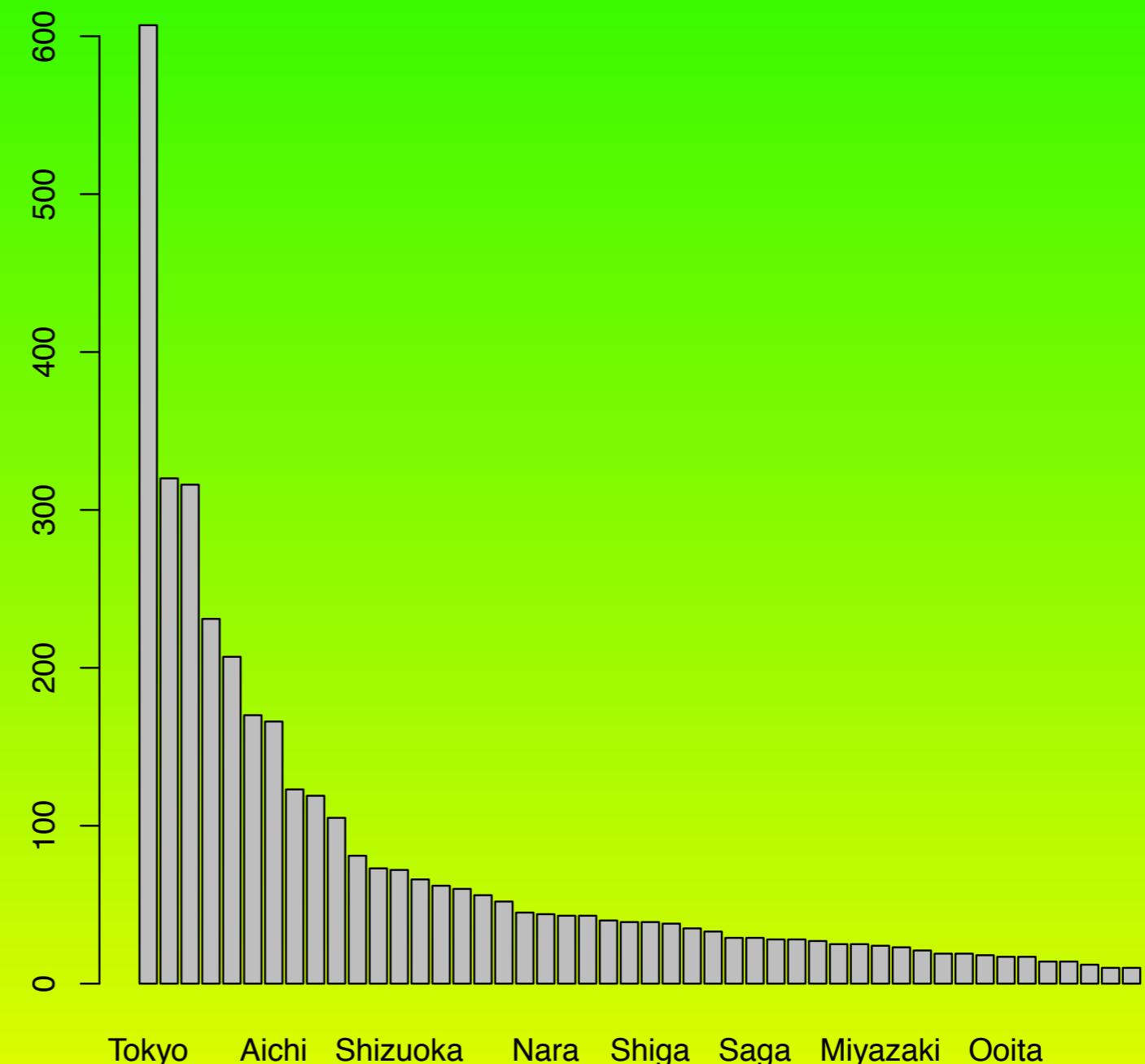
Q3への回答

- Q3. あなたの母語
は日本語ですか？



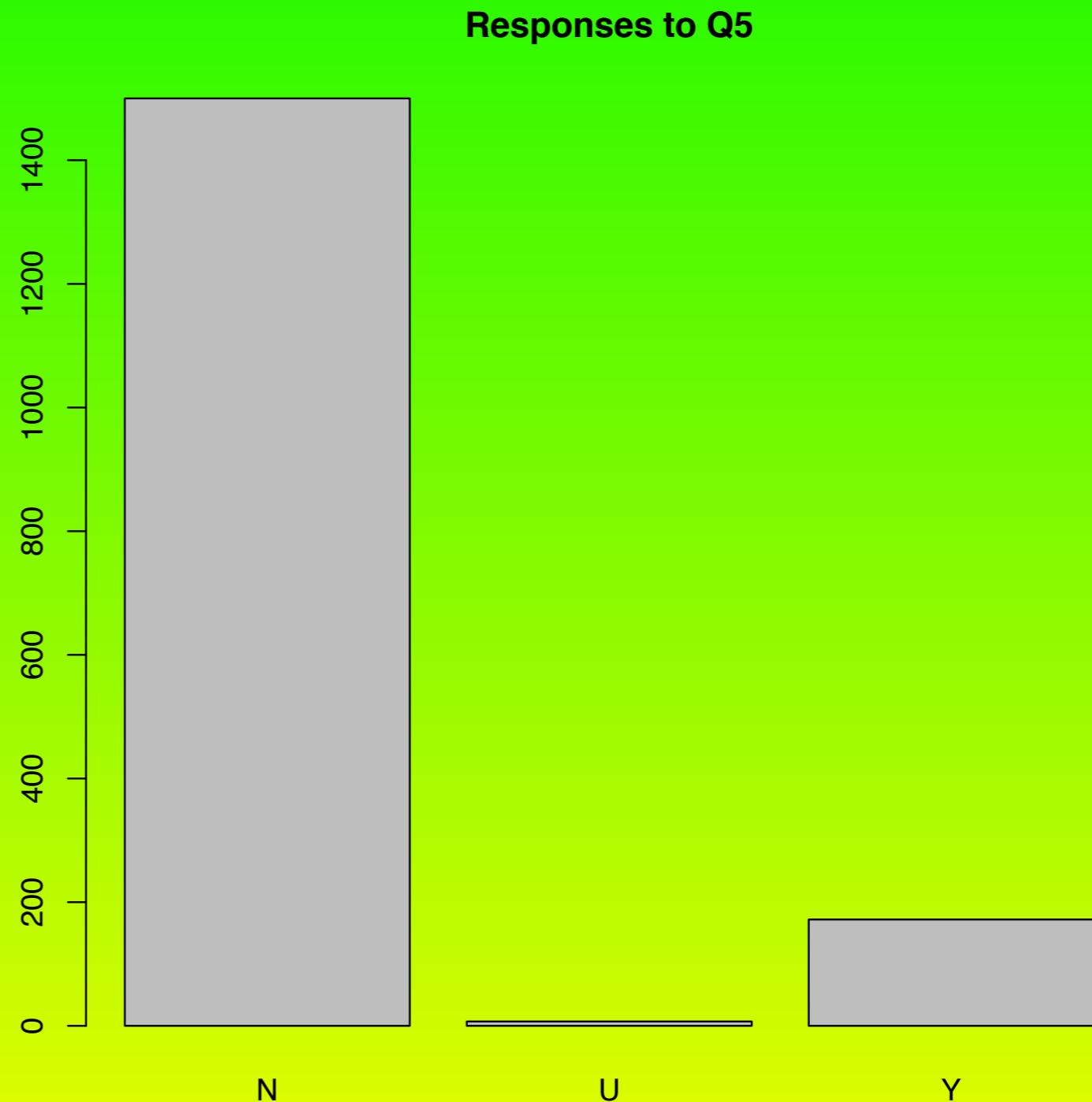
Q4への回答

- Q4. あなたは過去にどの地域（県単位）で暮らしたことがありますか？該当するものすべてを指定して下さい。



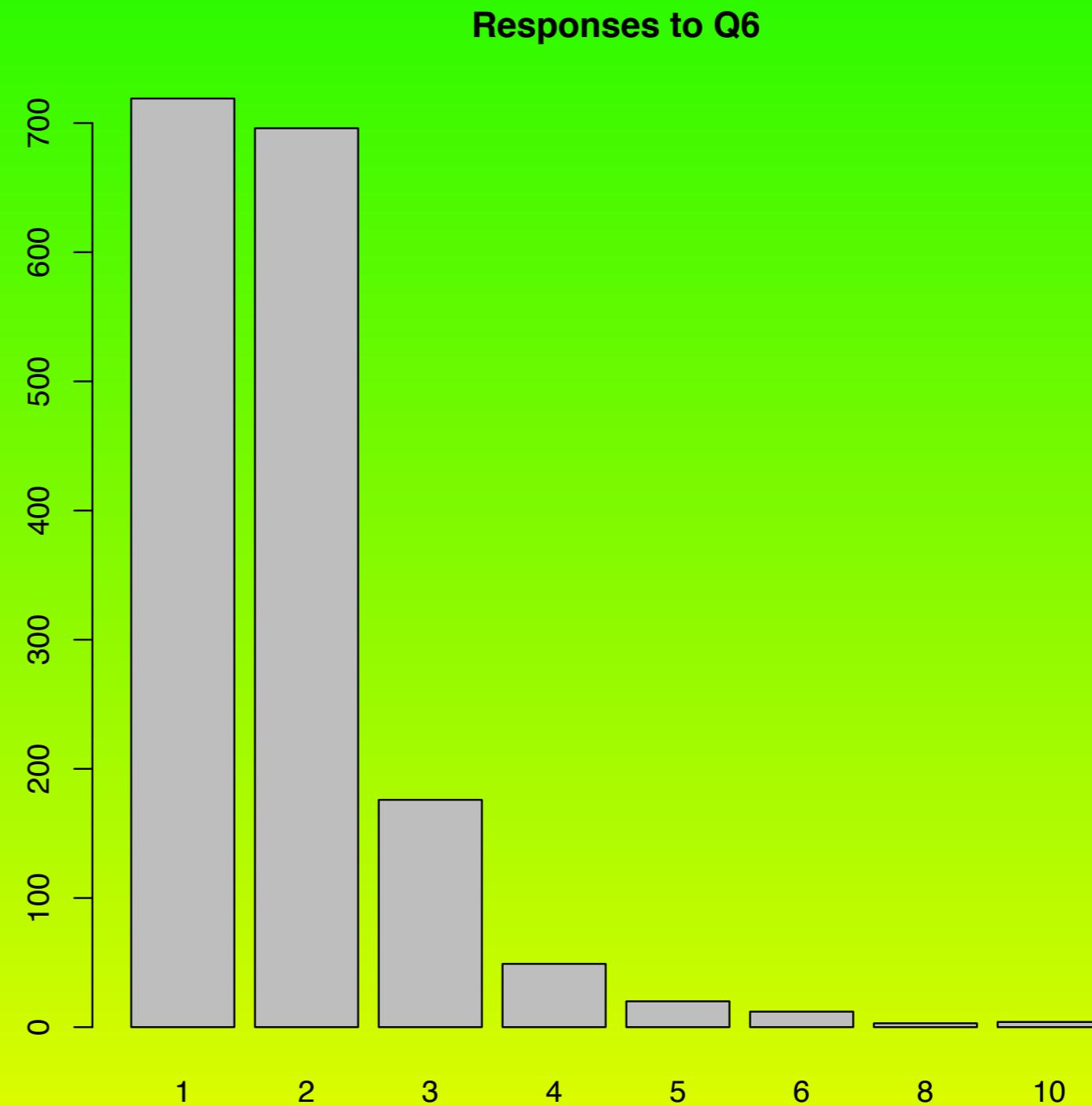
Q5への回答

- Q5. 過去に日本語が話されない地域で1年以上の生活をしたことがありますか？



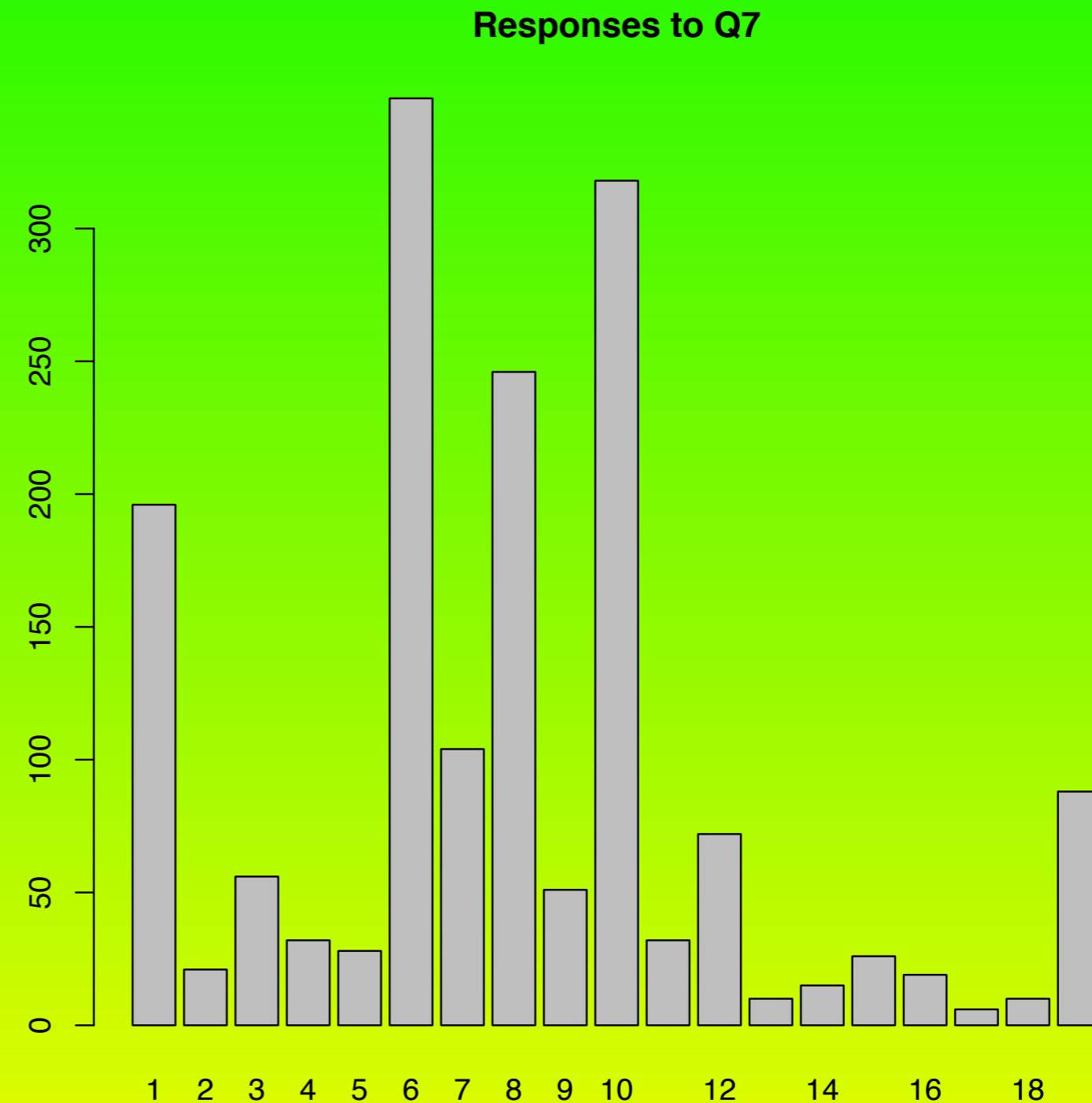
Q6への回答

- Q6. 学んだ事のある外国語の種類の数を概数で答えて下さい (期間の長さに関係なく、10個以上の場合は区別なしで)。



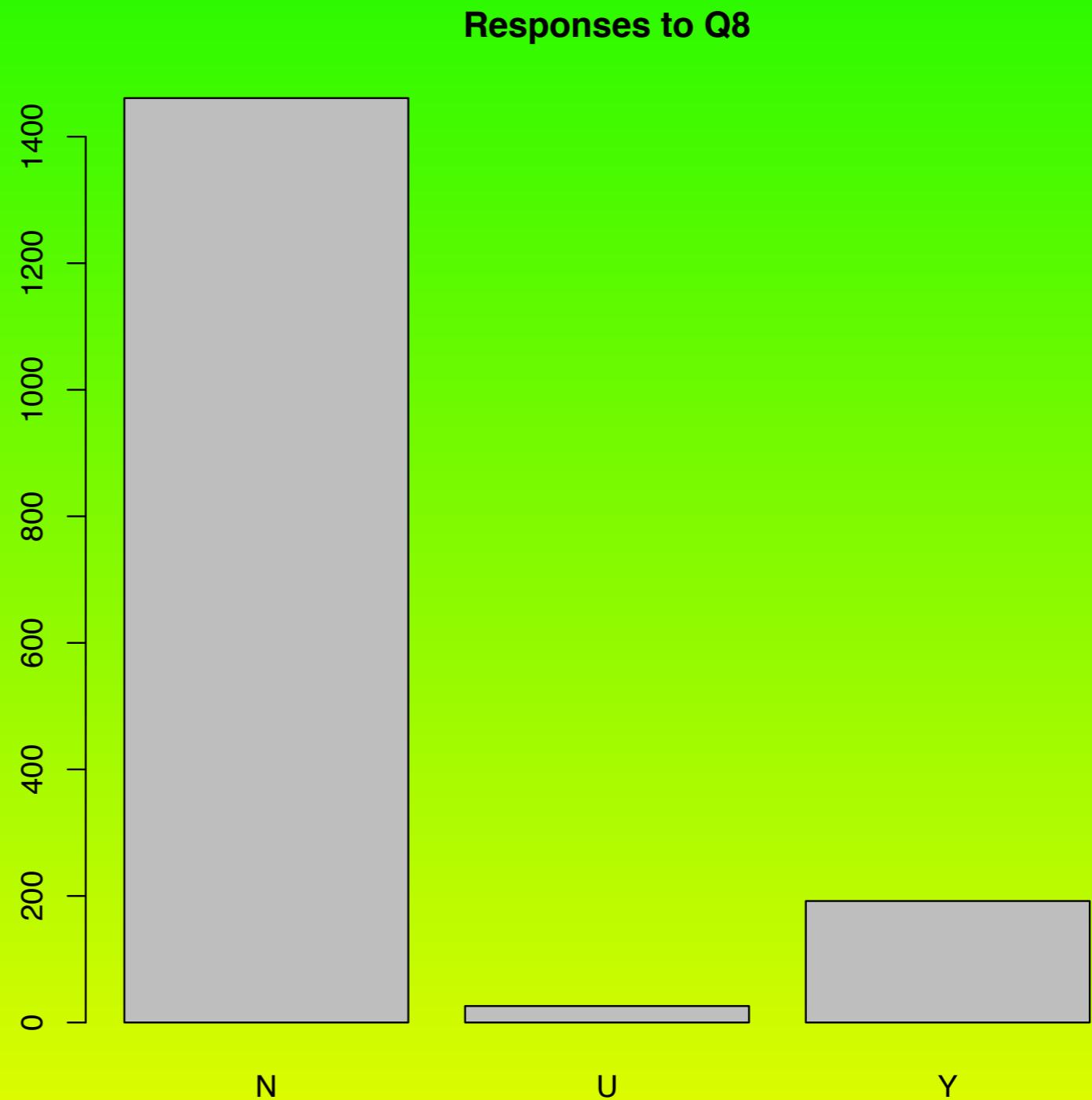
Q7への回答

- Q7. 英語を含めた異国語を学んだ期間の今までの合計を、およその年数で答えて下さい。
-



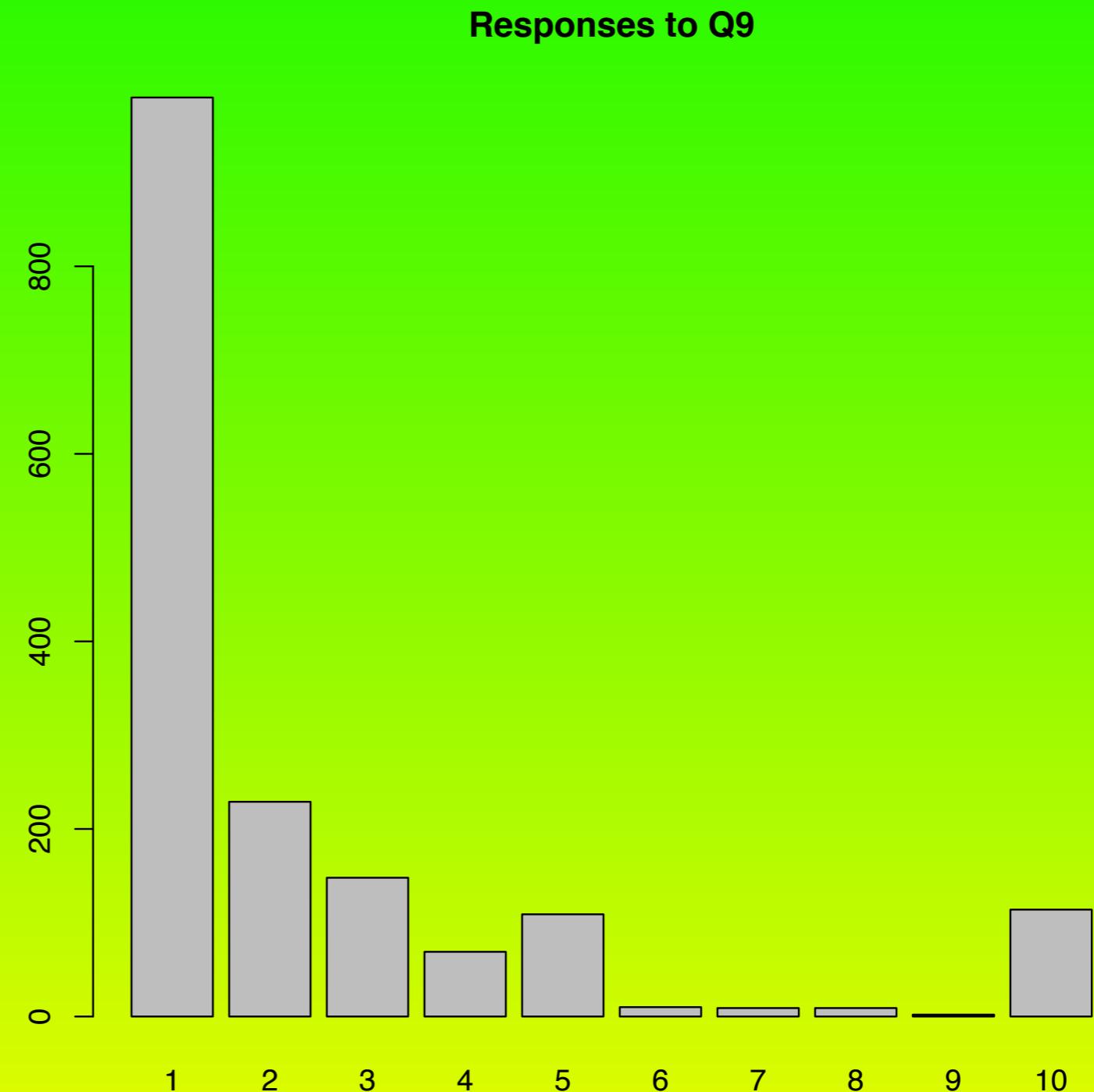
Q8への回答

- Q8. 外国語/異国語を話す人と日常的に接しますか？
-



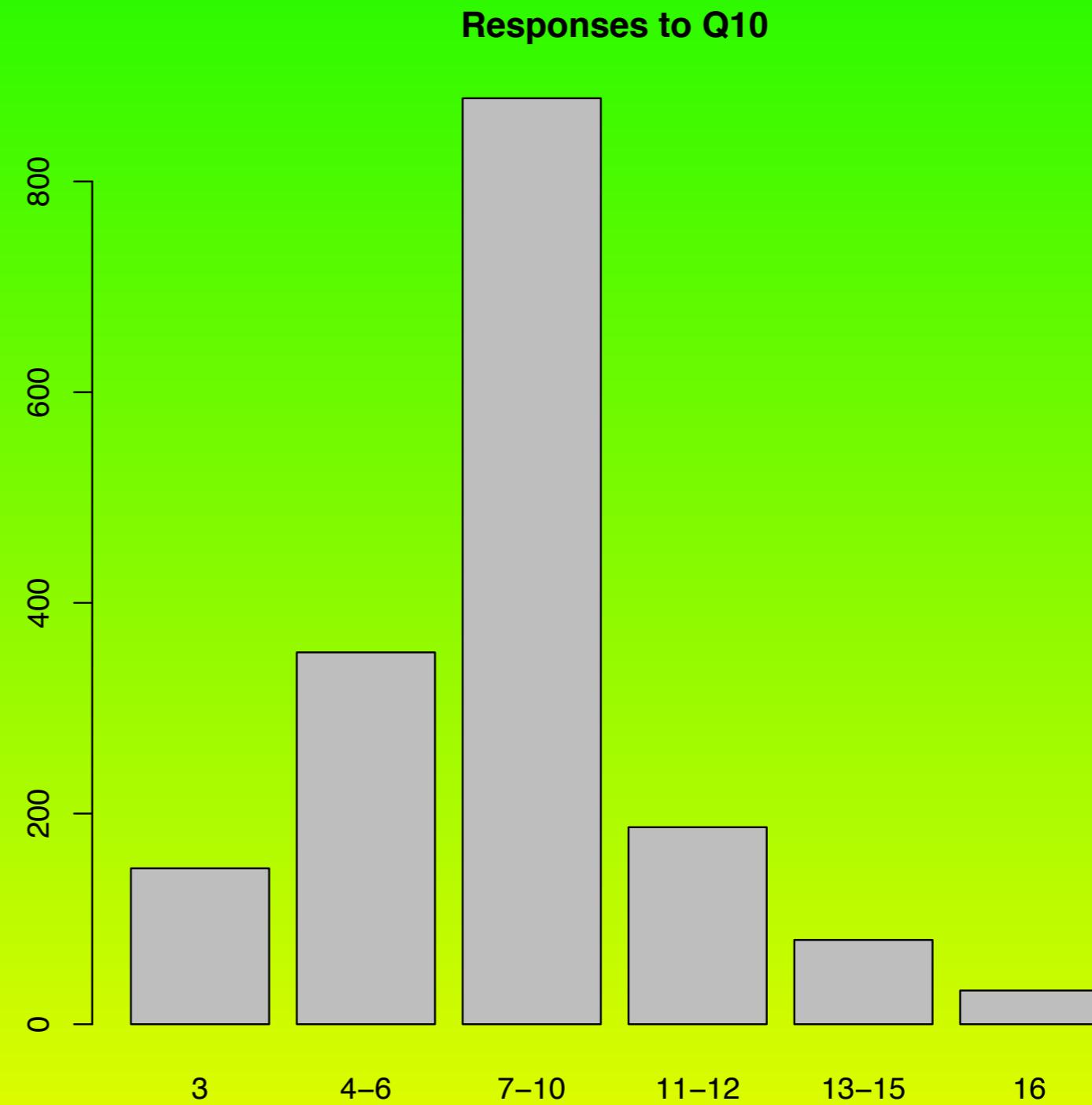
Q9への回答

- Q9. 一ヶ月当たりの読書量を、おおよその冊数で答えて下さい（画像中心の雑誌は含めません）。



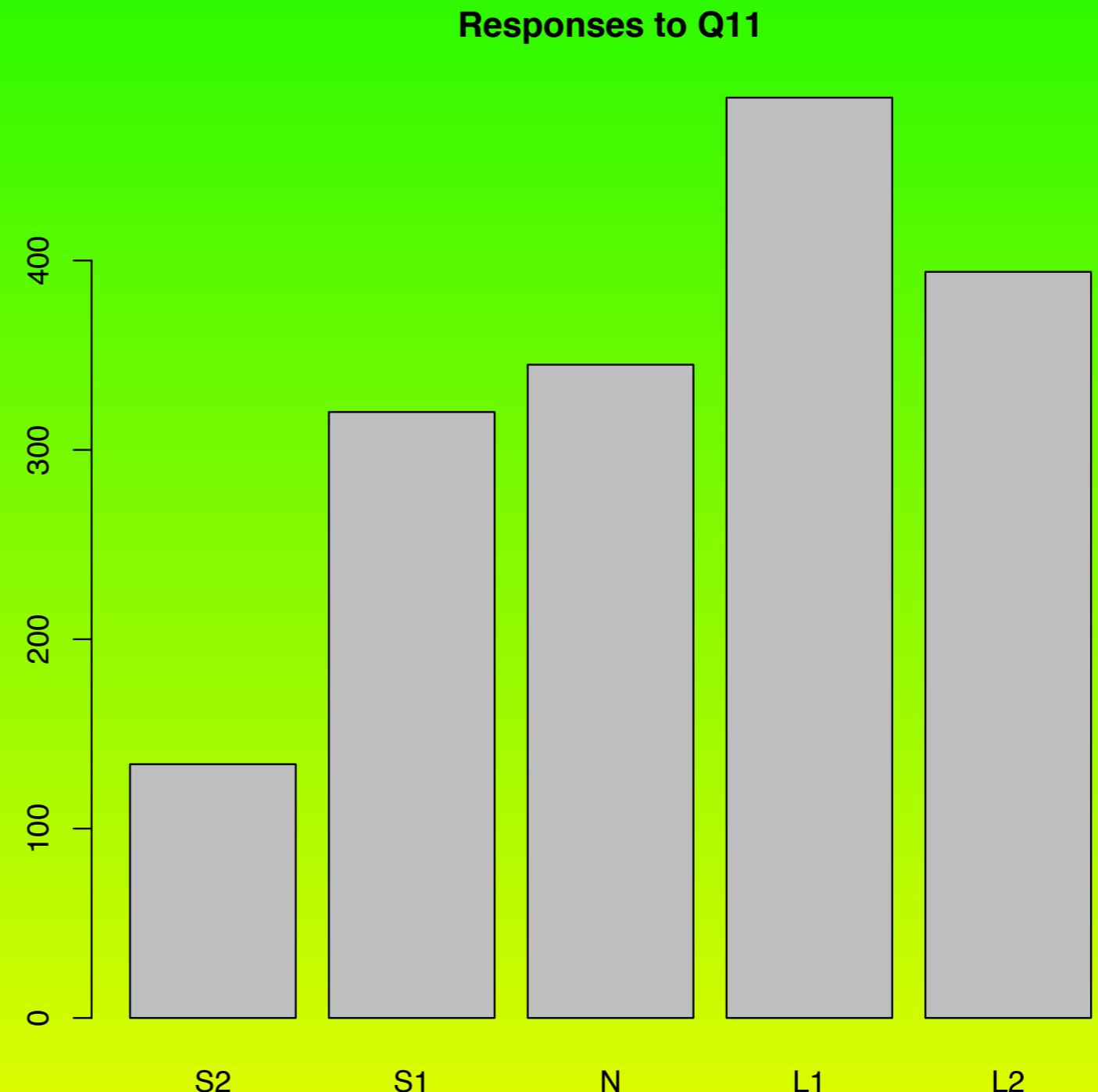
Q10への回答

- Q10. 小学校より後の教育機関でどれ位学んだか、おおよその年数を2桁の数字で答えて下さい。例えば普通大学の学部卒業の場合、10年(=中学で3年 + 高等学校で3年 + 大学で4年)です。



Q11への回答

- Q11. あなたは自分の考え方や感じ方が大きく分けて文系だと思いますか？理系だと思いますか？（自分の主観評価で構いません）



調査の第二節

- 30種類の刺激文の容認度を評定する課題が10種類 gr0, gr1, ..., gr9
- gr.i の刺激文は $\text{mod}(\text{S.ID}) = i$

4 セクション中 4 個目のセクション

日本語の文の評定作業

次に30種類の文をお見せします。個々の文を次の4つの基準で評定して下さい。

0. 違和感がなく自然に理解できる文,
1. 違和感を感じるが理解可能な文,
2. 違和感を感じて理解困難な文,
3. 不自然な理解不能な文

提示の順序は回答者ごとに変えてあります。文の先頭に付いている整理数字は無視して下さい。

s006: 青年が震災が救出時に飼い犬と助かった。*

	0	1	2	3	
違和感のない自然に理解可能な文	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	不自然で理解不能な文

刺激の評定者への提示

- Web 実験と紙ベースの実験で実施条件が異なる
 - Web 実験では Google Form を使っているので、11個の質問と30種類の刺激文の提示順は回答ごとにランダム化紙ベースの実験では、事前に順序の異なる A, B, C の三つの版を用意し、回答者はそのどれかに回答

評定課題 (mode 1)

- 次に30種類の文をお見せします。個々の文を次の4つの基準で評定して下さい。
 - 0. 違和感がなく自然に理解できる文
 - 1. 違和感を感じるが理解可能な文
 - 2. 違和感を感じて理解困難な文
 - 3. 不自然な理解不能な文
- 提示の順序は回答者ごとに変えてあります。文の先頭についている整理数字は無視して下さい。

評定課題 (mode 2)

- 次にお見せする30種類の文はコンピュータが自動作成したものです。一部の品質には難があり、それを見つけたいと思っています。そのために、個々の文を次の4つの基準で評定して下さい。
 - 0. 違和感がなく自然に理解できる文
 - 1. 違和感を感じるが理解可能な文
 - 2. 違和感を感じて理解困難な文
 - 3. 不自然な理解不能な文
- 提示の順序は回答者ごとに変えてあります。文の先頭に付いている整理数字は無視して下さい。

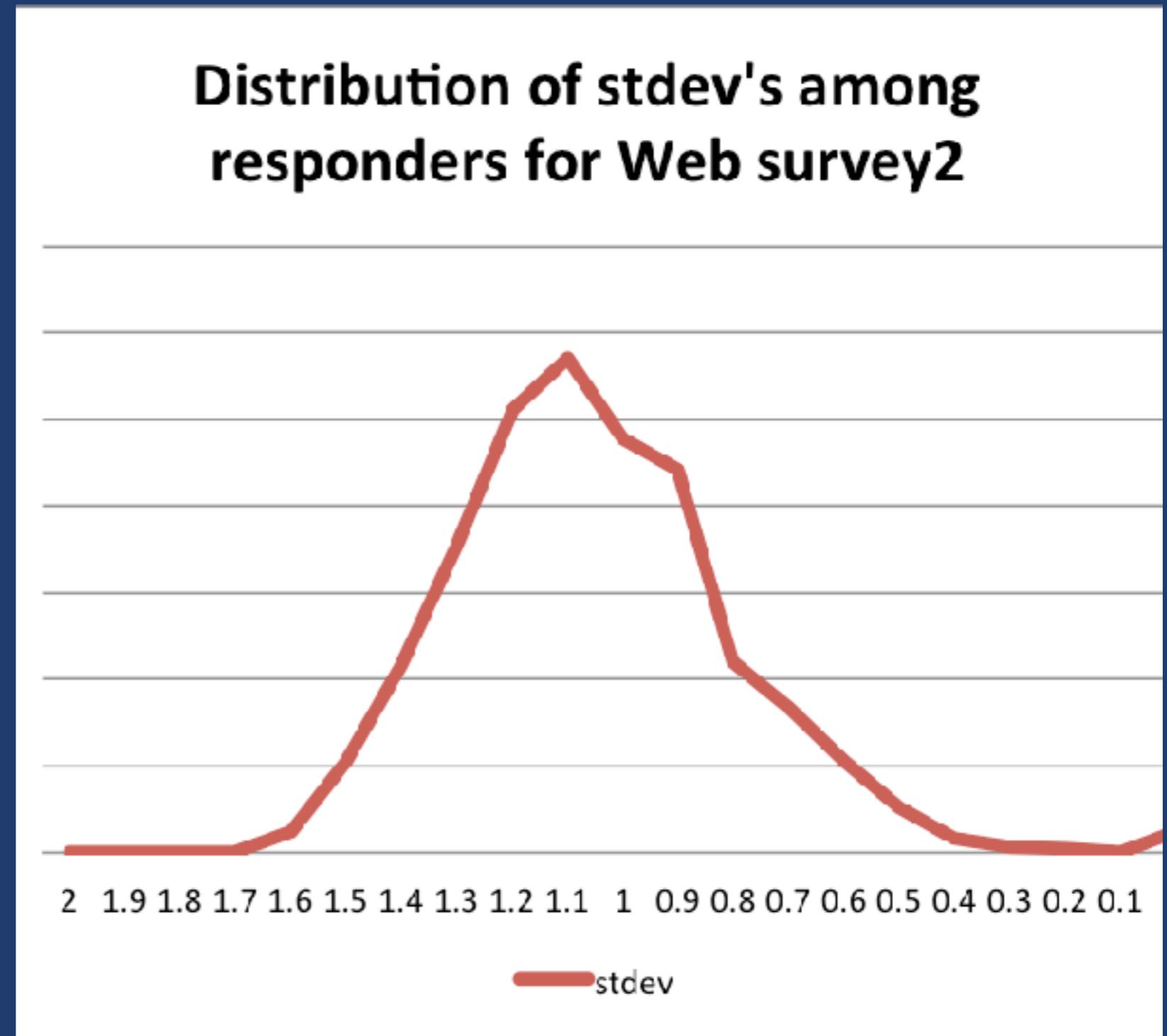
Web 調査の概略

- gr 内で、回答者の評定値の `stdev` の分布を考える
- `stdev` の `stdev` の分布で、 $0.5 < \text{stdev}(\text{stdev}) < 1.6$ の範囲にない回答者を除外

gr	count	removed	effective count
gr0	166	6	160
gr1	174	5	169
gr2	169	8	161
gr3	168	5	163
gr4	166	7	159
gr5	163	10	153
gr6	178	3	175
gr7	168	5	163
gr8	162	7	155
gr9	165	4	161
sum	1679	60	1619

回答者の評定値の $stdev$ の分布

- 反応の $stdev$ の $stdev$ が 0 の回答者はすべてに同じ評定値をつけている
- 要するに手抜きをしている (satisfice 現象)
- 反応の $stdev$ の $stdev$ が並外れて大きい回答者はデタラメに評定値をつけている可能性が大きい



収集した反応を一瞥

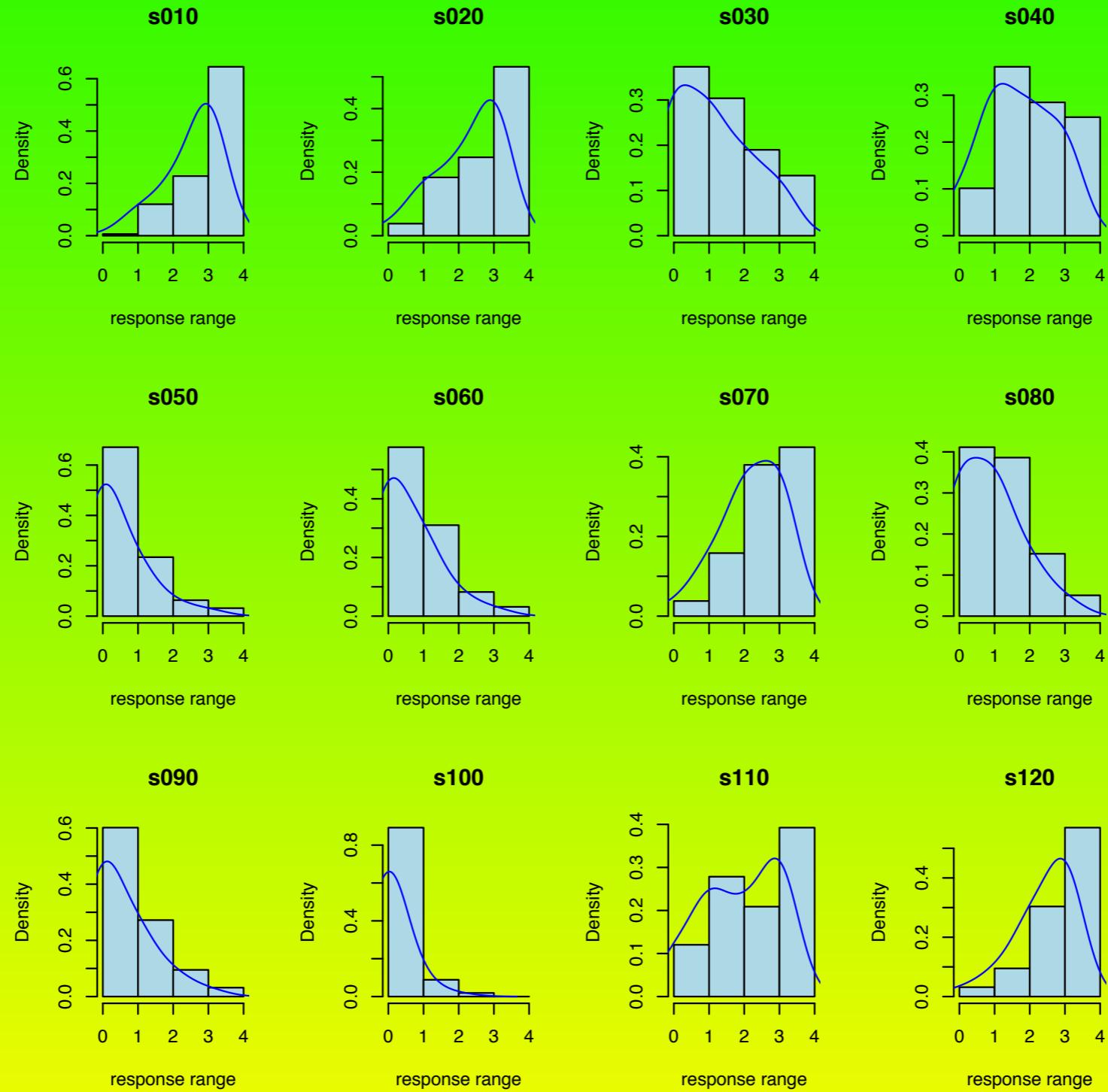
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
Timestamp	stage	R.ID	Q1.あなたは現在、何	Q9.一ヶ月当たりの読書量をおお	Q10.小学校より後の教育機関で	Q11.あなたは自分の考え方を	s010:担当者が指導で出張もさが	s020:都がサービスで自転車針	s030:伝書メ							
1 2018/11/08 5:18:45 PM GN	1	45398997	50才	1冊以下	(6年+) 4年から6年	とっても文系だと思います	3	3	3							
2 2018/11/08 5:23:27 PM GN	1	81949566	48才	1冊以下	(6年+) 15年以上	どちらとも言えない	3	3	3							
3 2018/11/08 5:24:35 PM GN	1	21230961	24才	1冊以下	(6年+) 4年から6年	どちらとも言えない	0	0	0							
4 2018/11/08 5:25:16 PM GN	1	32778919	28才	1冊以下	(6年+) 4年から6年	文系っぽいと思う	3	1	1							
5 2018/11/08 5:25:20 PM GN	1	14697308	57才	2冊くらい	(6年+) 7年から10年	とっても文系だと思います	3	3	3							
6 2018/11/08 5:27:23 PM GN	1	47660142	35才	1冊以下	(6年+) 4年から6年	とっても文系だと思います	3	3	3							
7 2018/11/08 5:28:49 PM GN	1	71734856	65才	1冊以下	(6年+) 4年から6年	文系っぽいと思う	1	0	0							
8 2018/11/08 5:36:10 PM GN	1	82998866	27才	5冊くらい	(6年+) 7年から10年	文系っぽいと思う	3	1	1							
9 2018/11/08 5:40:44 PM GN	1	25835289	29才	3冊くらい	(6年+) 7年から10年	文系っぽいと思う	3	3	3							
10 2018/11/08 5:42:00 PM GN	1	44439170	69才	3冊くらい	(6年+) 11年から12年	どちらとも言えない	2	2	2							
148 2018/11/15 8:53:29 PM GN	2	99453680	59才	1冊以下	(6年+) 4年から6年	文系っぽいと思う	3	1	1							
149 2018/11/15 8:55:52 PM GN	2	27601948	59才	3冊くらい	(6年+) 13年から15年	文系っぽいと思う	3	2	2							
150 2018/11/15 9:32:22 PM GN	2	89791228	46才	4冊くらい	(6年+) 4年から6年	どちらとも言えない	2	3	3							
151 2018/11/15 9:40:40 PM GN	2	99245334	41才	1冊以下	(6年+) 4年から6年	とっても文系だと思います	3	3	3							
152 2018/11/15 11:57:52 PM G	2	89549417	34才	3冊くらい	(6年+) 13年から15年	とっても文系だと思います	2	1	1							
153 2018/11/16 12:33:09 AM G	2	58919690	67才	1冊以下	(6年+) 7年から10年	理系っぽいと思う	1	1	1							
154 2018/11/16 1:47:26 AM G	2	14143023	26才	7冊くらい	(6年+) 11年から12年	理系っぽいと思う	3	3	3							
155 2018/11/16 3:15:41 AM G	2	18055472	33才	1冊以下	(6年+) 7年から10年	文系っぽいと思う	1	1	1							
156 2018/11/16 3:56:49 AM G	2	23161987	48才	1冊以下	(6年+) 4年から6年	どちらとも言えない	3	3	3							
157 2018/11/16 8:09:47 AM G	2	53866897	56才	1冊以下	(6年+) 4年から6年	とっても文系だと思います	2	2	2							
158 2018/11/16 8:16:00 AM G	2	76301573	27才	1冊以下	(6年+) 7年から10年	文系っぽいと思う	3	3	3							
159 2018/11/16 9:41:27 AM G	2	60521880	58才	1冊以下	(6年+) 13年から15年	とっても文系だと思います	3	3	3							
160 2018/11/16 6:39:26 PM GN	2	36452322	24才	1冊以下	(6年+) 7年から10年	とっても文系だと思います	2	3	3							
161 2018/11/16 8:08:05 PM GN	2	43530403	31才	10冊以上	(6年+) 4年から6年	理系っぽいと思う	3	3	3							
162 2018/11/17 10:31:51 PM G	2	39567944	45才	1冊以下	(6年+) 4年から6年	文系っぽいと思う	3	3	3							
163 2018/11/19 10:45:45 PM G	2	88436544	29才	1冊以下	(6年+) 7年から10年	文系っぽいと思う	2	1	1							
164 2018/11/20 6:51:56 PM GN	2	43778840	61才	1冊以下	(6年+) 7年から10年	文系っぽいと思う	3	0	0							
165 2018/11/20 7:55:07 PM GN	2	70810256	42才	1冊以下	(6年+) 7年から10年	理系っぽいと思う	3	3	3							
166 2018/11/20 10:35:09 PM G	2	48161520	27才	1冊以下	(6年+) 11年から12年	どちらとも言えない	3	3	3							
167 2018/11/22 2:21:56 PM GN	2	52793046	59才	10冊以上	(6年+) 11年から12年	とっても理系だと思います	3	3	3							
168 2018/11/23 10:58:43 AM G	2	61516069	54才	3冊くらい	(6年+) 7年から10年	文系っぽいと思う	3	2	2							
169 2018/11/23 12:43:50 PM G	2	57025183	64才	1冊以下	(6年+) 7年から10年	文系っぽいと思う	3	1	1							
170 2018/11/23 2:23:52 PM GN	2	83809383	57才	1冊以下	(6年+) 7年から10年	とっても文系だと思います	1	1	1							
171 2018/11/23 7:14:27 PM GN	2	46674200	52才	1冊以下	(6年+) 7年から10年	文系っぽいと思う	1	1	1							
172 2018/11/23 7:48:12 PM GN	2	55304939	67才	2冊くらい	(6年+) 7年から10年	文系っぽいと思う	3	1	1							
173 2018/11/23 9:12:49 PM GN	2	57582451	57才	1冊以下	(6年+) 7年から10年	理系っぽいと思う	2	1	1							
174 2018/11/23 10:49:21 PM G	2	45941560	51才	1冊以下	(6年+) 7年から10年	とっても文系だと思います	3	3	3							
175 2018/11/24 1:08:41 AM G	2	51553173	40才	1冊以下	(6年+) 4年から6年	理系っぽいと思う	3	3	3							
176																

Histograms of Stimulus-Wise Responses

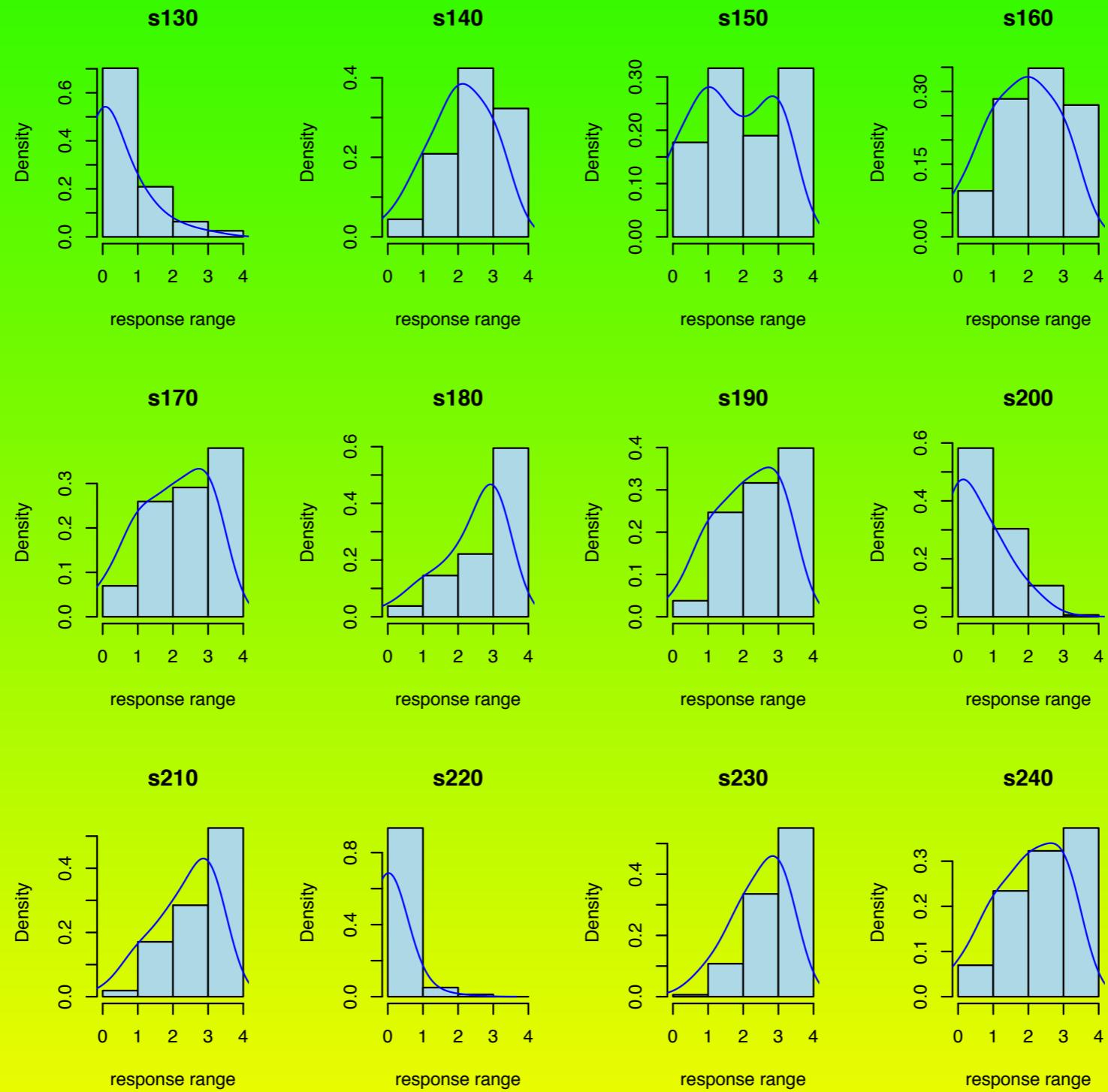
282種類の文に対する評定結果の簡単な分析



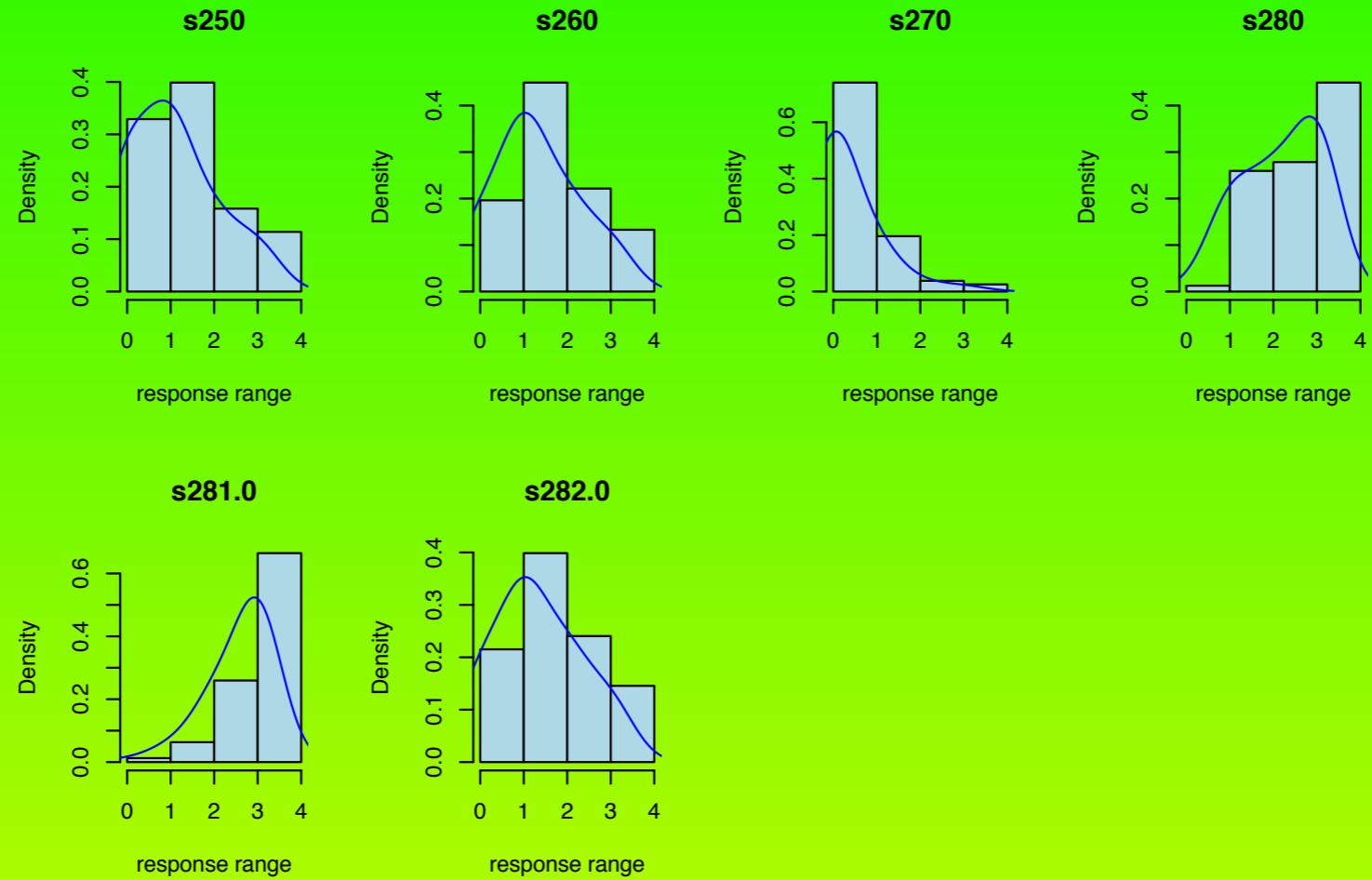
gr0 の s010, s020, ..., s120 [1/3]



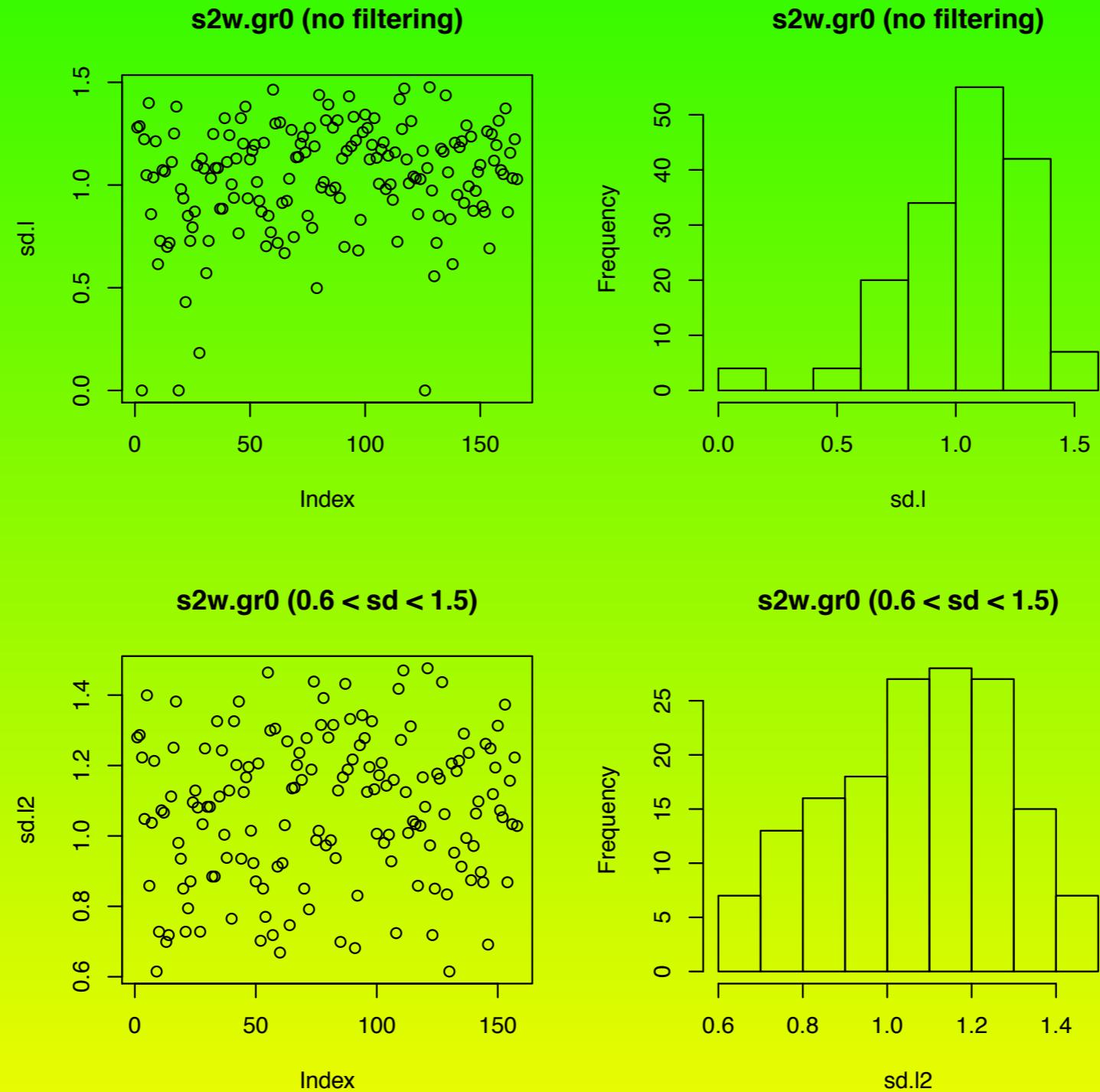
gr0 の s130, s140, ..., s240 [2/3]



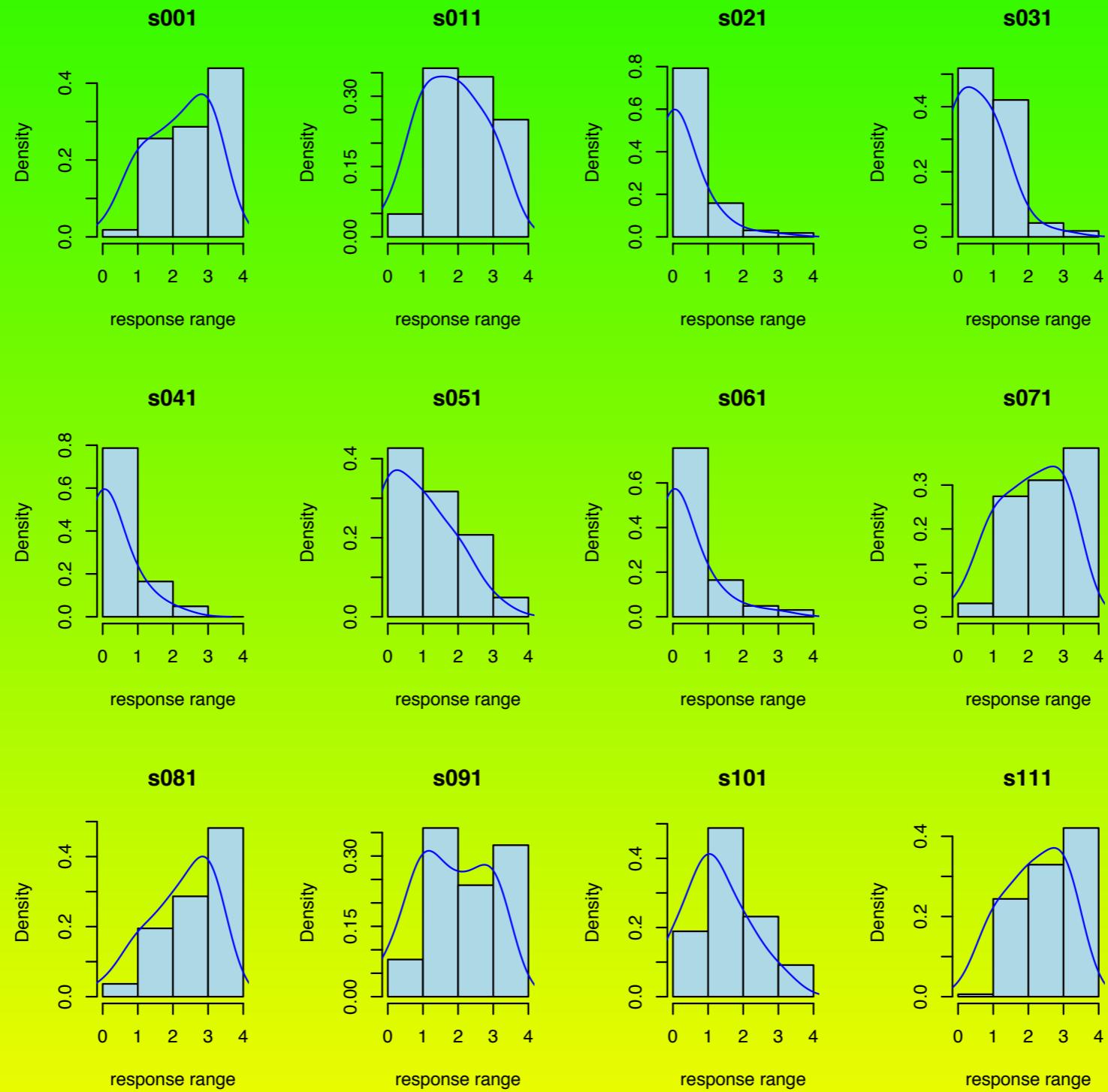
gr0 の s250, s260, ..., s282.0 [3/3]



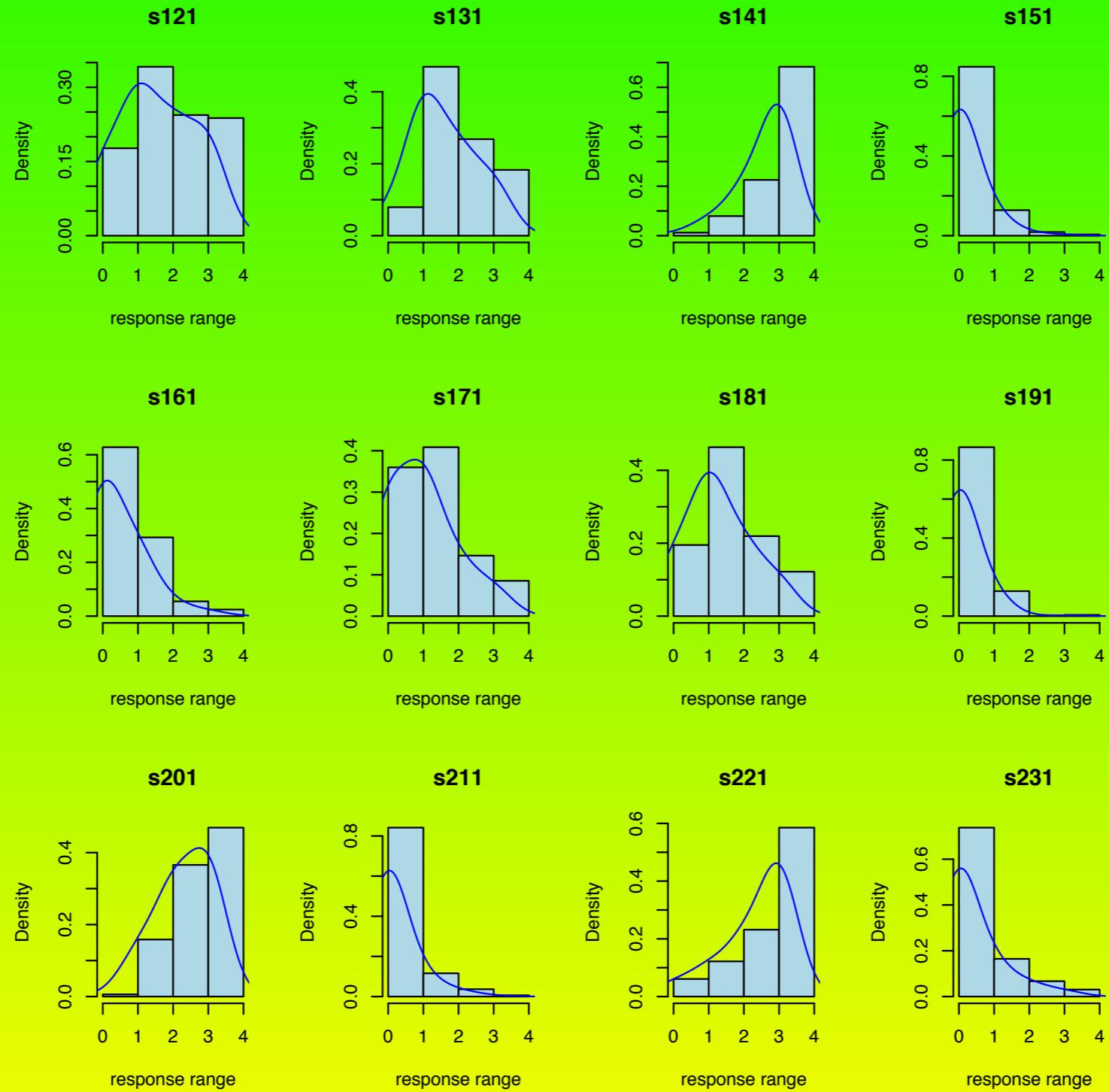
gr0 の s010, s020, … , s281.0, s282.0 の分析の前処理



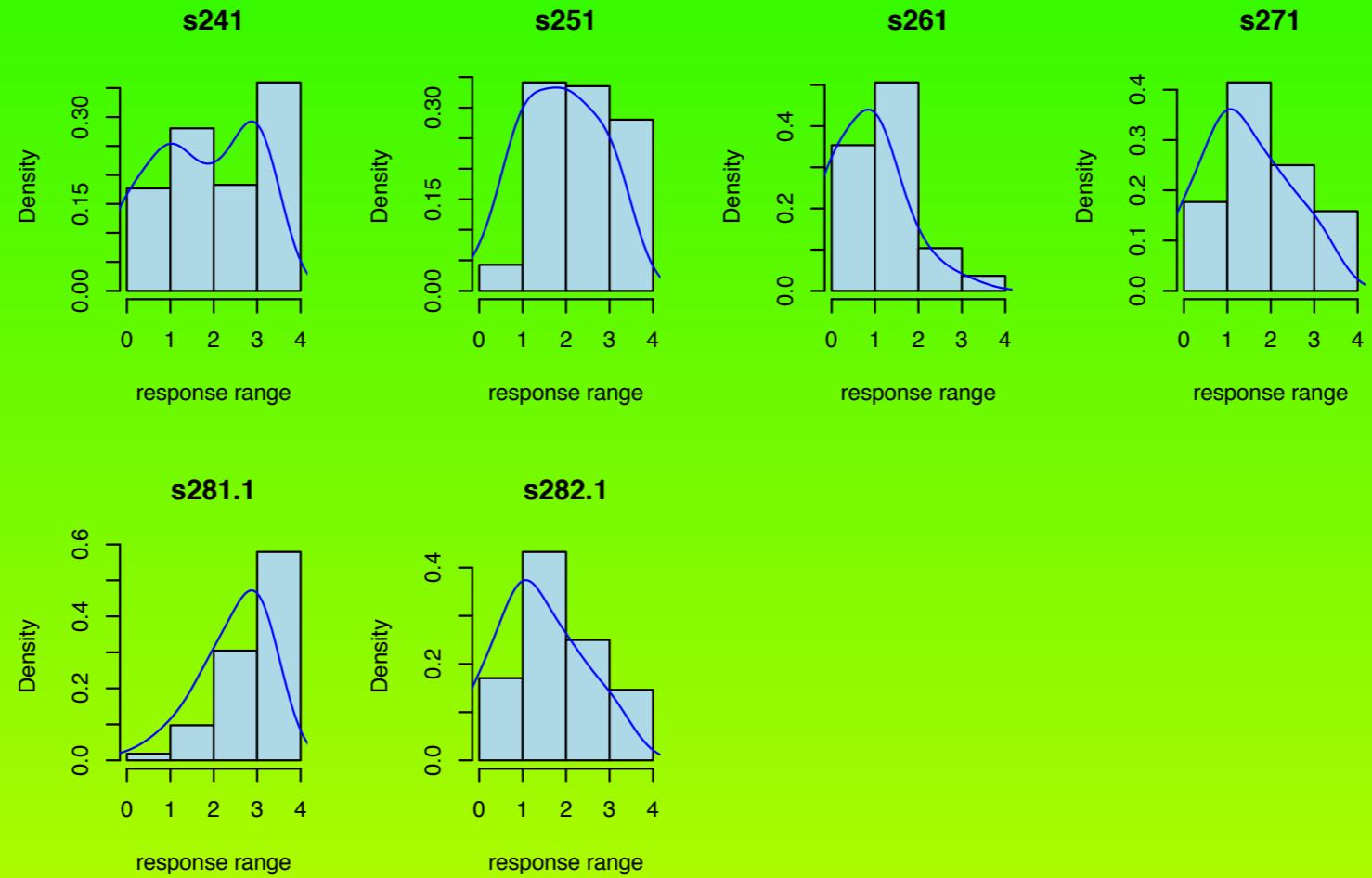
gr1 の s001, s011, ..., s111 [1/3]



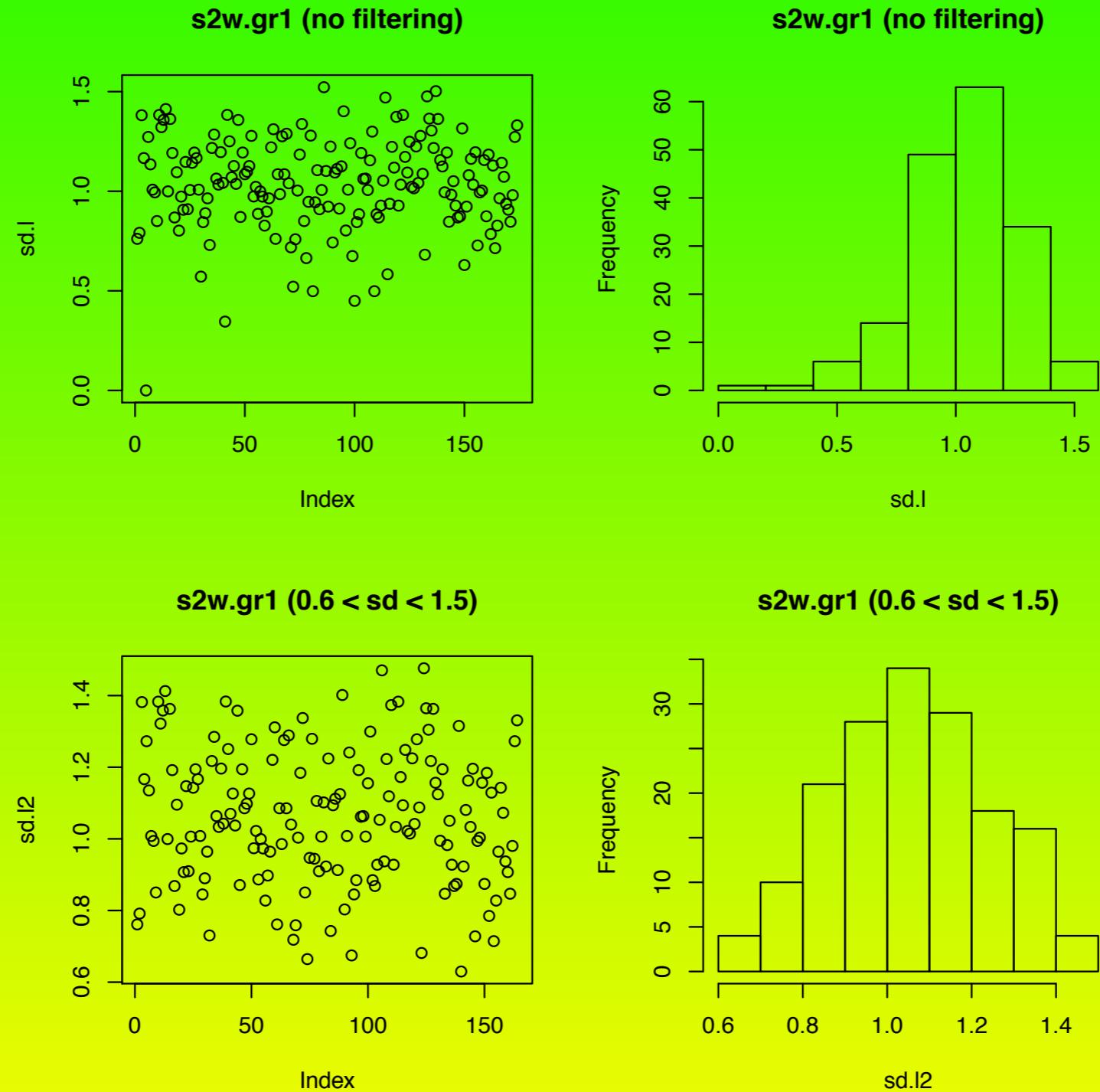
gr1 の s121, s131, ..., s231 [2/3]



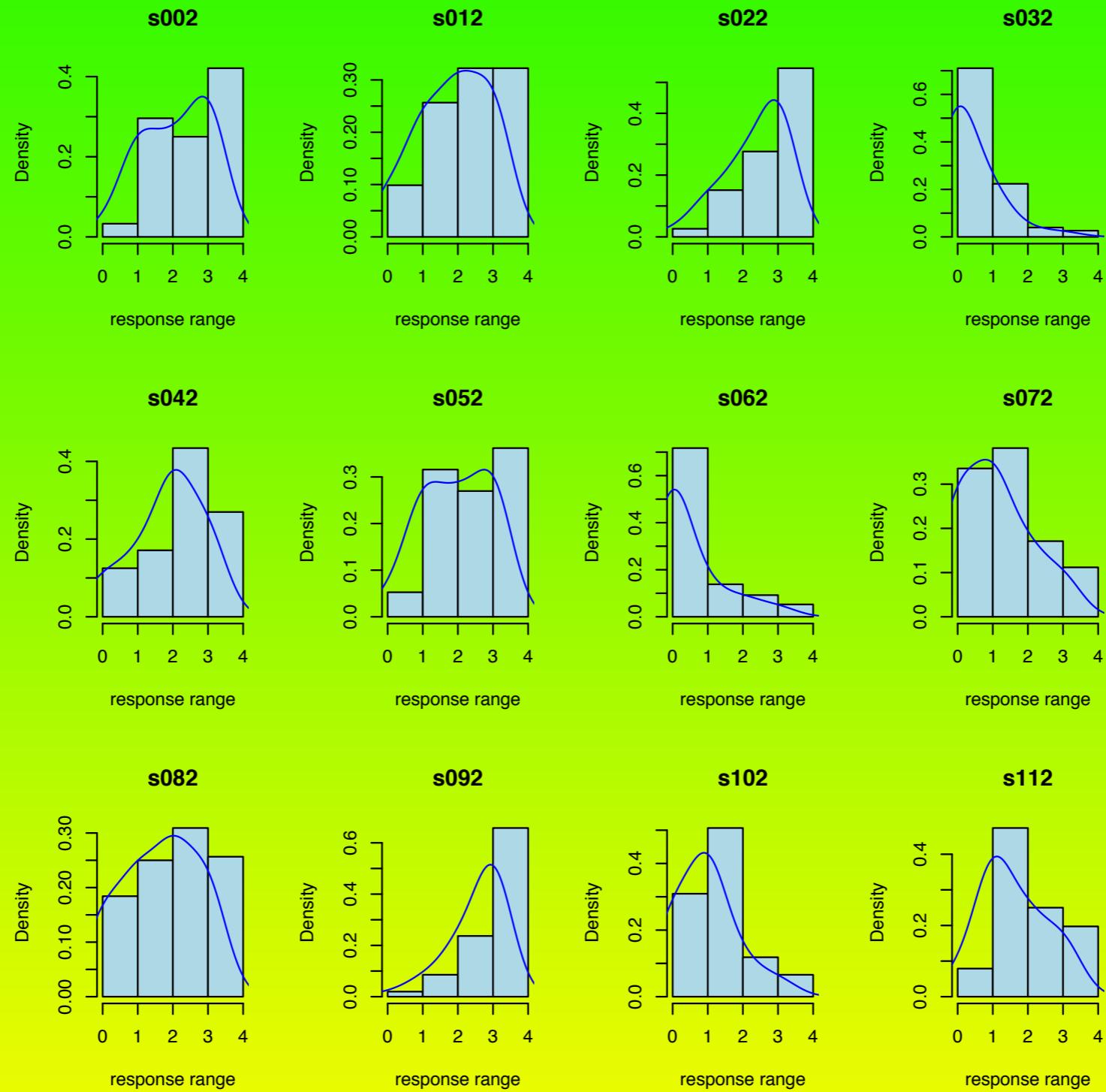
gr1 の s241, s251, ..., s282.1 [3/3]



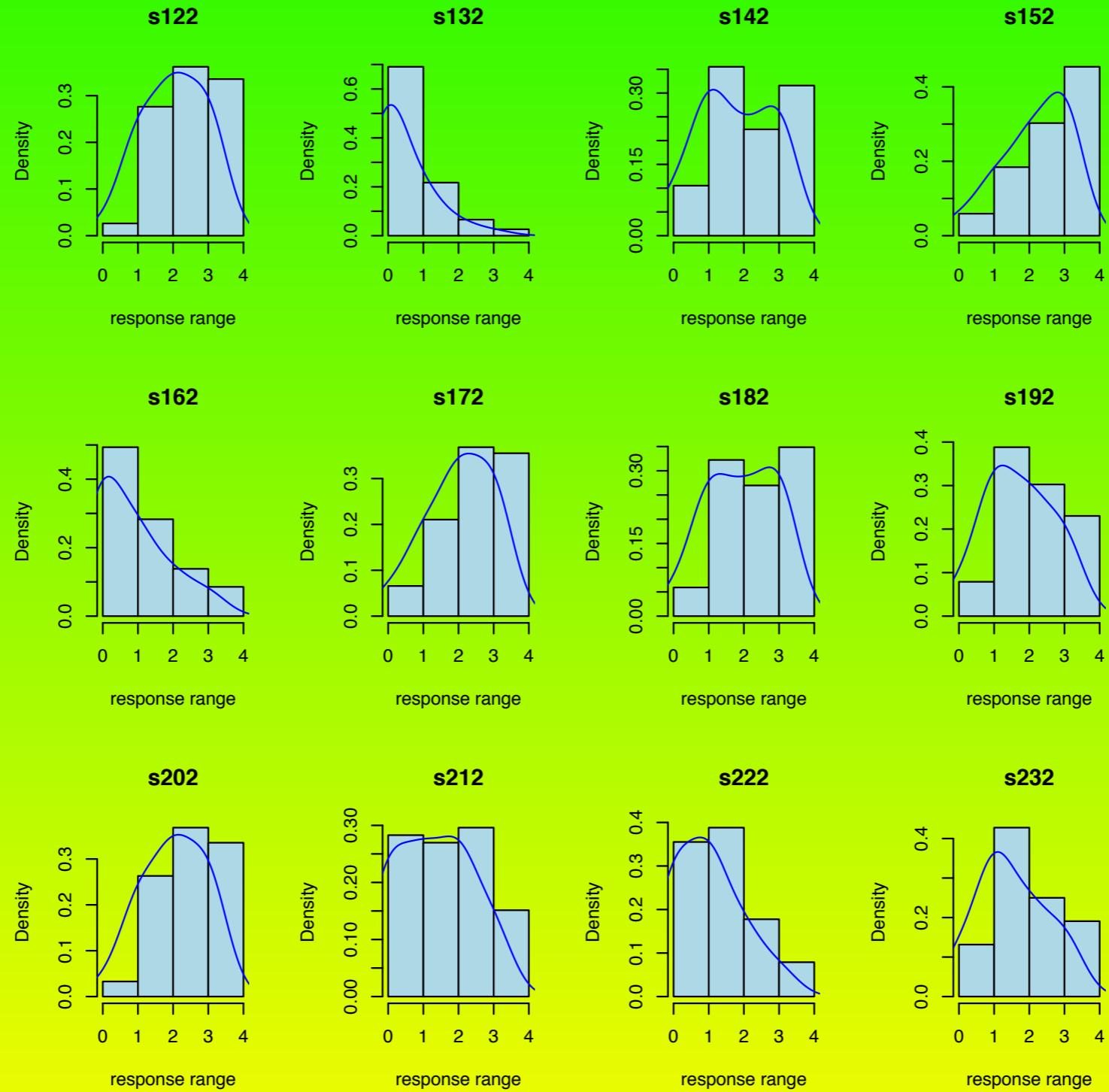
gr1 の s001, s011, … , s281.1, s282.1 の分析の前処理



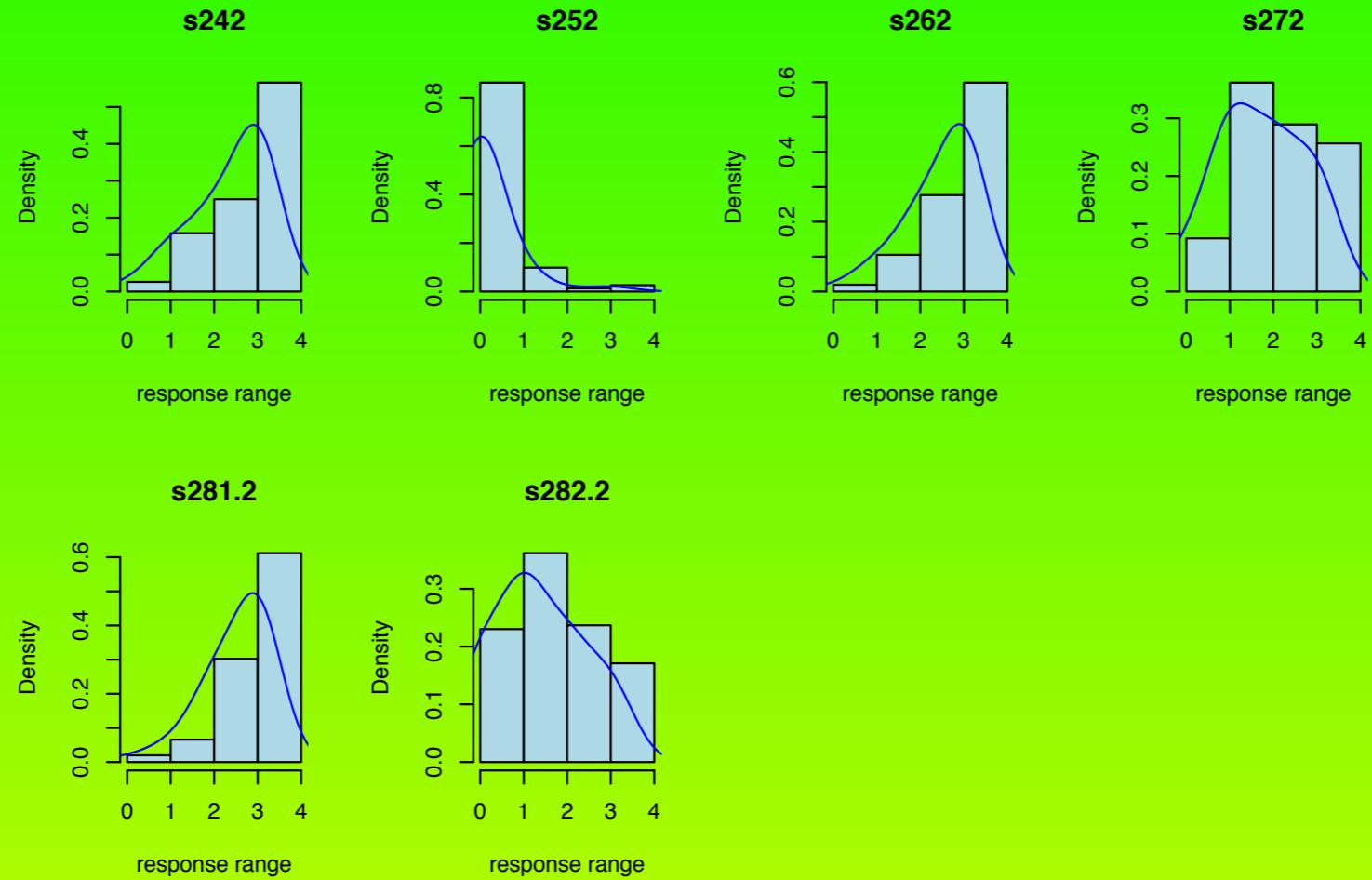
gr2 の s002, s012, ..., s112 [1/3]



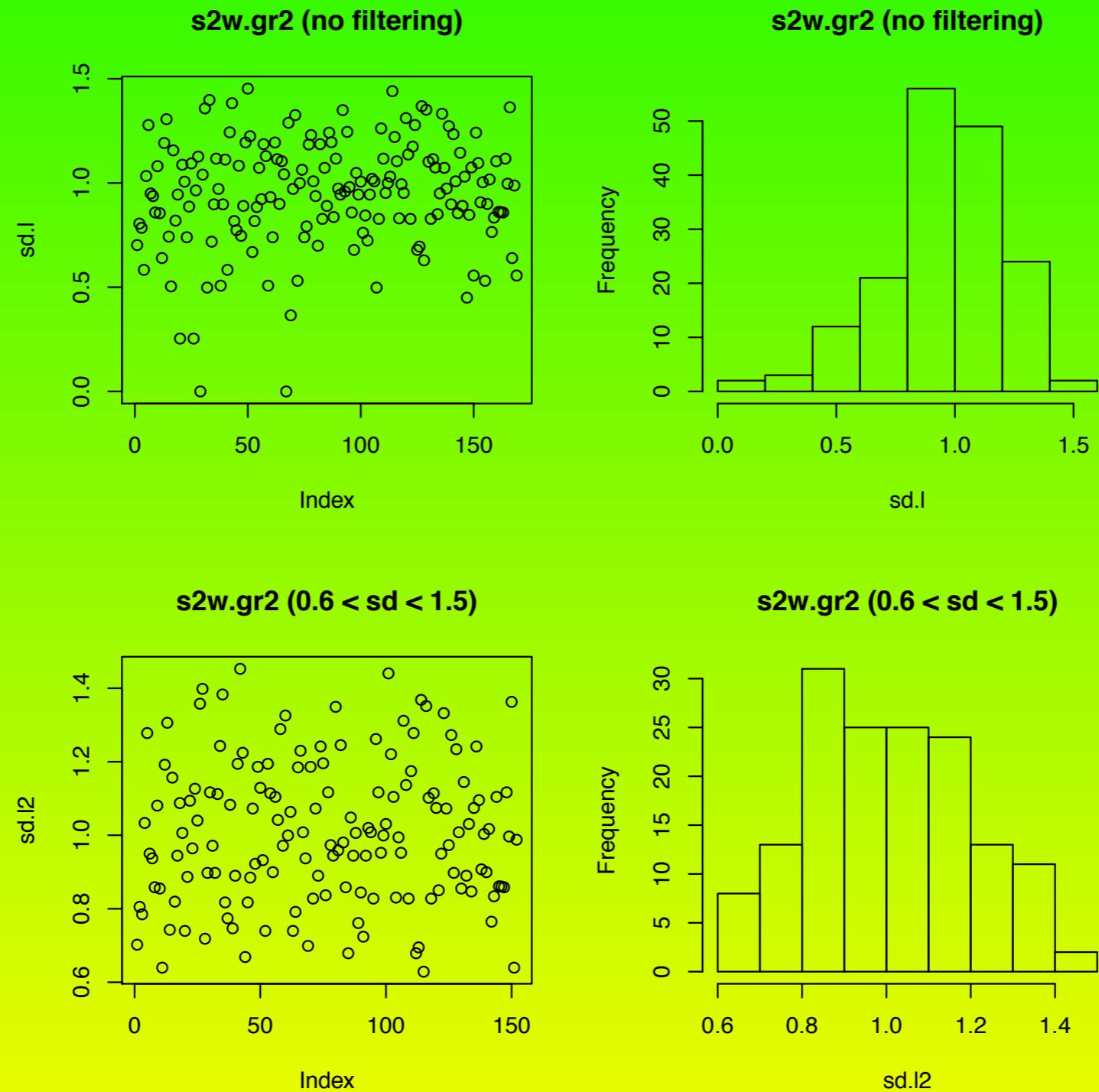
gr2 の s122, s132, ..., s232 [2/3]



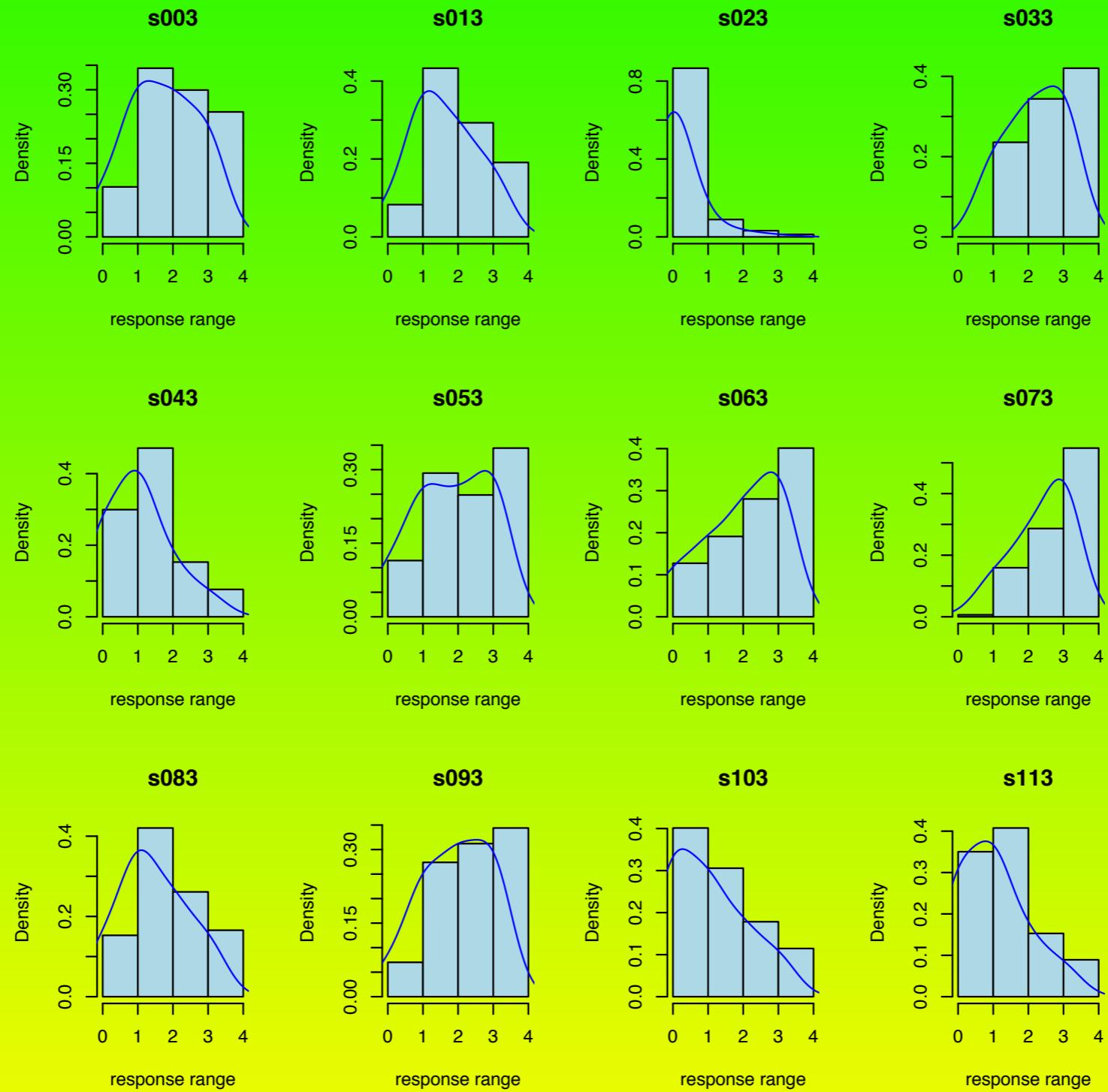
gr2 の s242, s252, ..., s282.2 [3/3]



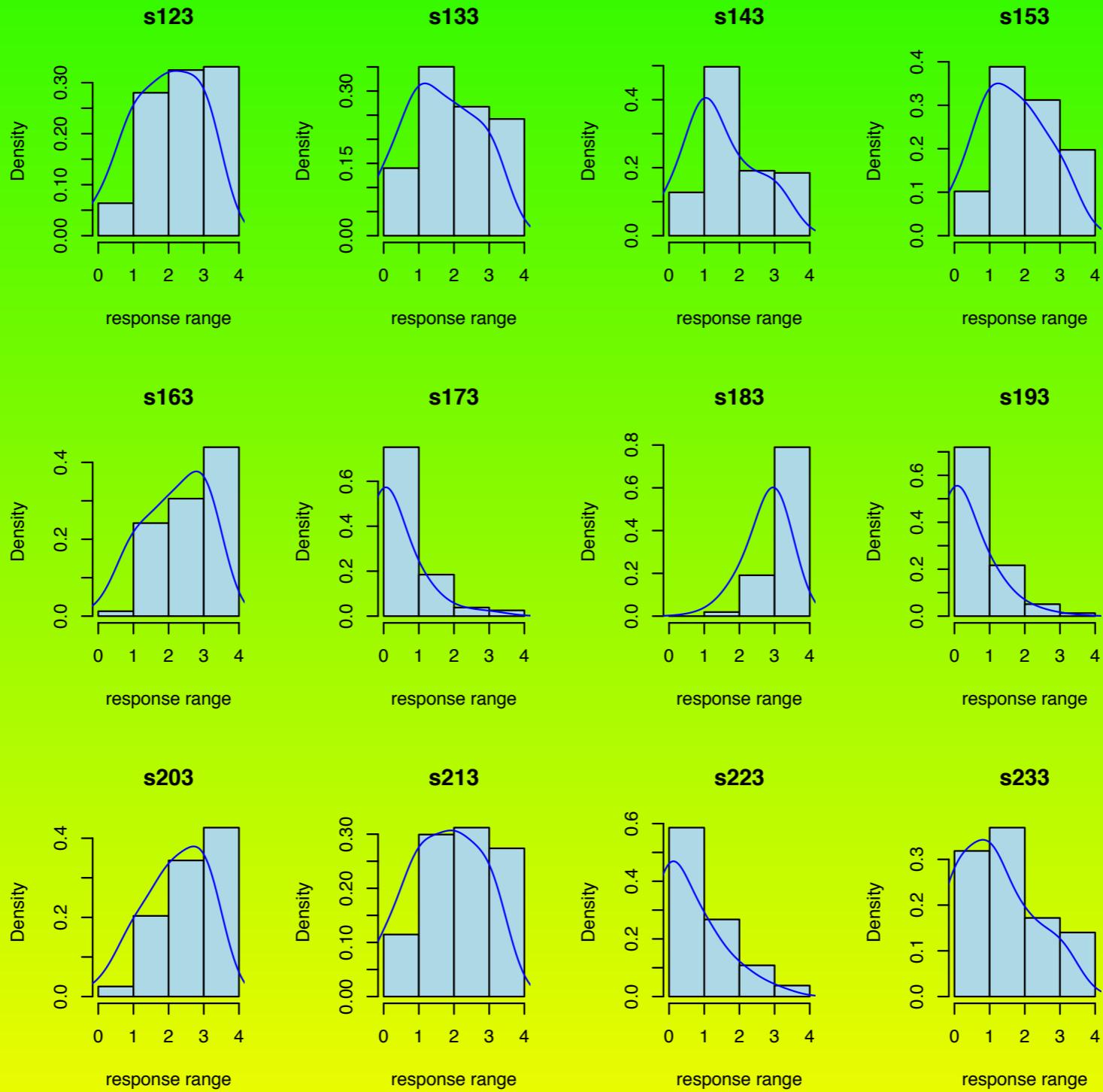
gr2 の s002, s012, … , s281.2, s282.2 の分析の前処理



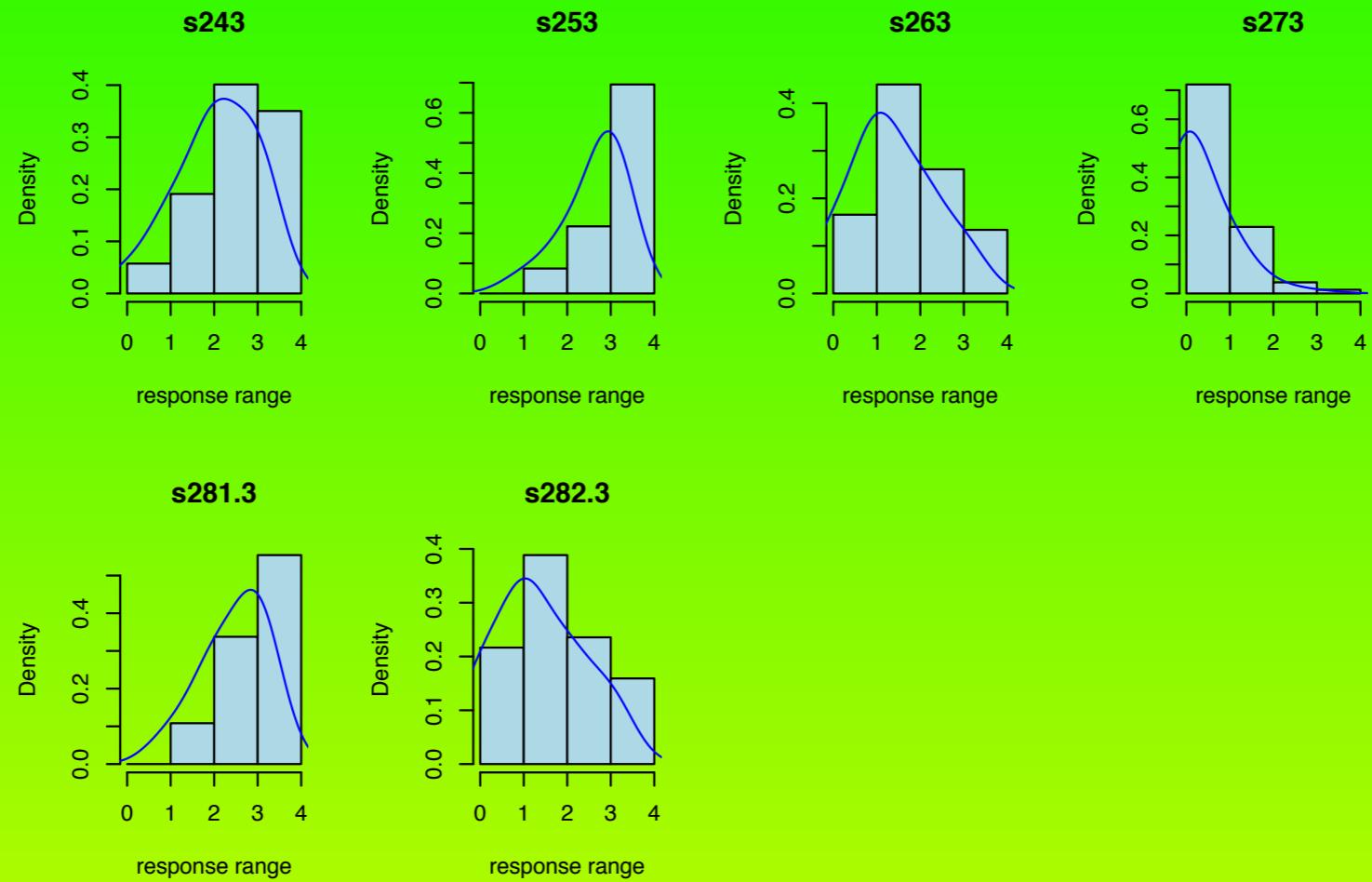
gr3 の s003, s013, ..., s113 [1/3]



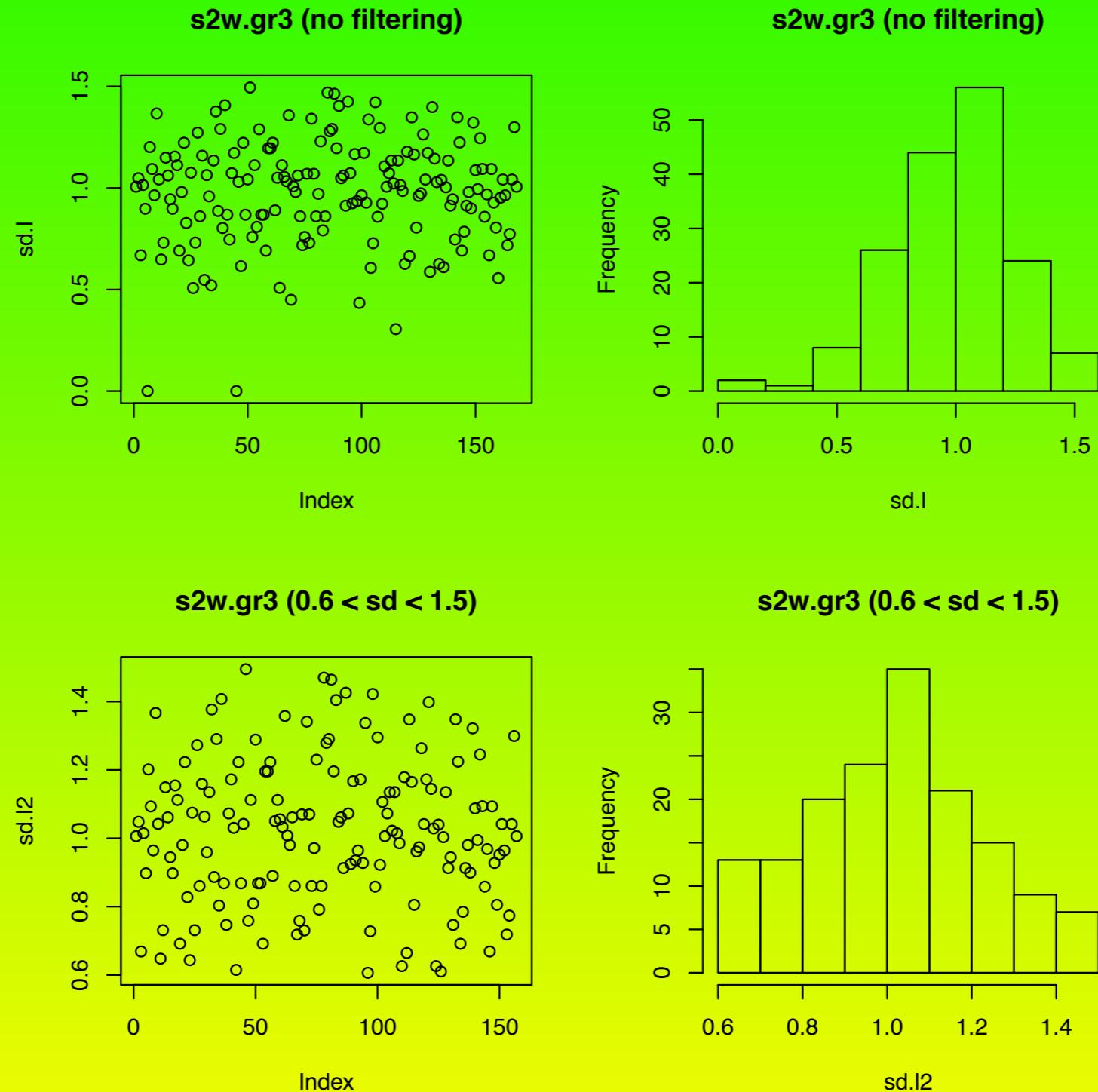
gr3 の s123, s133, ..., s233 [2/3]



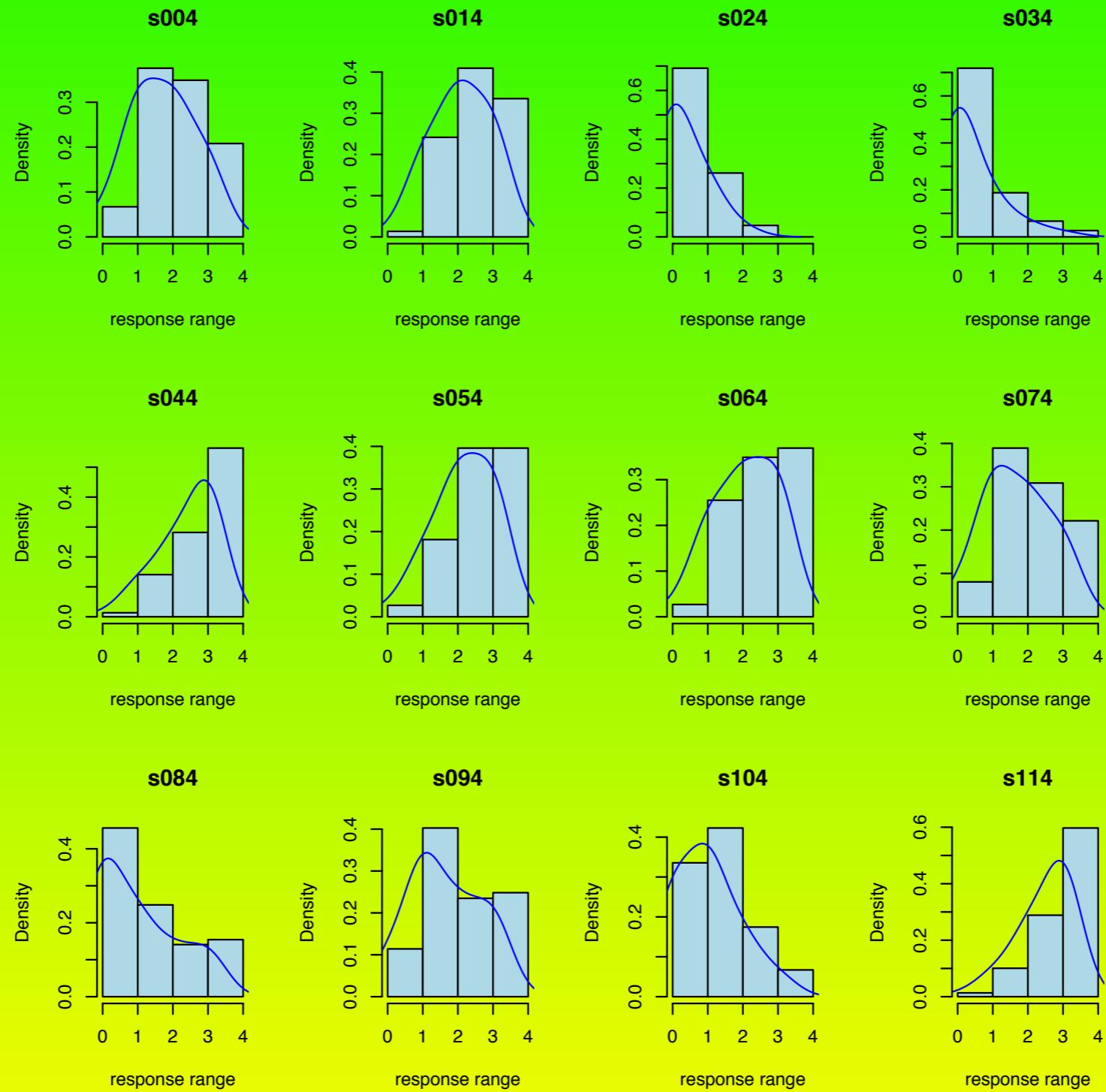
gr3 の s243, s253, ..., s282.3 [3/3]



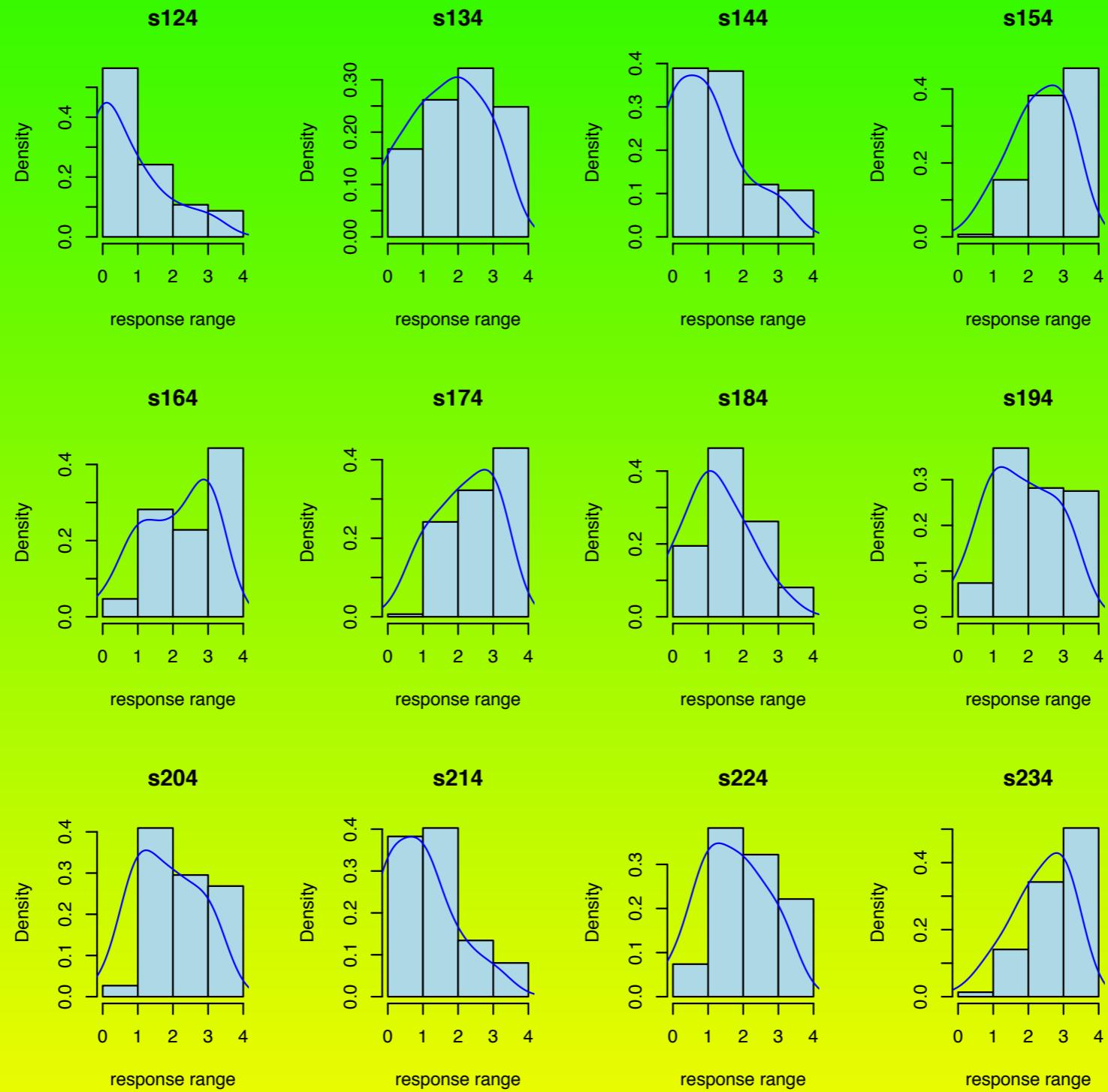
gr3 の s003, s013, … , s281.3, s282.3 の分析の前処理



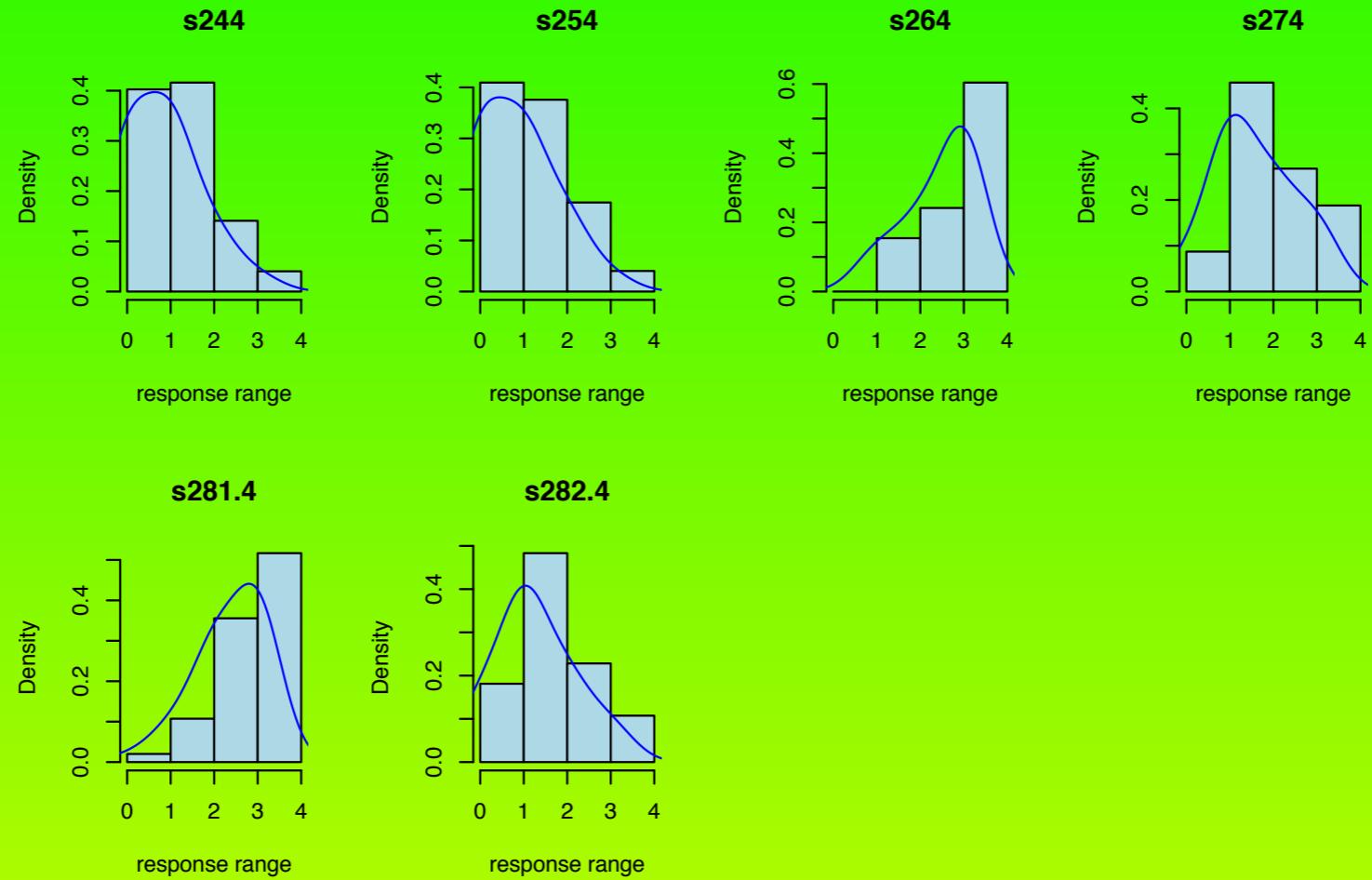
gr4 の s004, s014, ..., s114 [1/3]



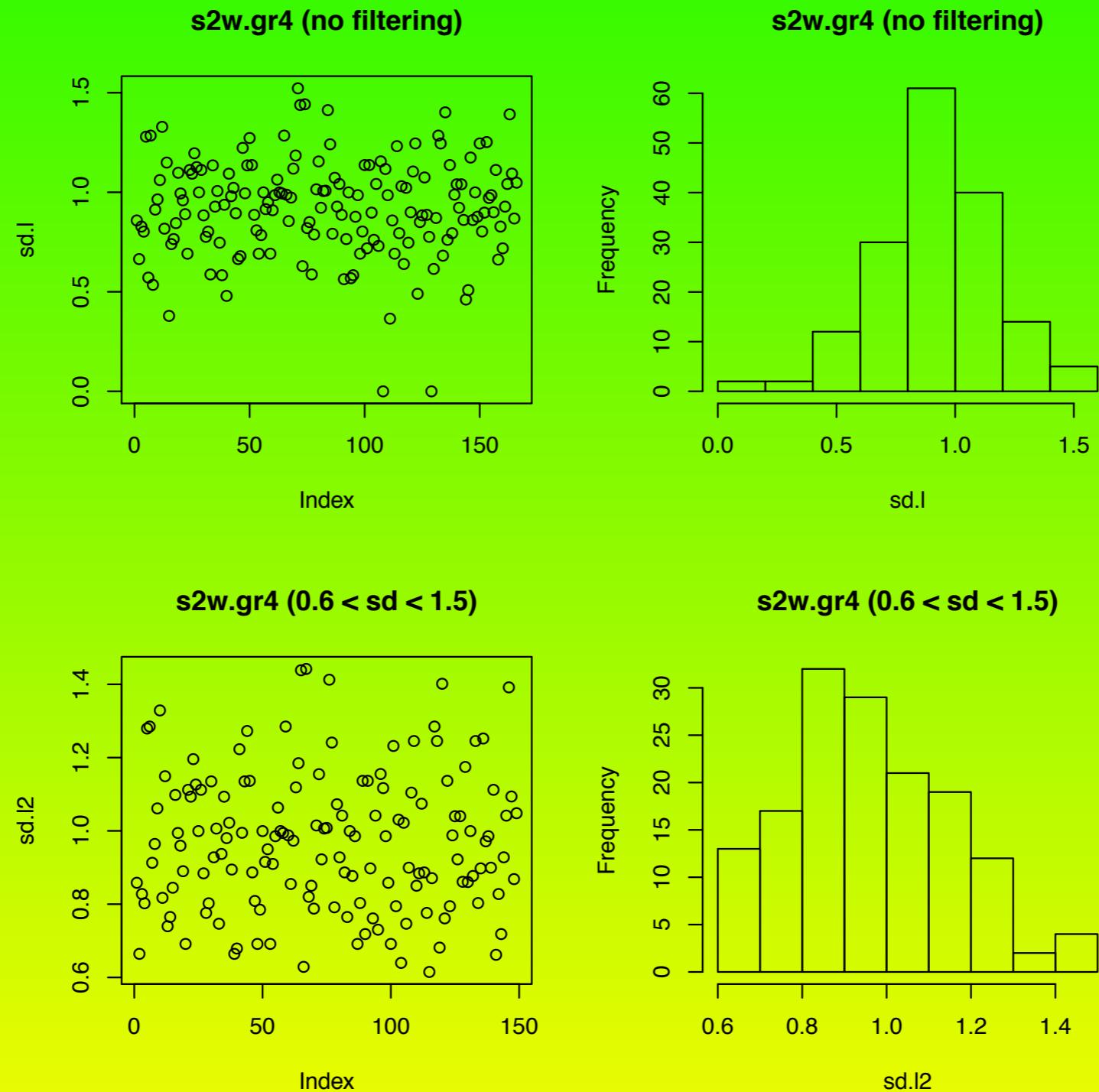
gr4 の $s_{124}, s_{134}, \dots, s_{234}$ [2/3]



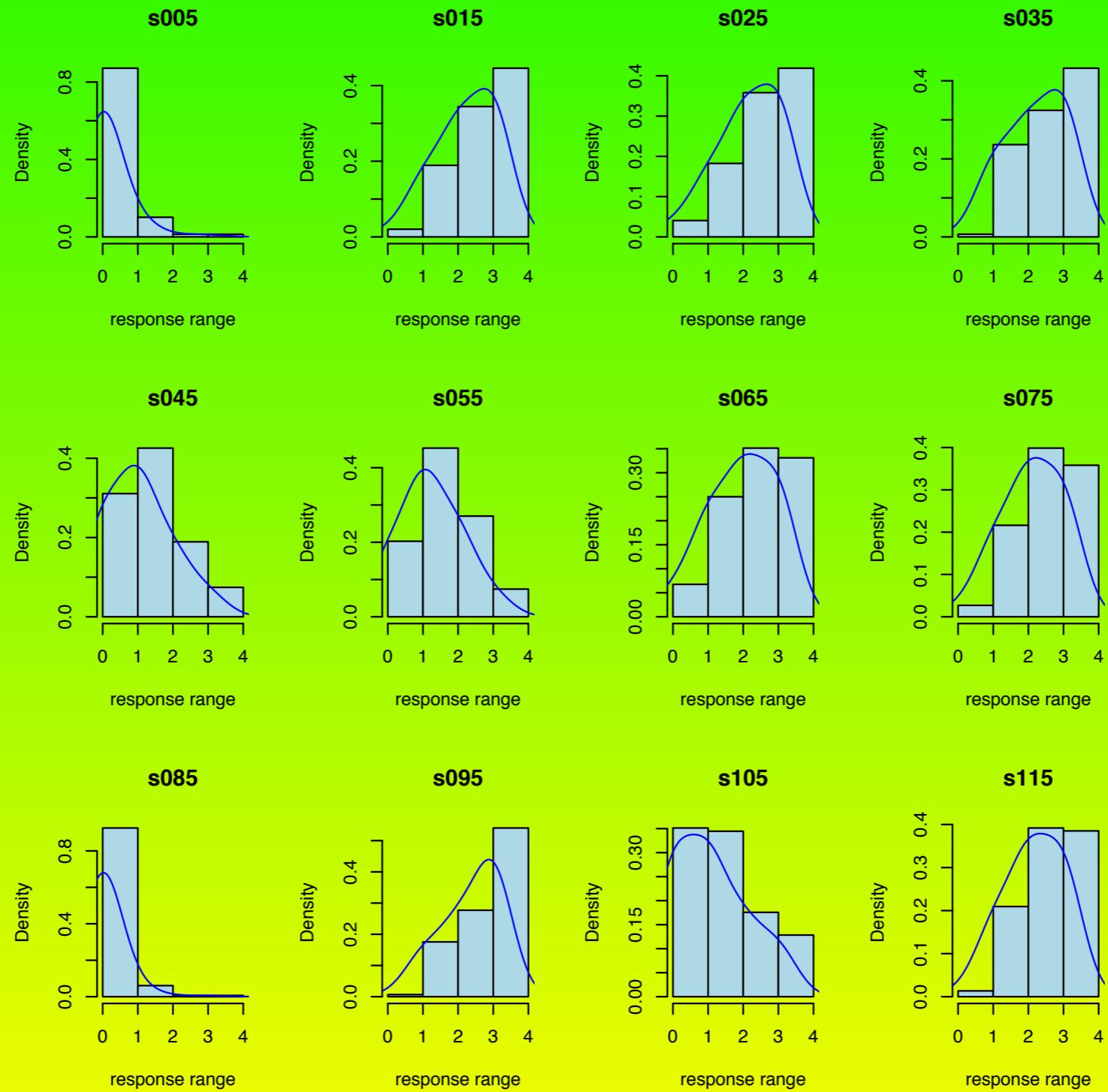
gr4 の s244, s254, ..., s282.4 [3/3]



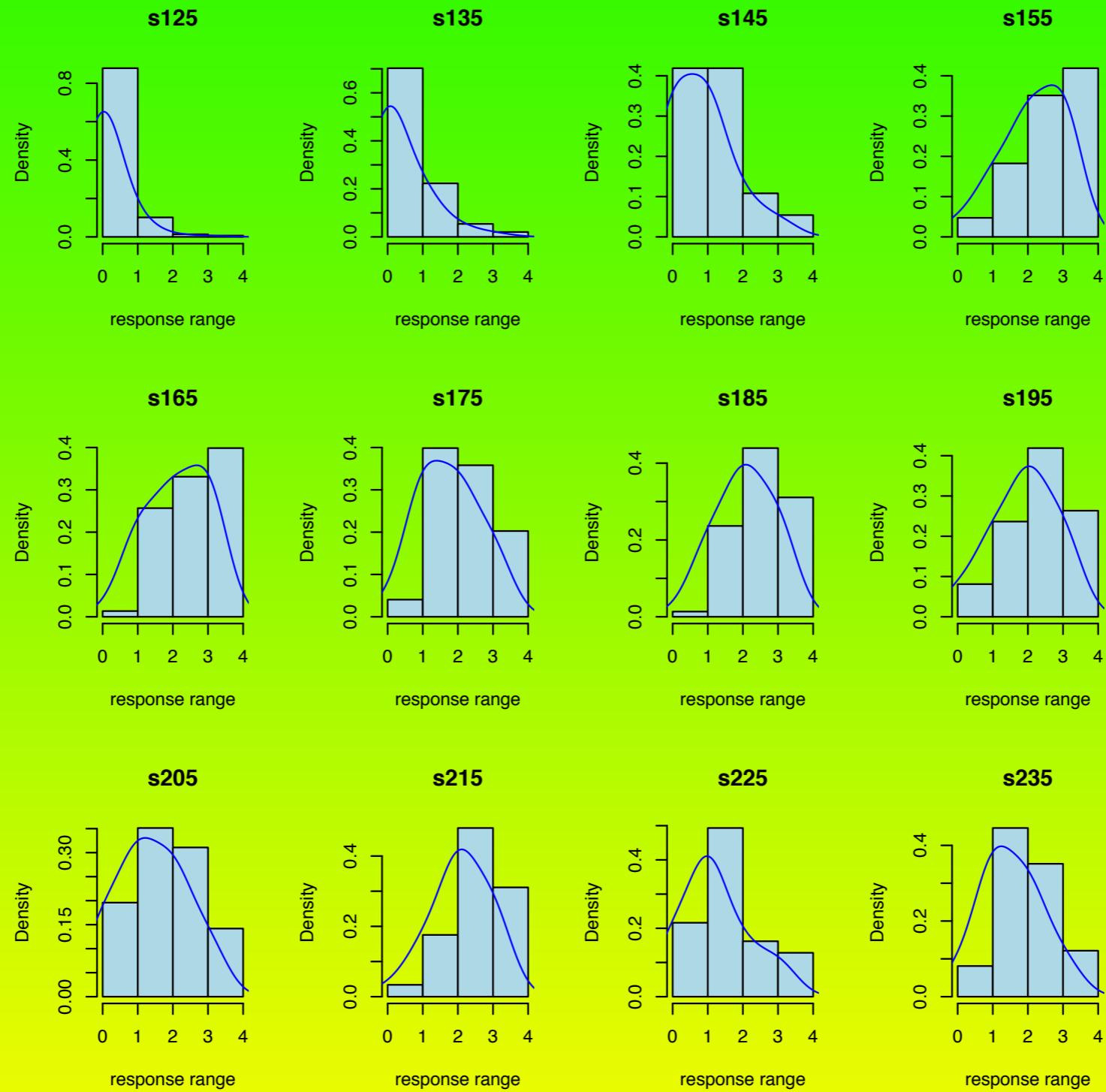
gr4 の s004, s014, … , s281.4, s282.4 の分析の前処理



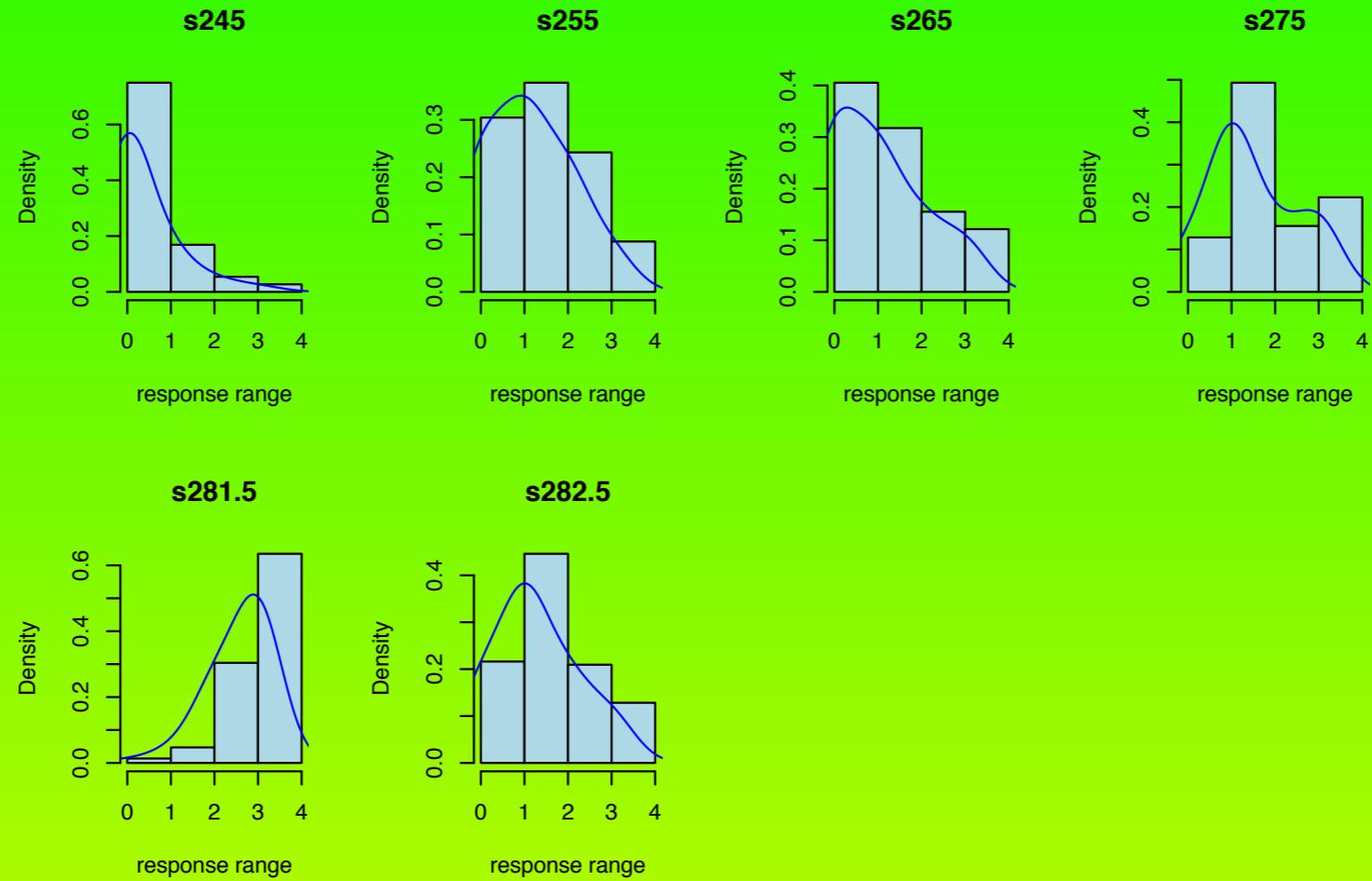
gr5 の s005, s015, ..., s115 [1/3]



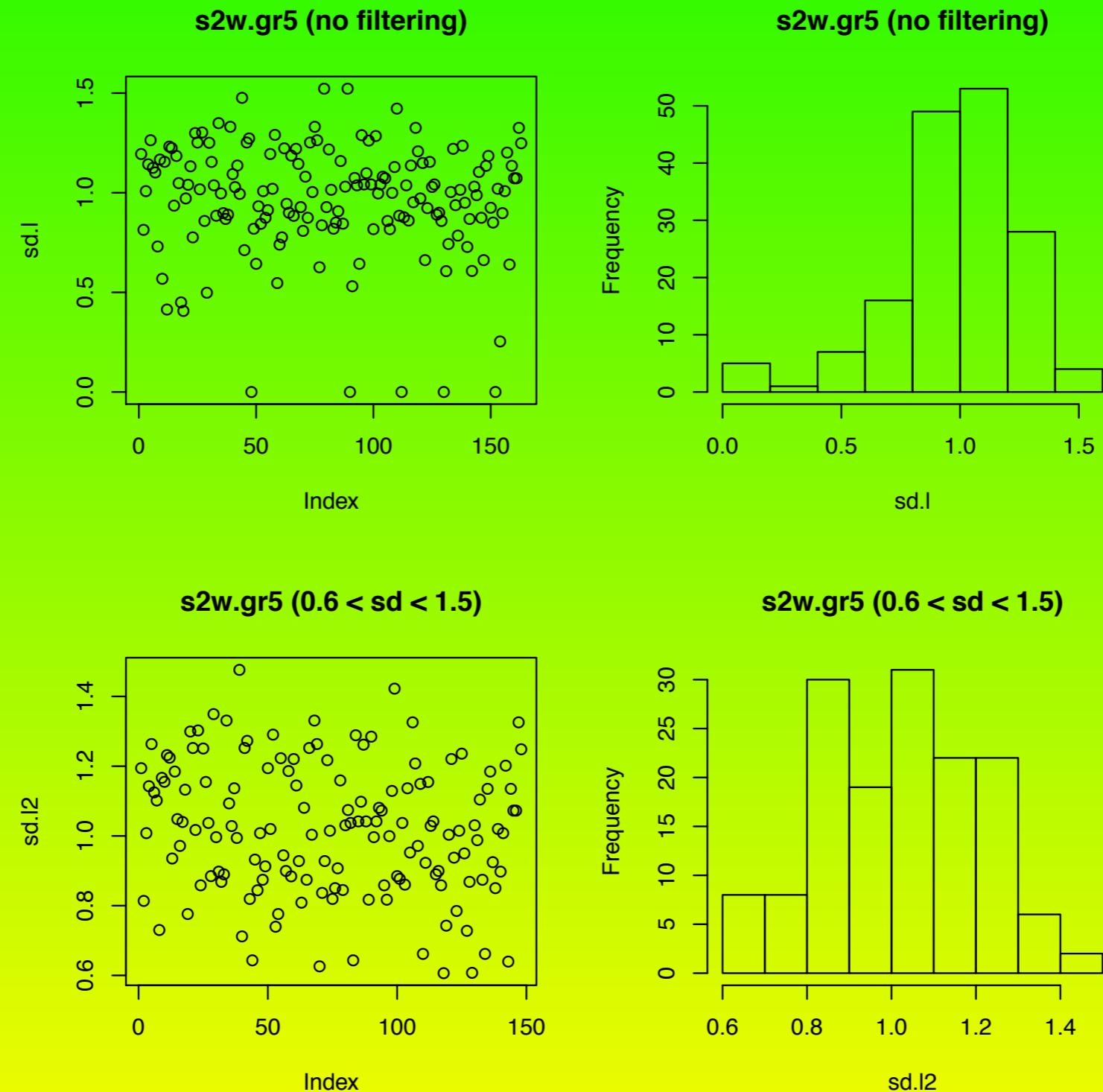
gr5 の s125, s135, ..., s235 [2/3]



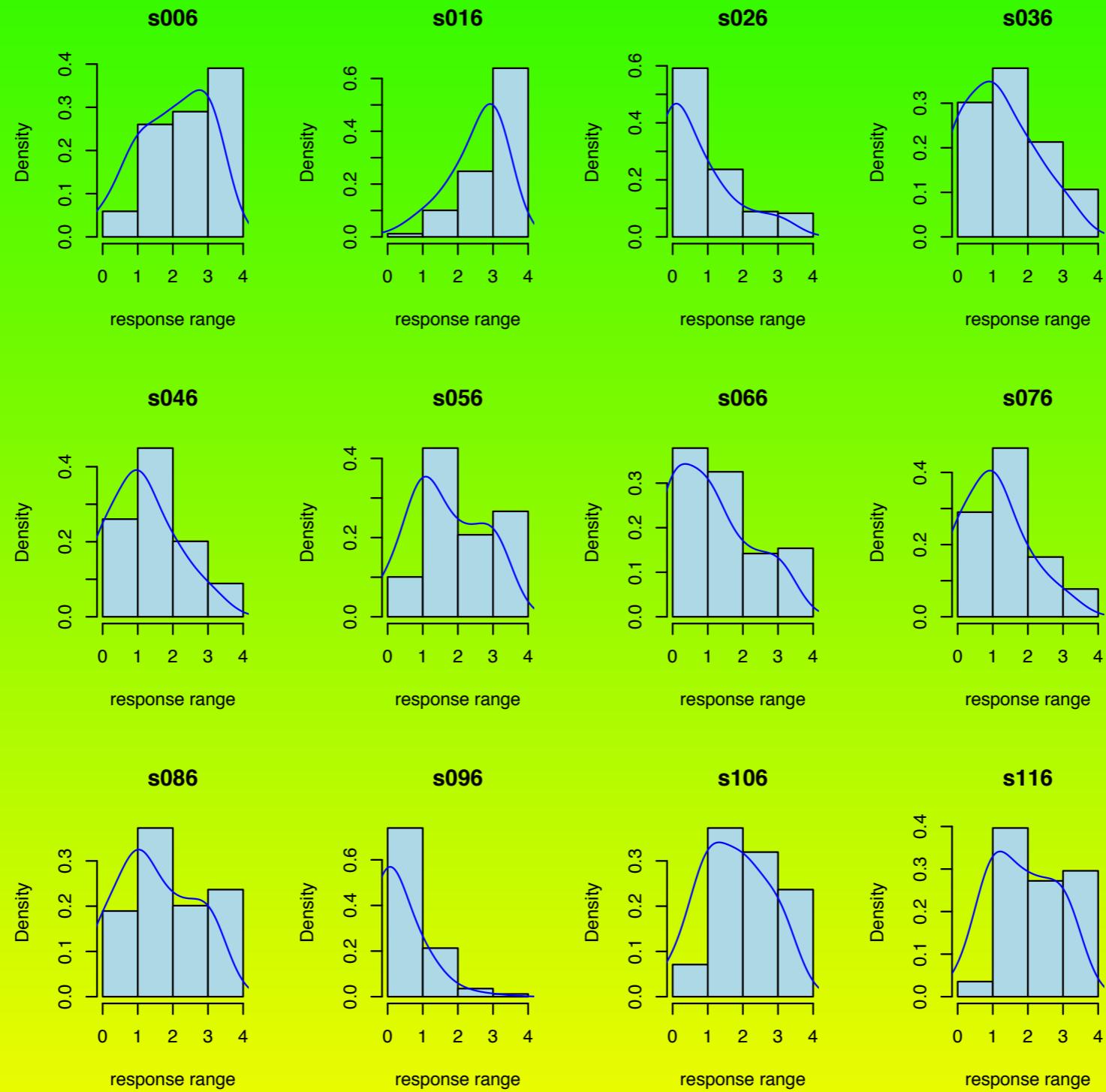
gr5 の s245, s255, ..., s282.5 [3/3]



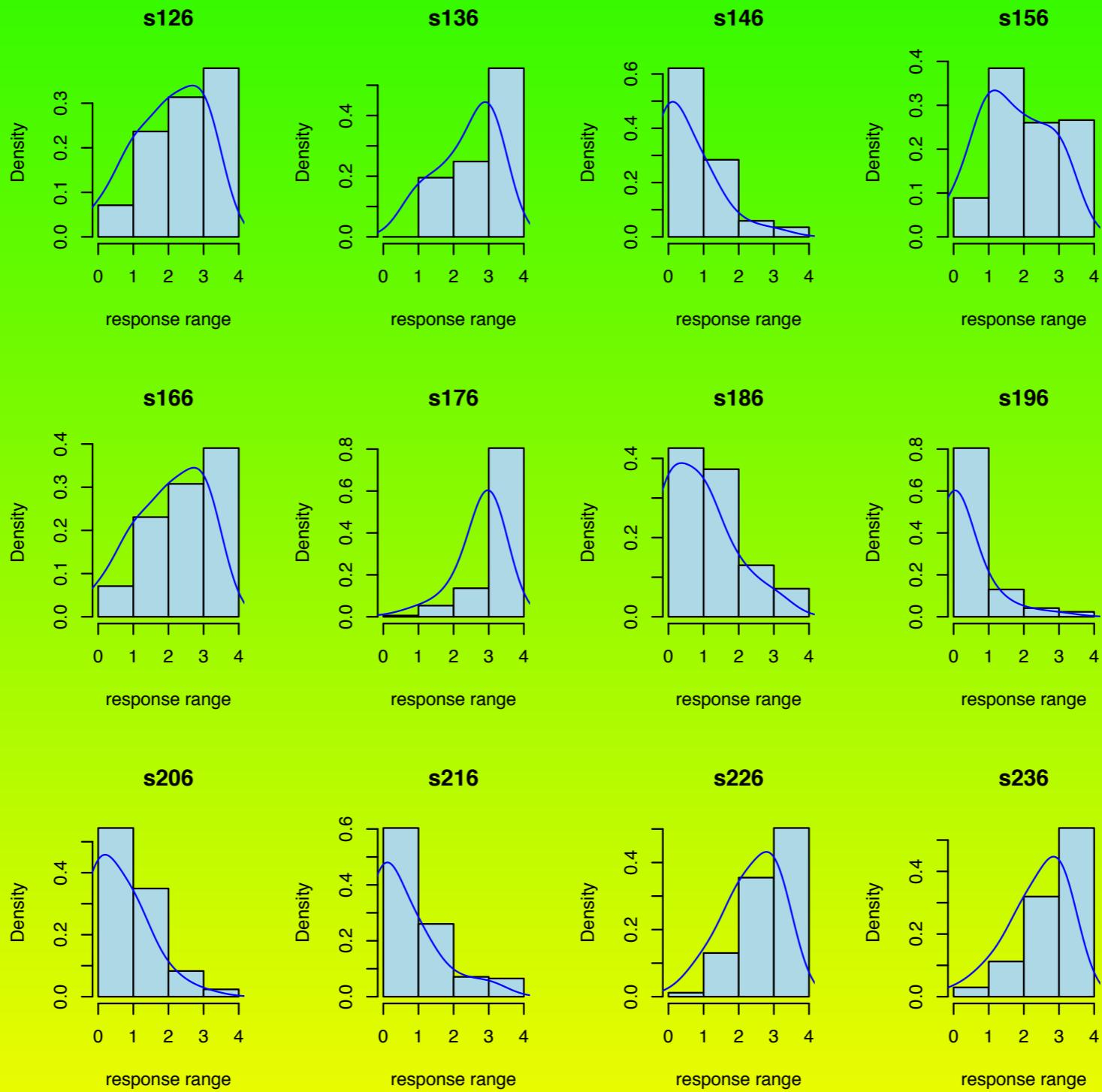
gr5 の s005, s015, … , s281.5, s282.5 の分析の前処理



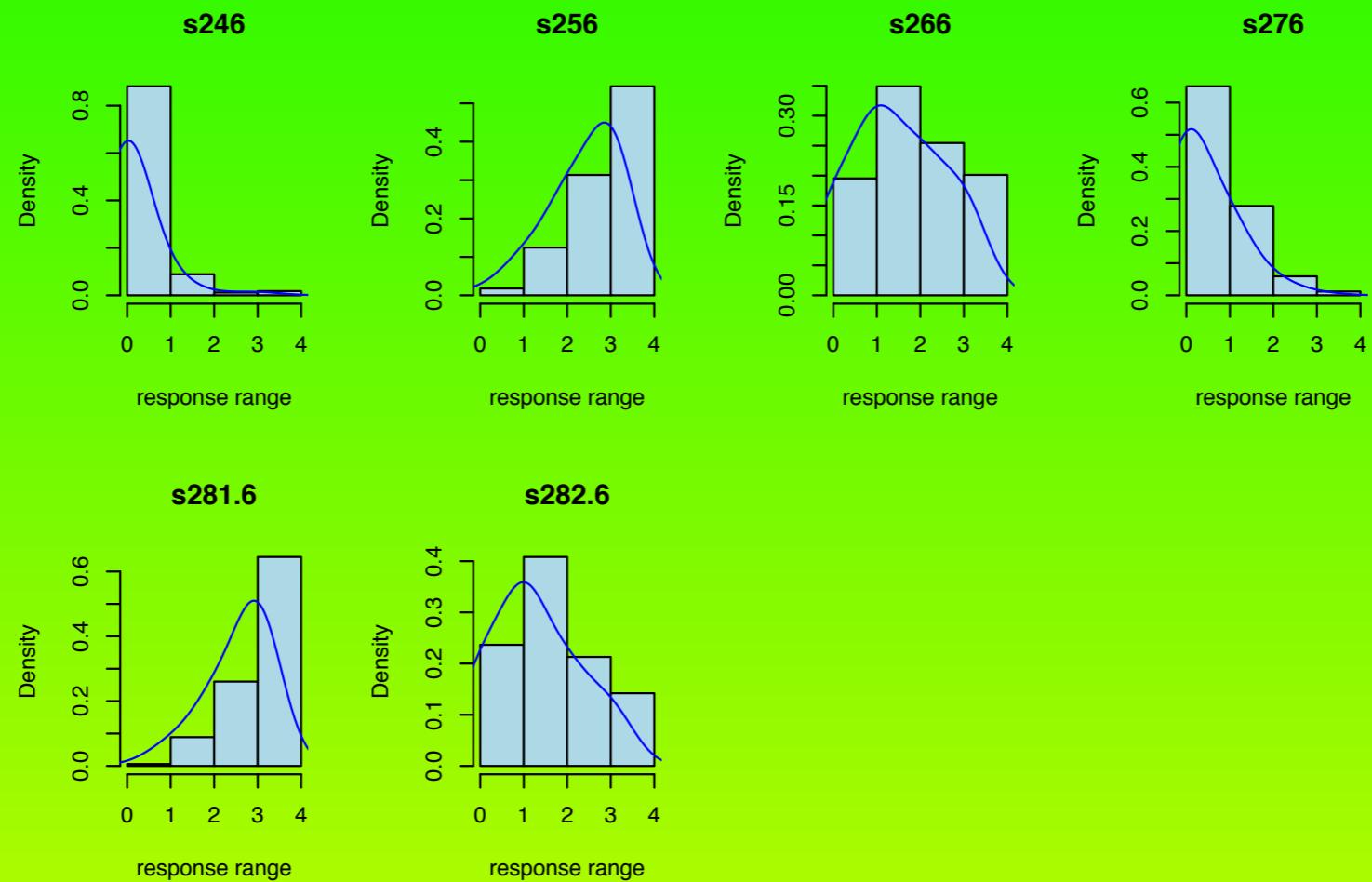
gr6 の s006, s016, ..., s116 [1/3]



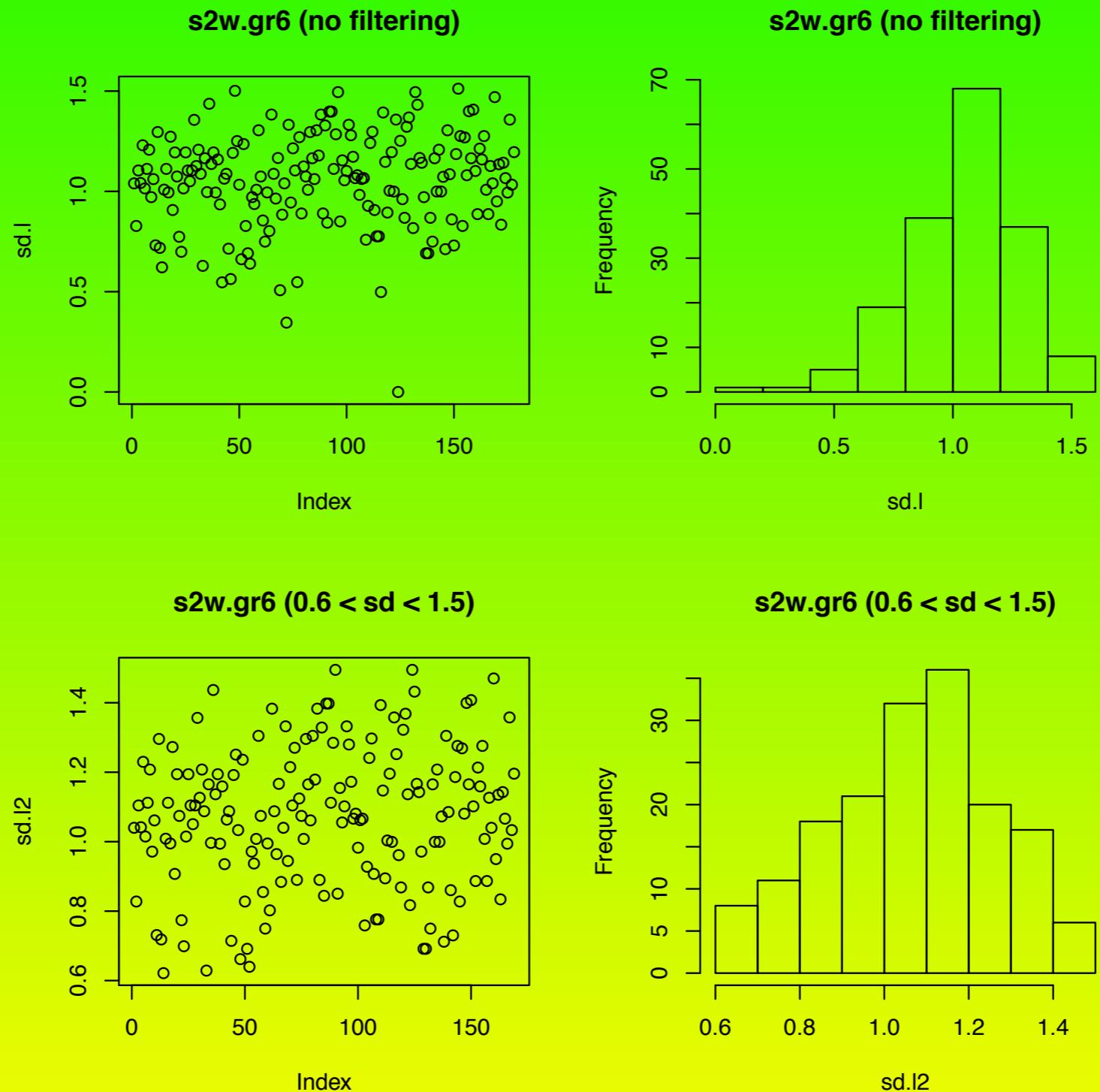
gr6 の s126, s136, ..., s236 [2/3]



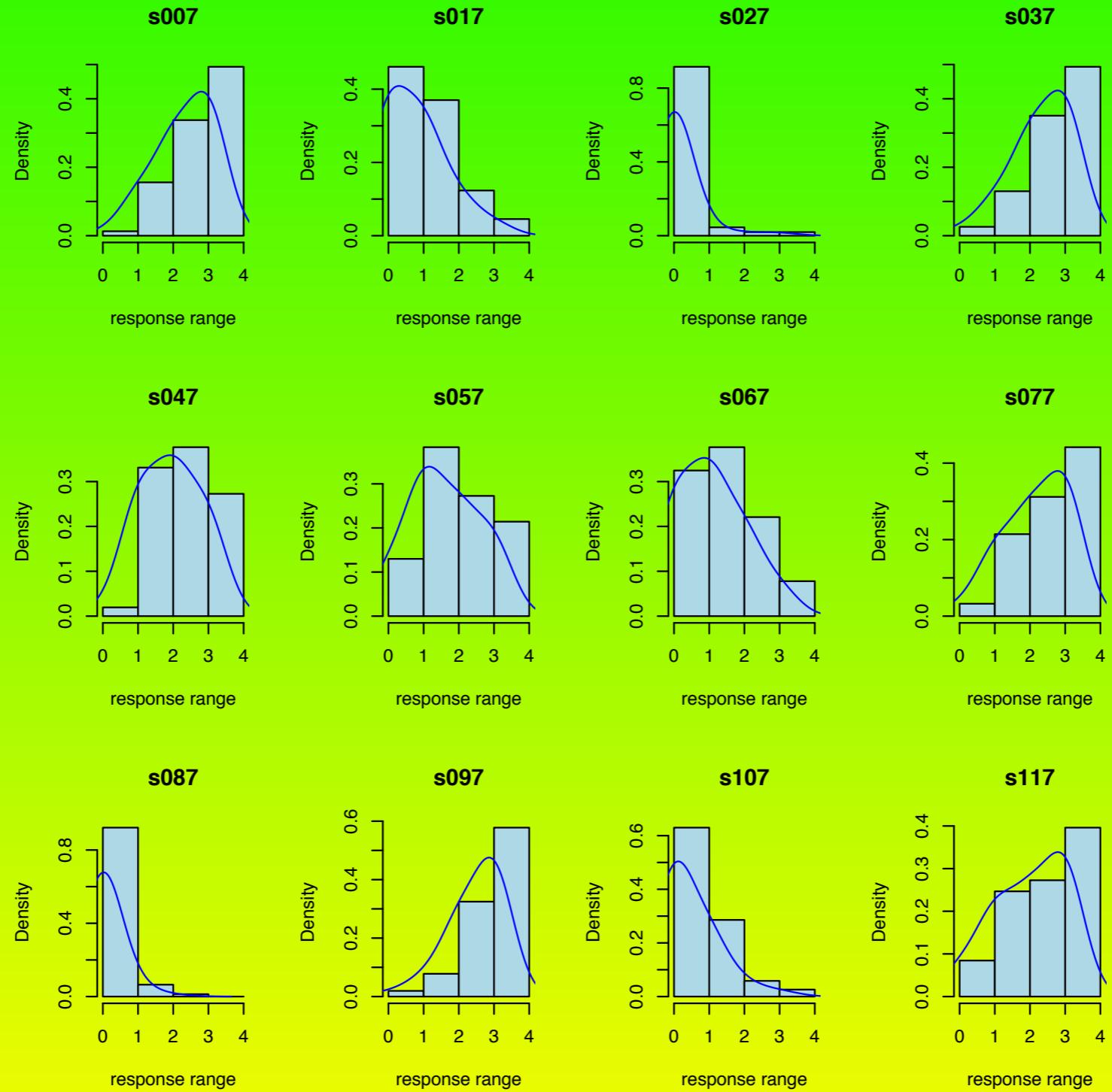
gr6 の s246, s136, ..., s282.6 [3/3]



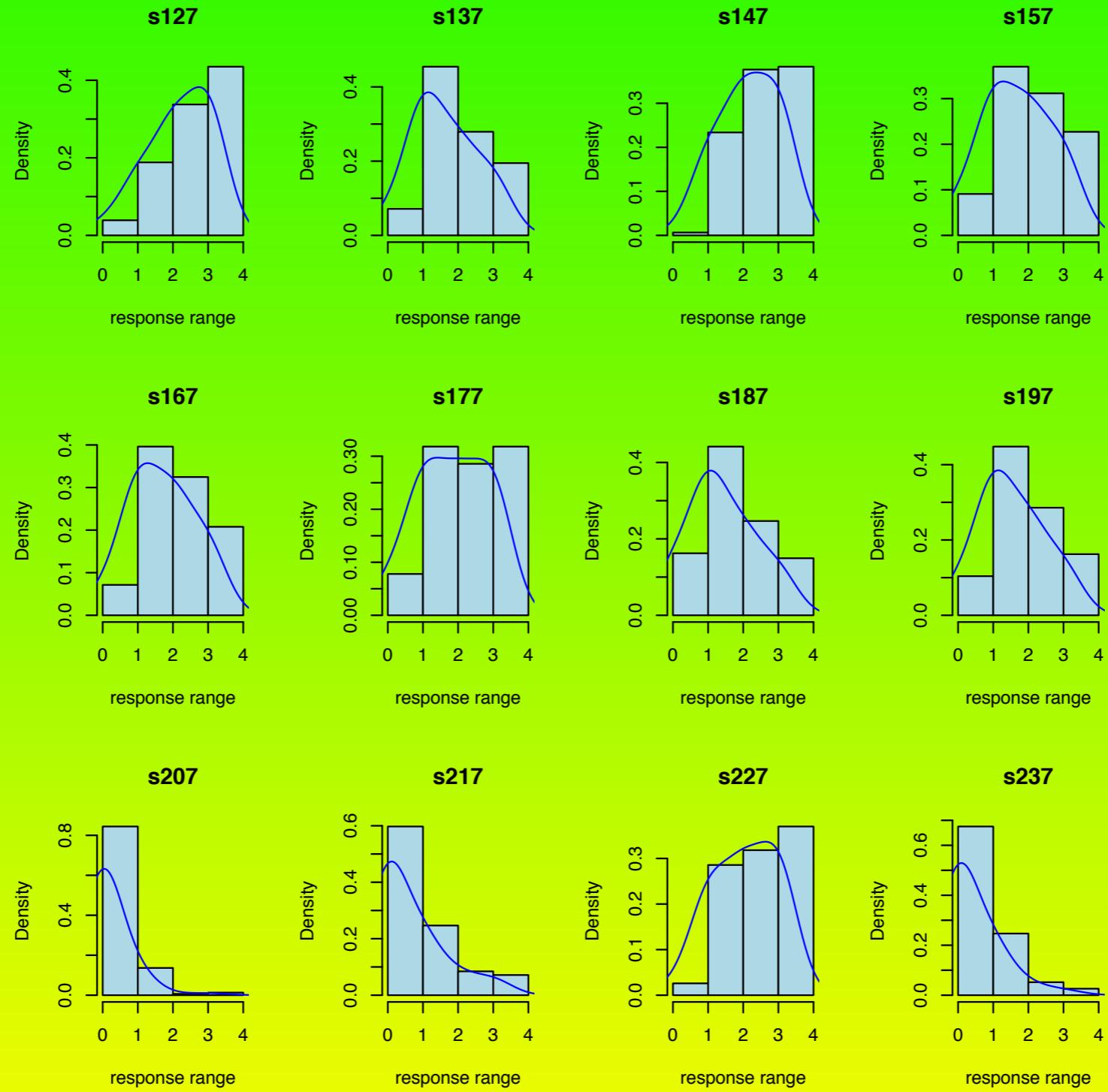
gr6 の s006, s016, … , s281.6, s282.6 の分析の前処理



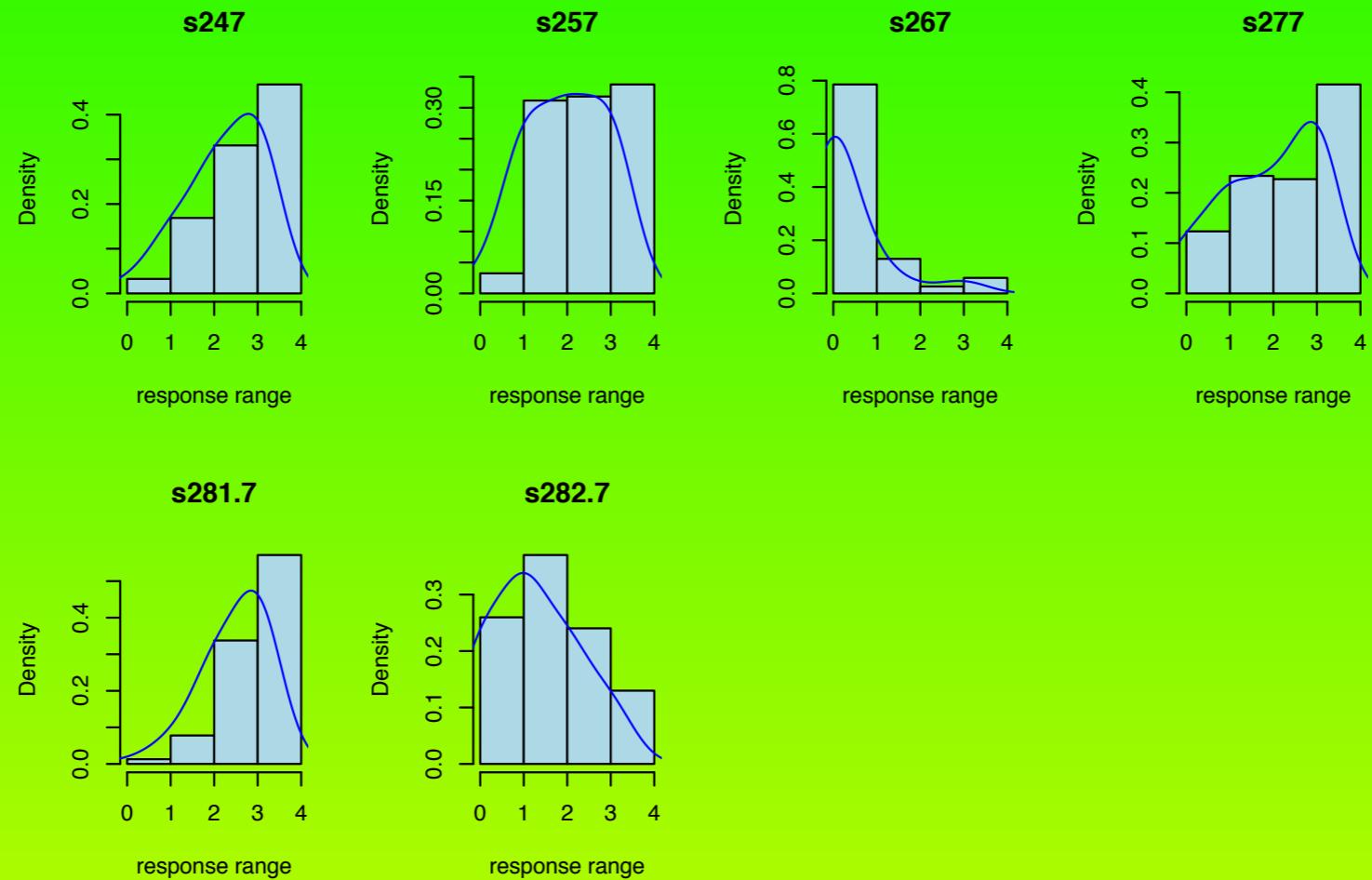
gr7 の s007, s017, ..., s117 [1/3]



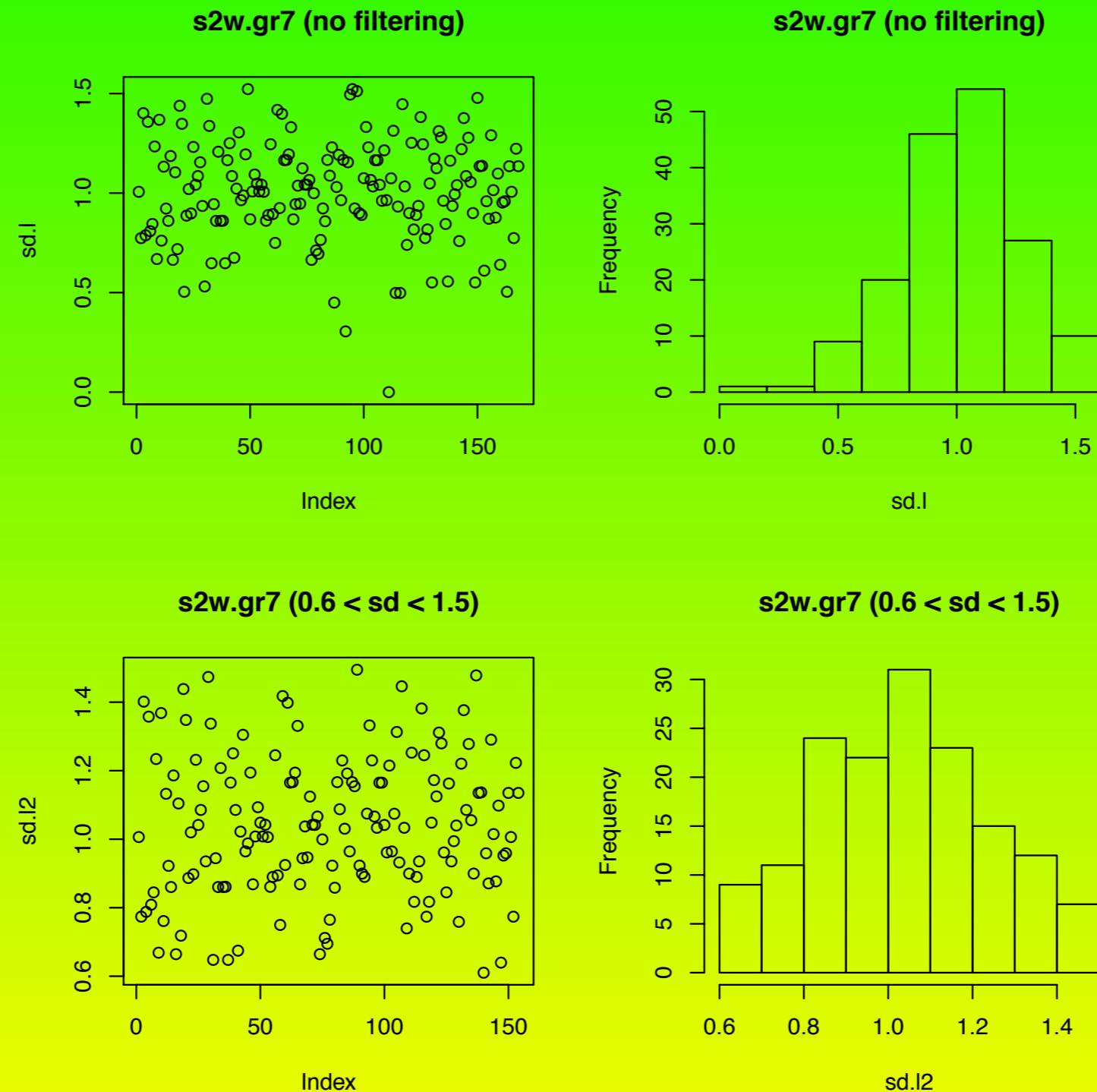
gr7 の s127, s137, ..., s237 [2/3]



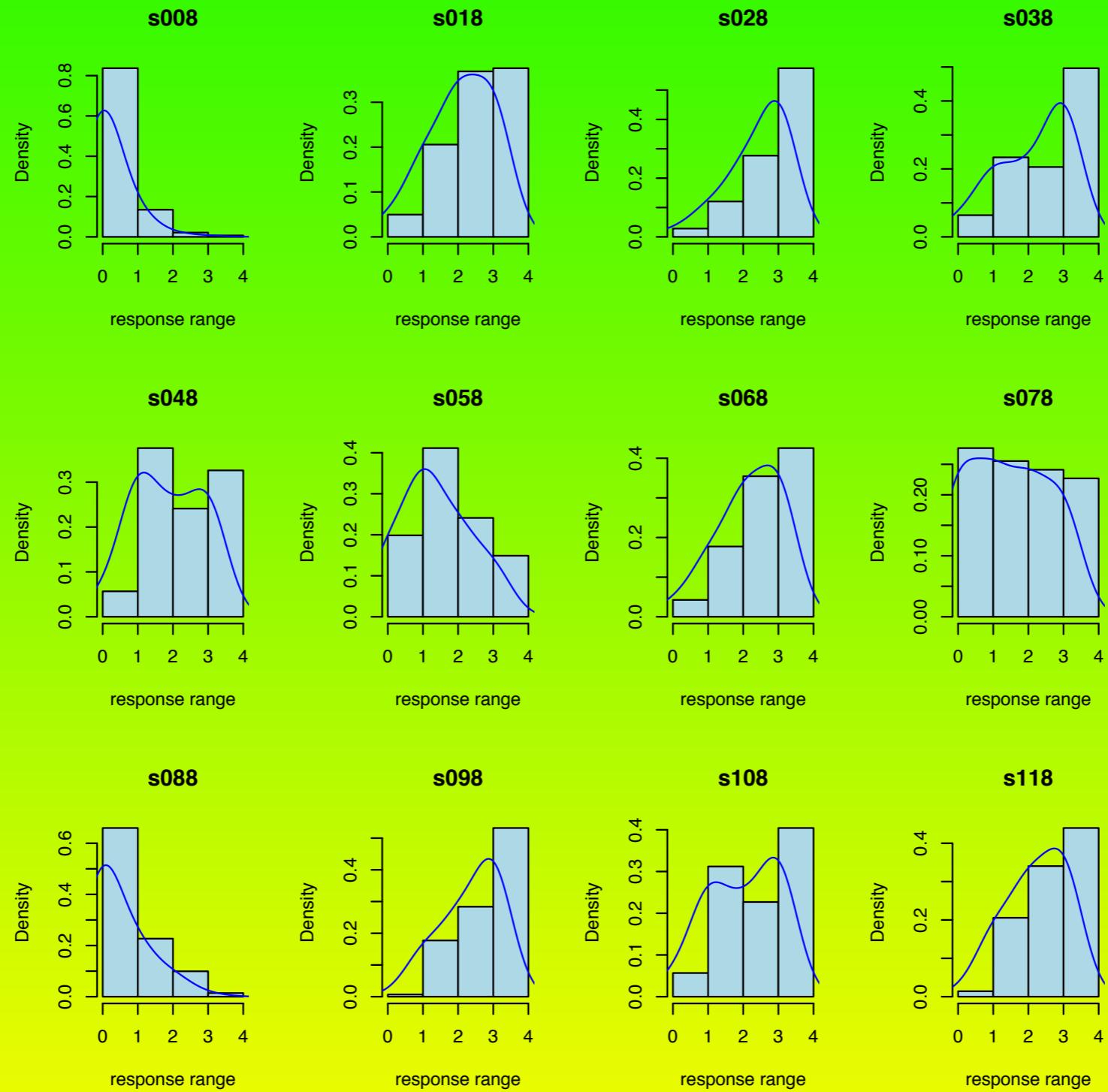
gr7 の s247, s257, ..., s282.7 [3/3]



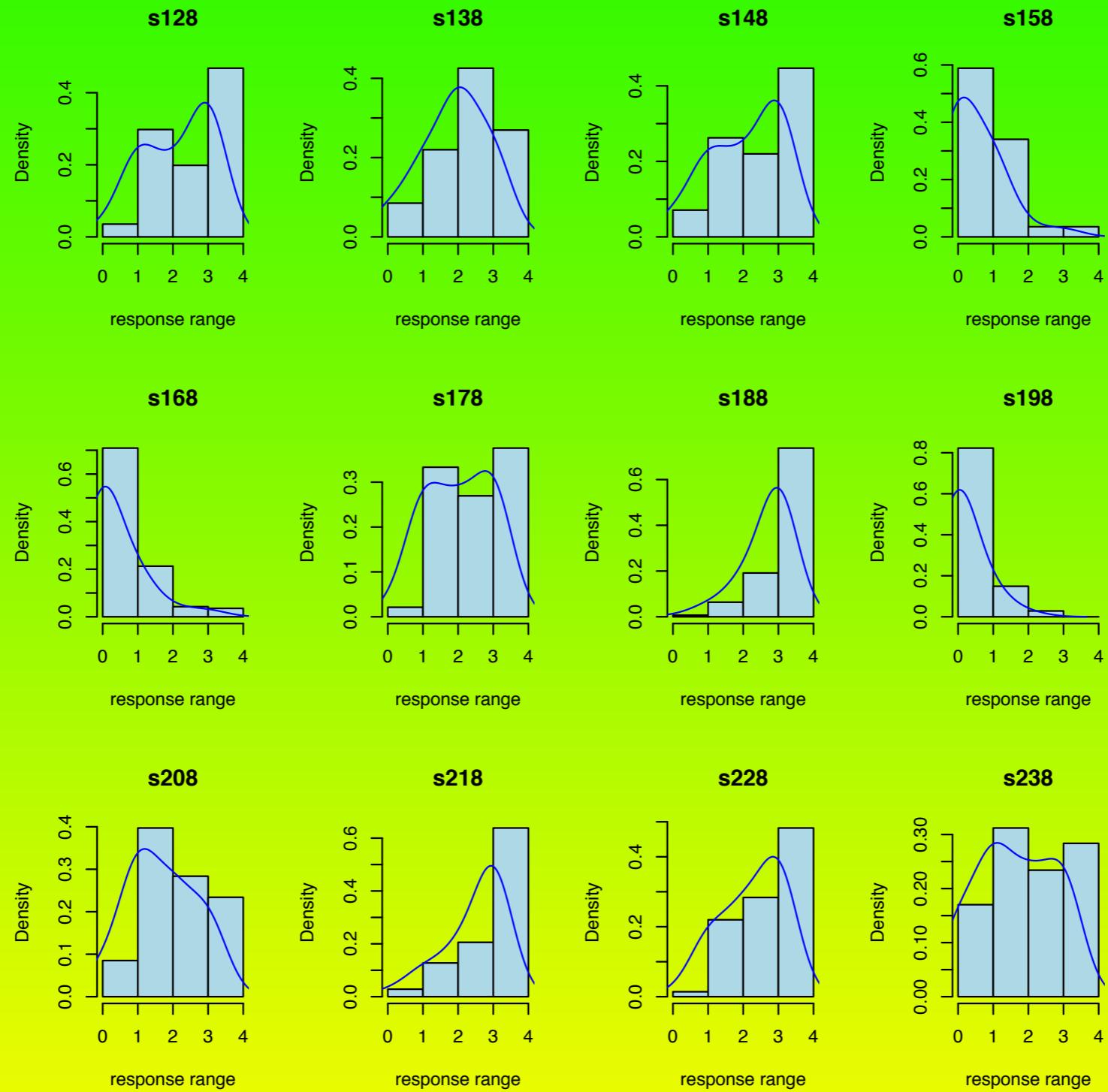
gr7 の s007, s017, … , s281.7, s282.7 の分析の前処理



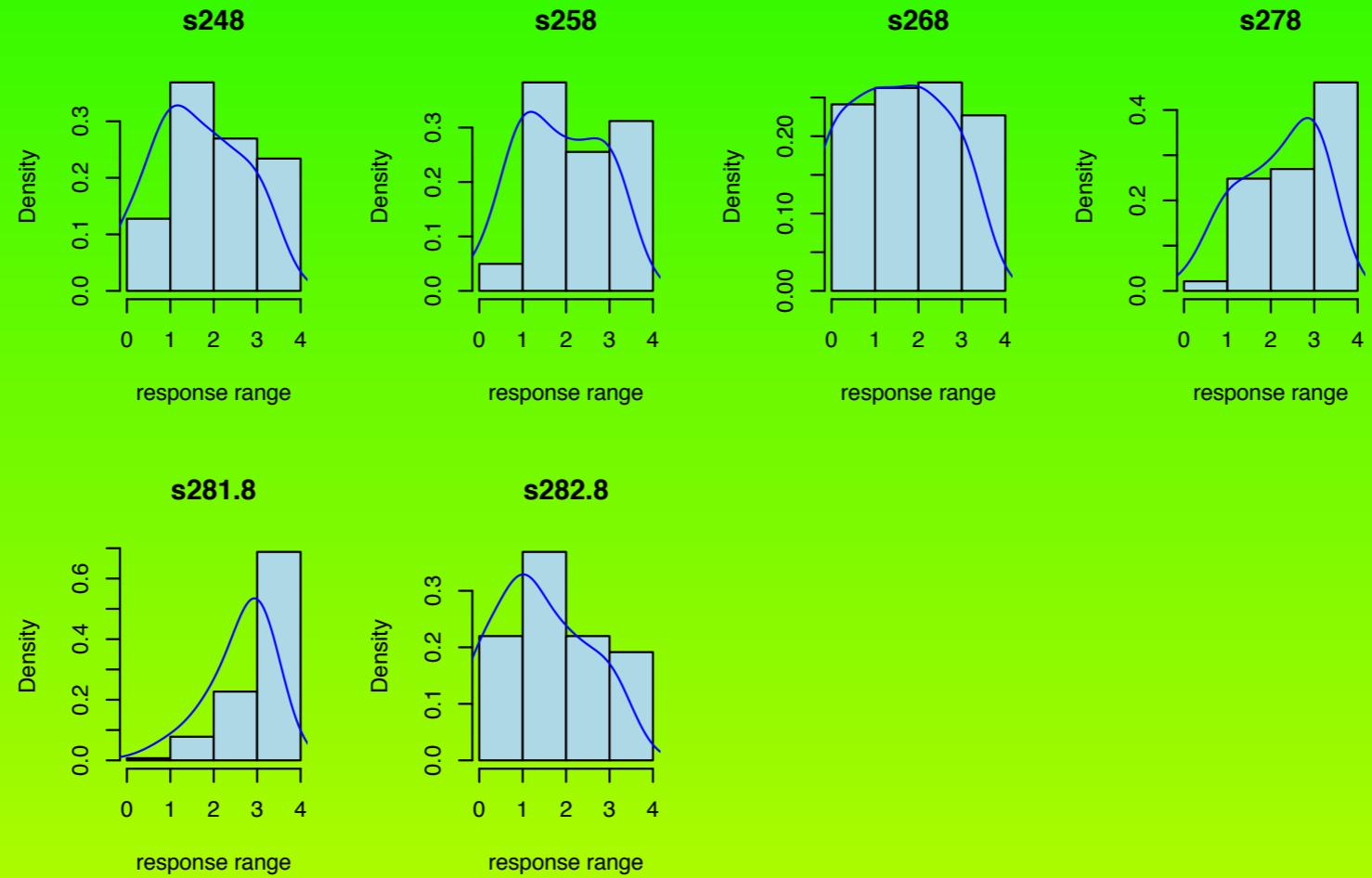
gr8 の s008, s018, ..., s118 [1/3]



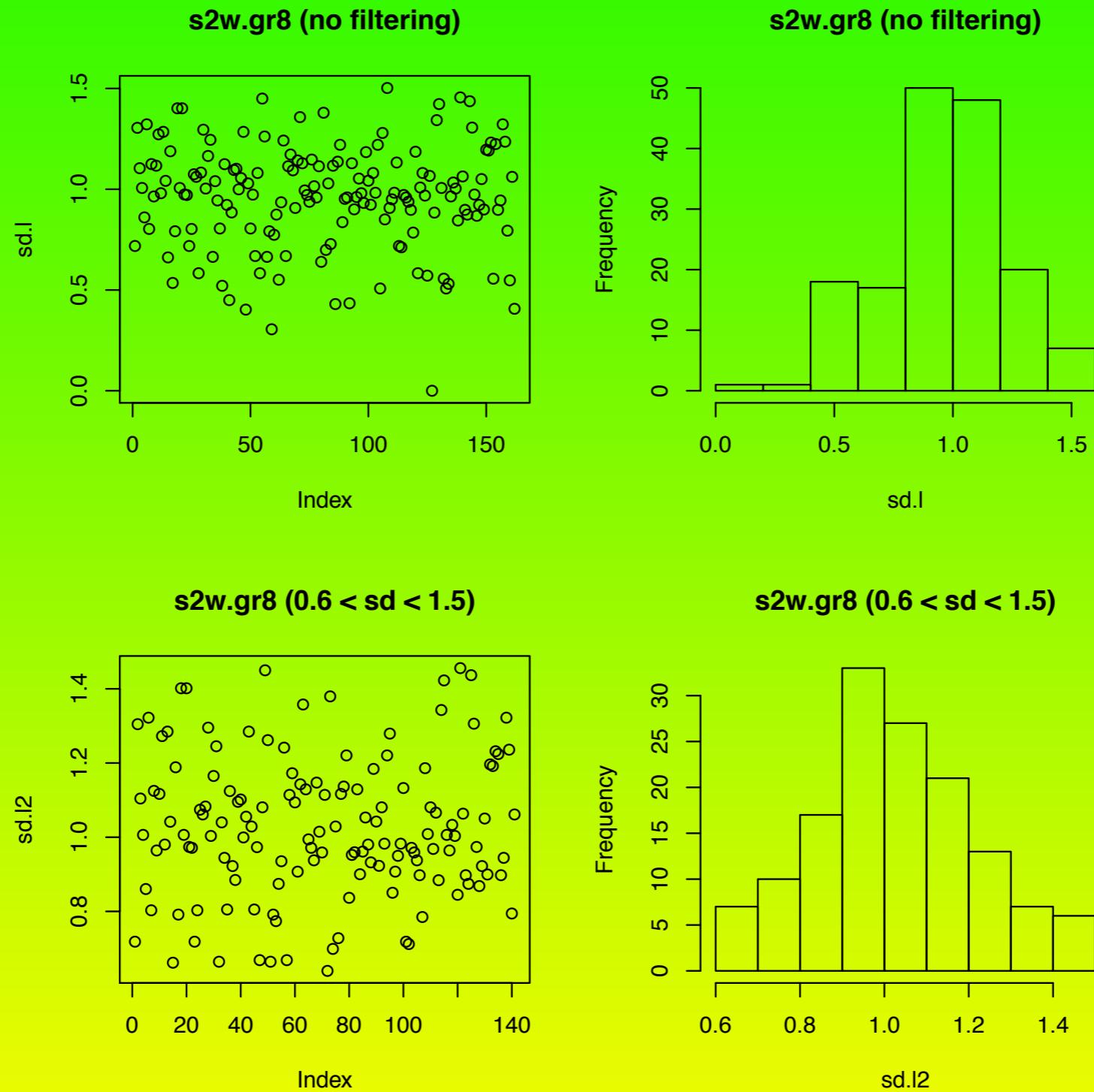
gr8 の s128, s138, ..., s238 [2/3]



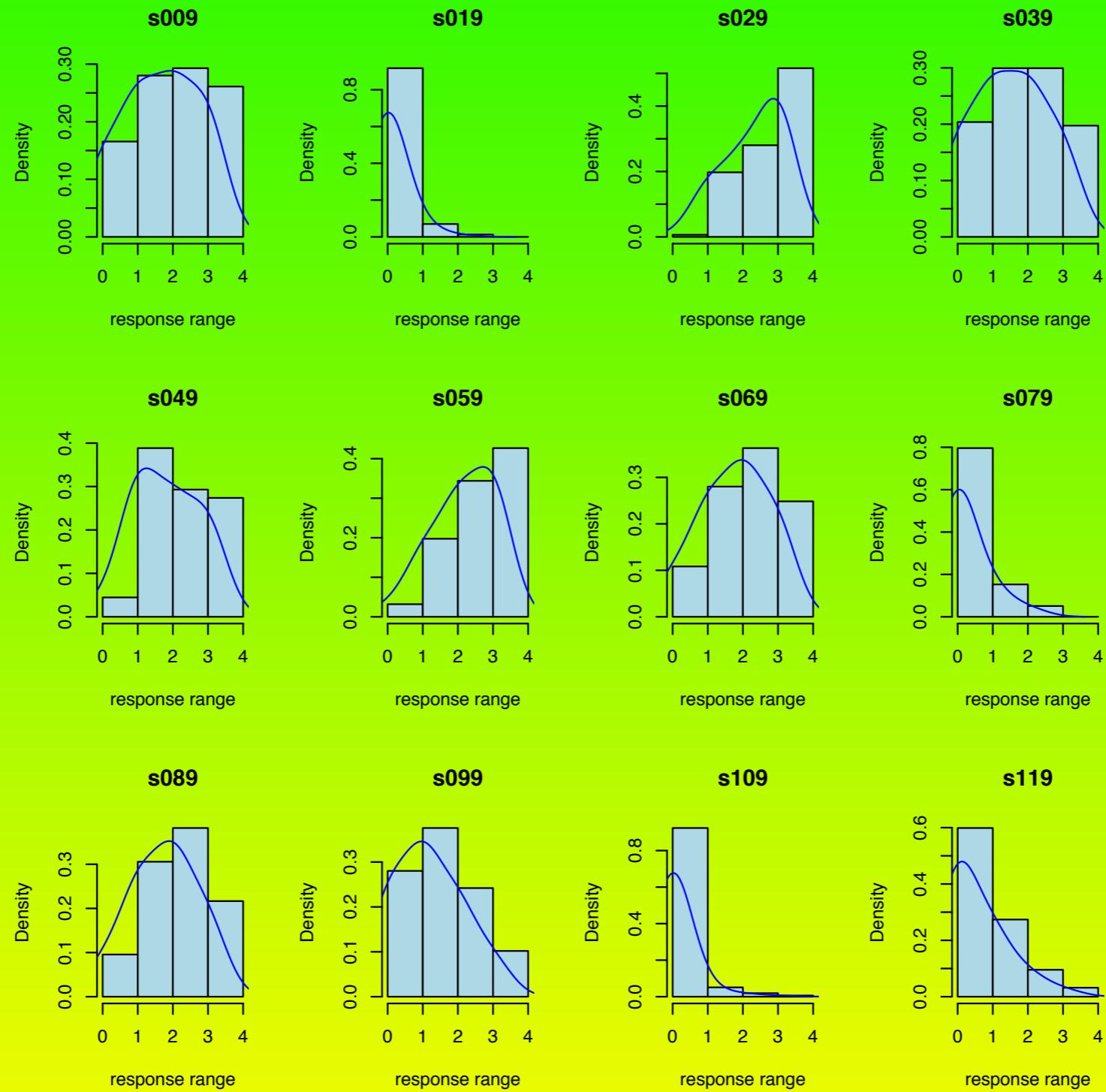
gr8 の s248, s258, ..., s282.8 [3/3]



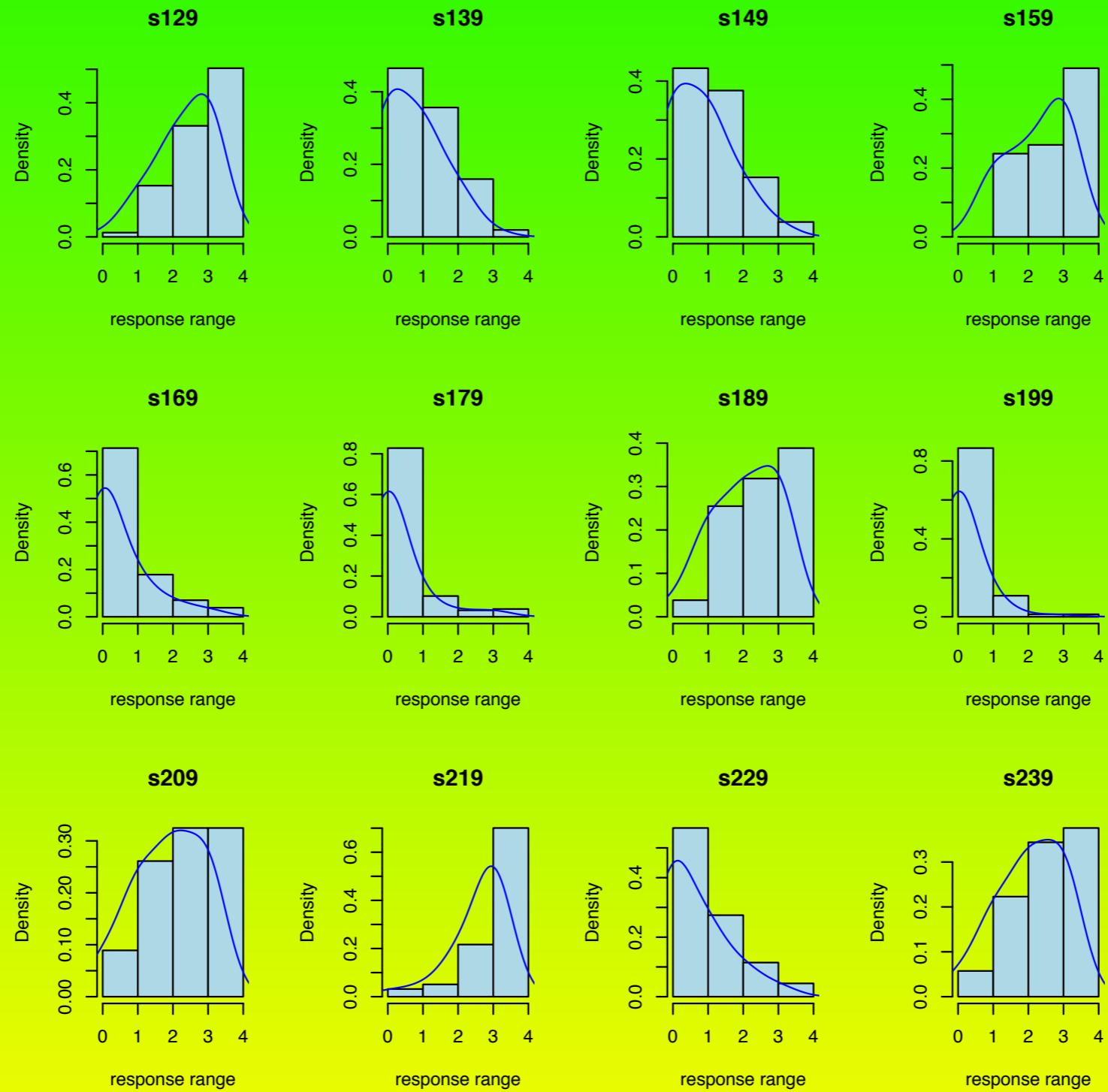
gr8 の s008, s018, … , s281.8, s282.8 の分析の前処理



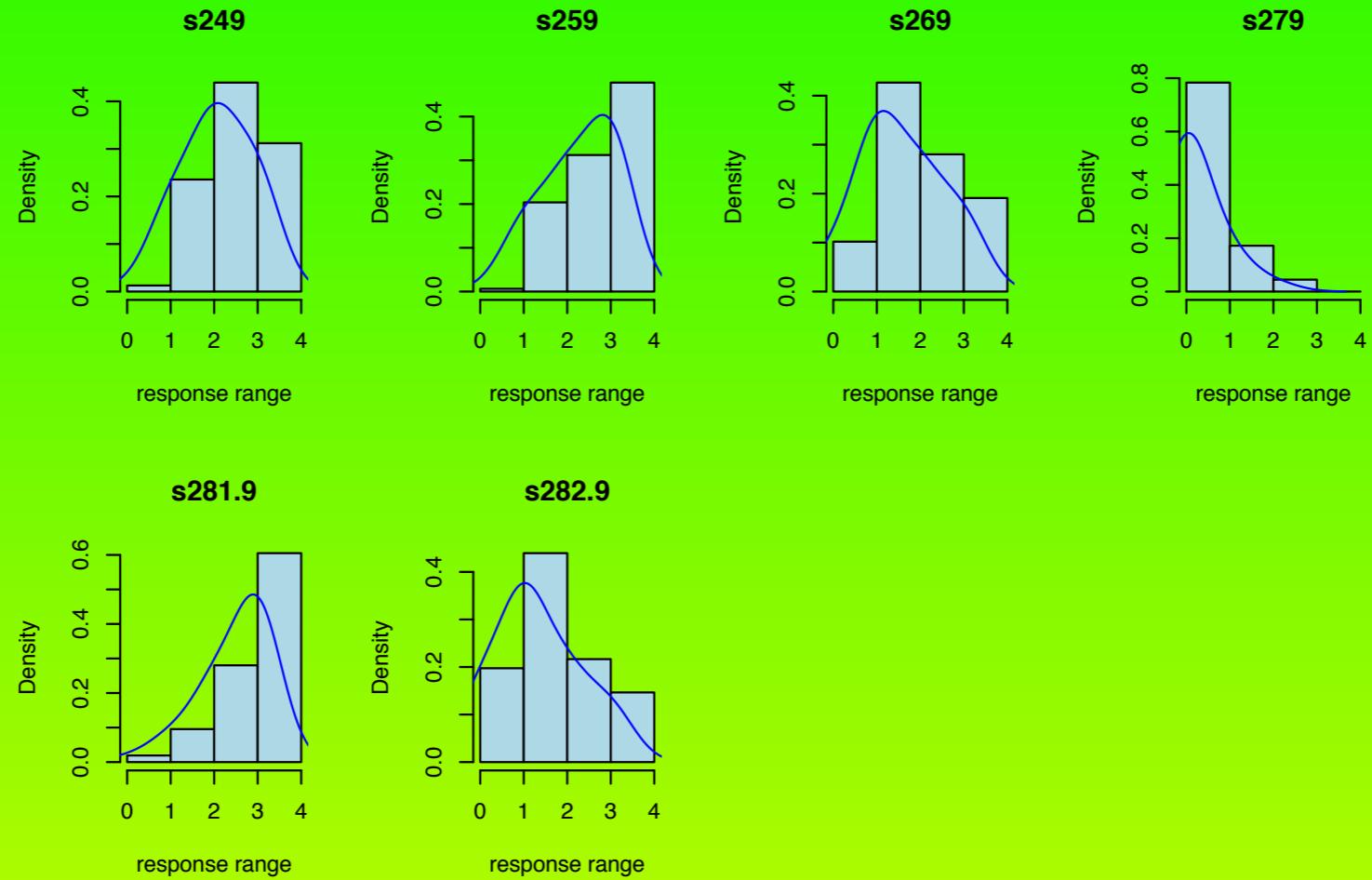
gr9 の s009, s019, ..., s119 [1/3]



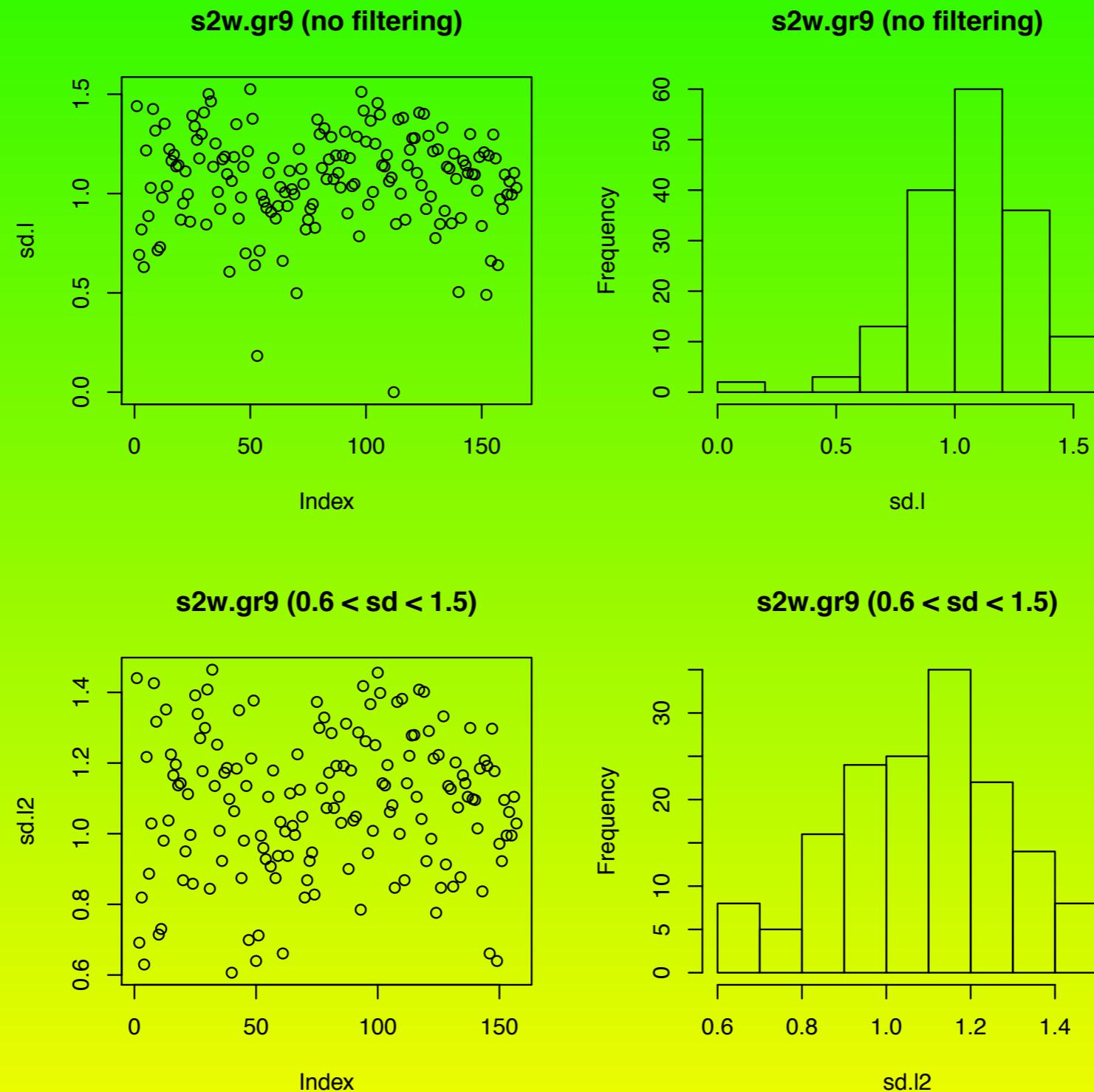
gr9 の s129, s139, ..., s239 [2/3]



gr9 の s249, s259, ..., s282.9 [3/3]



gr9 の s009, s019, … , s281.9, s282.9 の分析の前処理



注意

- ・グラフが表わしているのは、個々の文の容認度評定=反応ポテンシャルの形状
- ・これは解釈する側の主観性を表わしたものではない
- ・これは解釈する側の性質でなく、解釈される対象=文の性質
- ・1文につき約160名の評定者による反応の平均なので、ランダムなものとは考えにくい

わかった事

- 知見 1
 - 個々の文に対する反応パターン (= 容認度評定) は言語学が理想化しているように一様ではない
 - つまり, 容認できる/できないの区別 (is-a 容認度が高い/低いの区別) は一様でない
- 知見 2
 - ただ $r[0,1]$, $r[1,2]$, $r[2,3]$, $r[3,\infty)$ の区域での反応分布から見て幾つかの類型があるのは明らか
 - つまり, 反応ポテンシャルのタイプロジーがある
 - しかも単峰性とは限らない (二峰性や一様性の分布もある)
 - 理論には非单峰性が厄介

次の課題

- ・文の反応のタイプロジー/類型化が理論的には可能
- ・ただし定量的解析が必要
- ・定性的解析（直観に基づく分類）だけでは無理

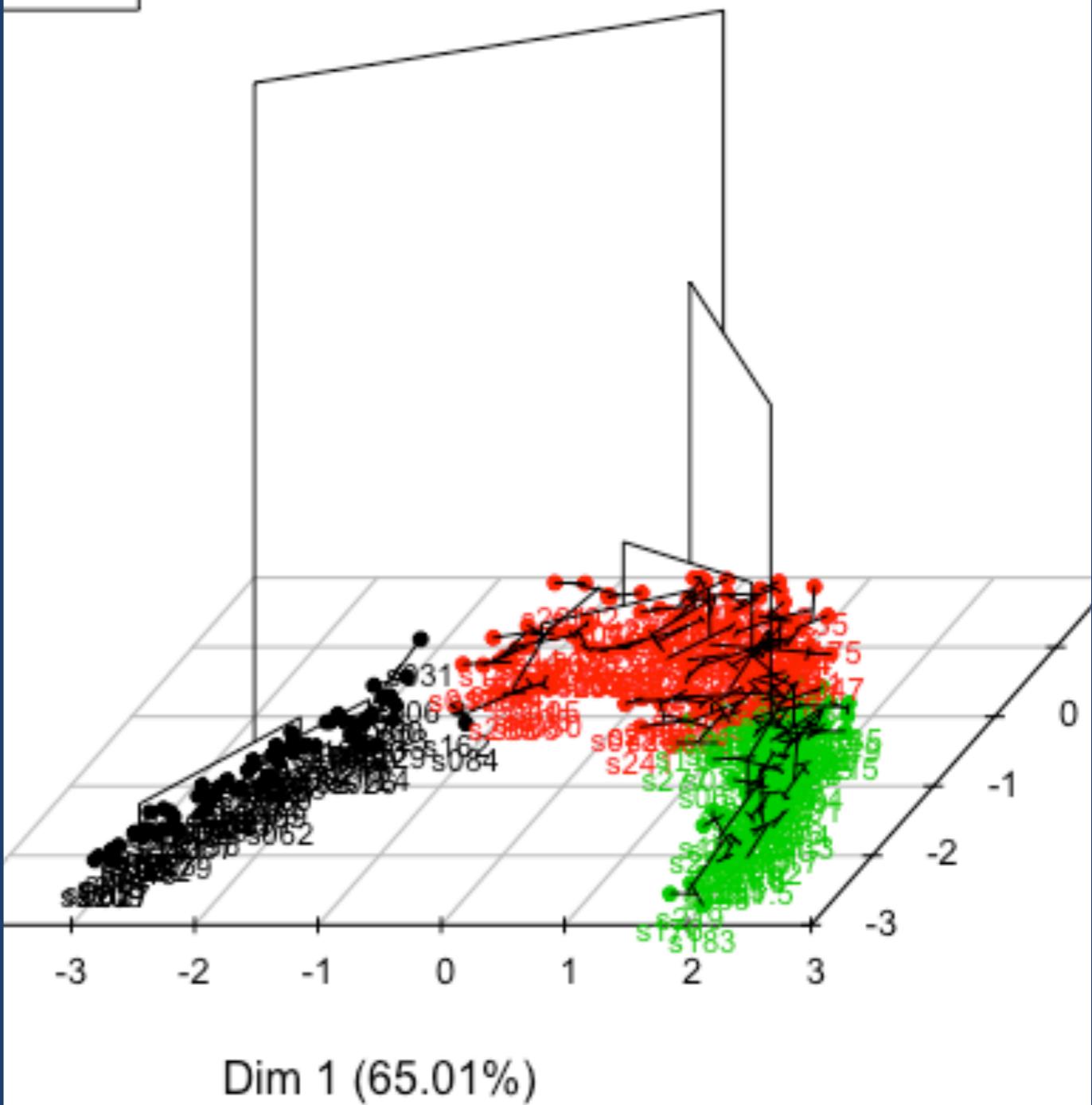
多変量解析を応用して解決

- まず反応のクラスタリングをしてグループ分けしてみる
 - Hierarchical Clustering (HC) を使って
 - データの標準化あり vs なし
- その後に次元を削減し、地図を作ってみる
 - Principal Component Analysis (PCA), Semi-supervised Local Fisher Discriminat Analysis (SELFA), PCA3D を使って

反応ポテンシャルの分類

Hierarchical clustering on the factor map

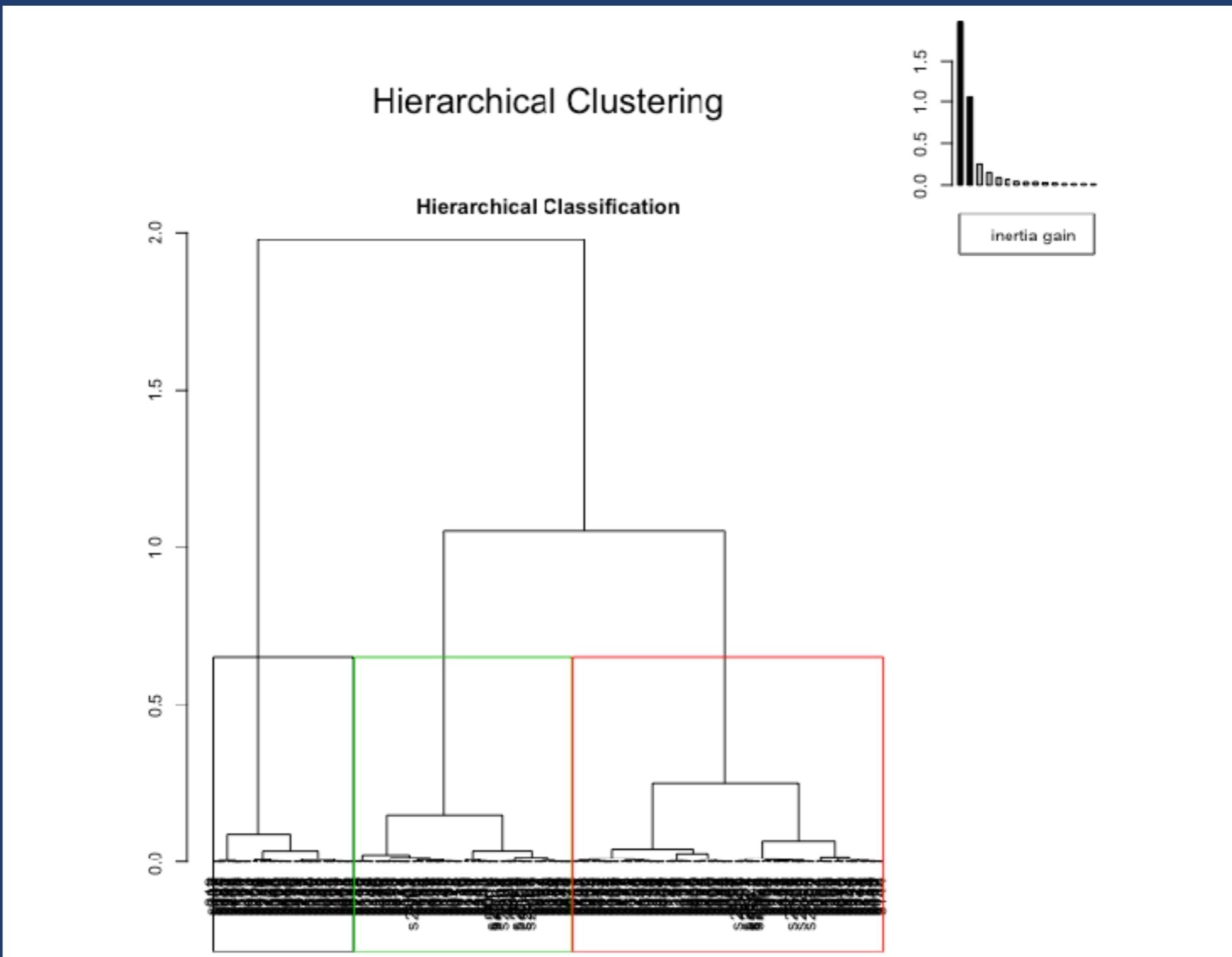
Cluster 1
Cluster 2
Cluster 3



手順

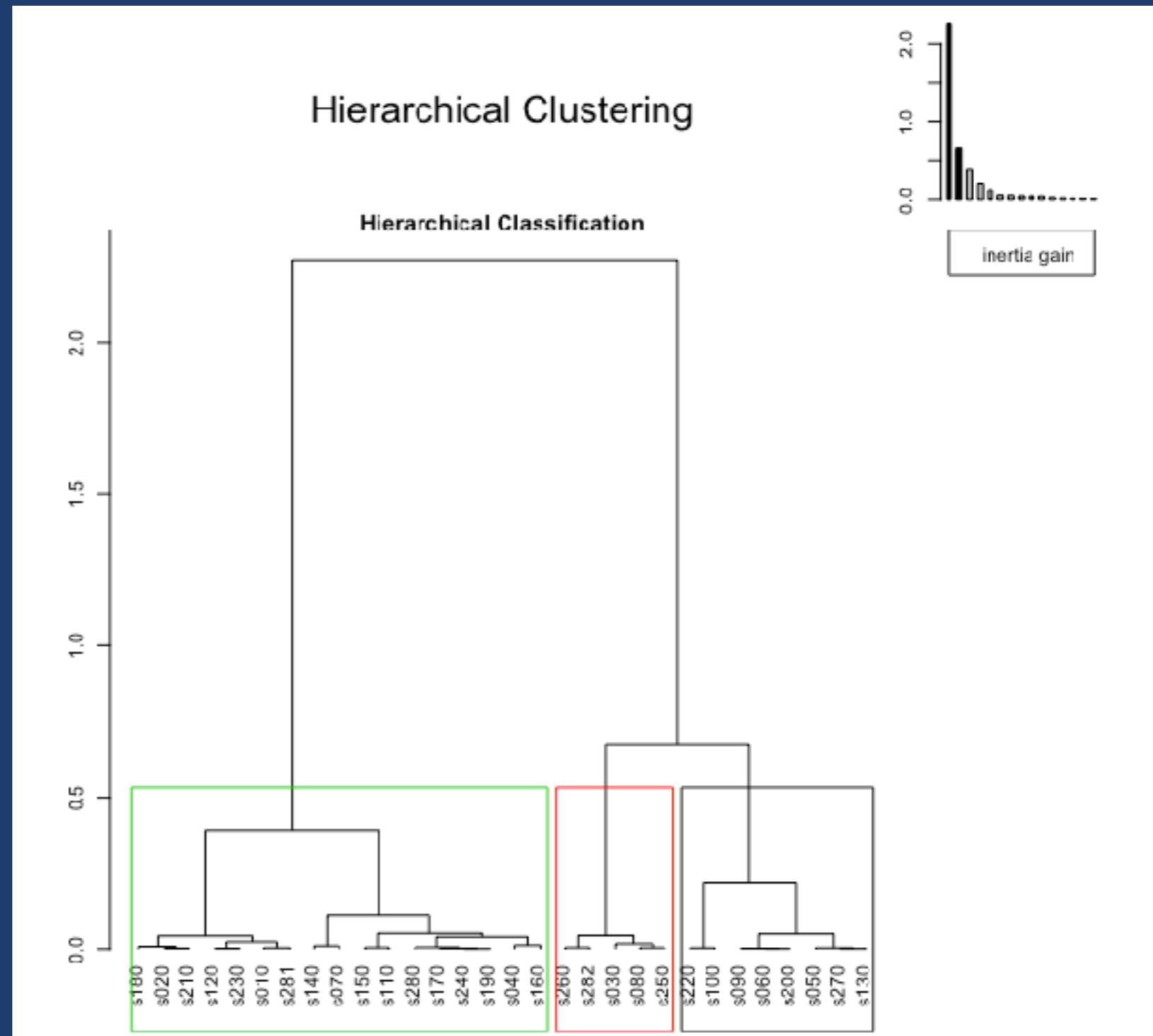
- 反応ポテンシャルの定義
- 4つの反応区域
 - `range[0,1], range[1,2], range[2,3], range[3,∞)` のそれぞれの選択確率の分布
 - $\pi = (\text{prob}[0,1], \text{prob}[1,2], \text{prob}[2,3], \text{prob}[3,∞))$
 - を求める
 - π は確率分布ベクトルで標準化
 - 標準化されていないと、異なる `gr` に属する事例を比較できない
- 使う多変量解析
 - Hierarchical Clustering (HC)
 - `x-means` 法: `k-means` の k の最適値を自動的に見つける
 - Principal Component Analysis (PCA)
 - HC と PCA のかけ合わせ
 - を実行
 - R の `FactoMineR` (Le, et al. 2008) パッケージを利用

反応ポテンシャルの Hierarchical Clustering



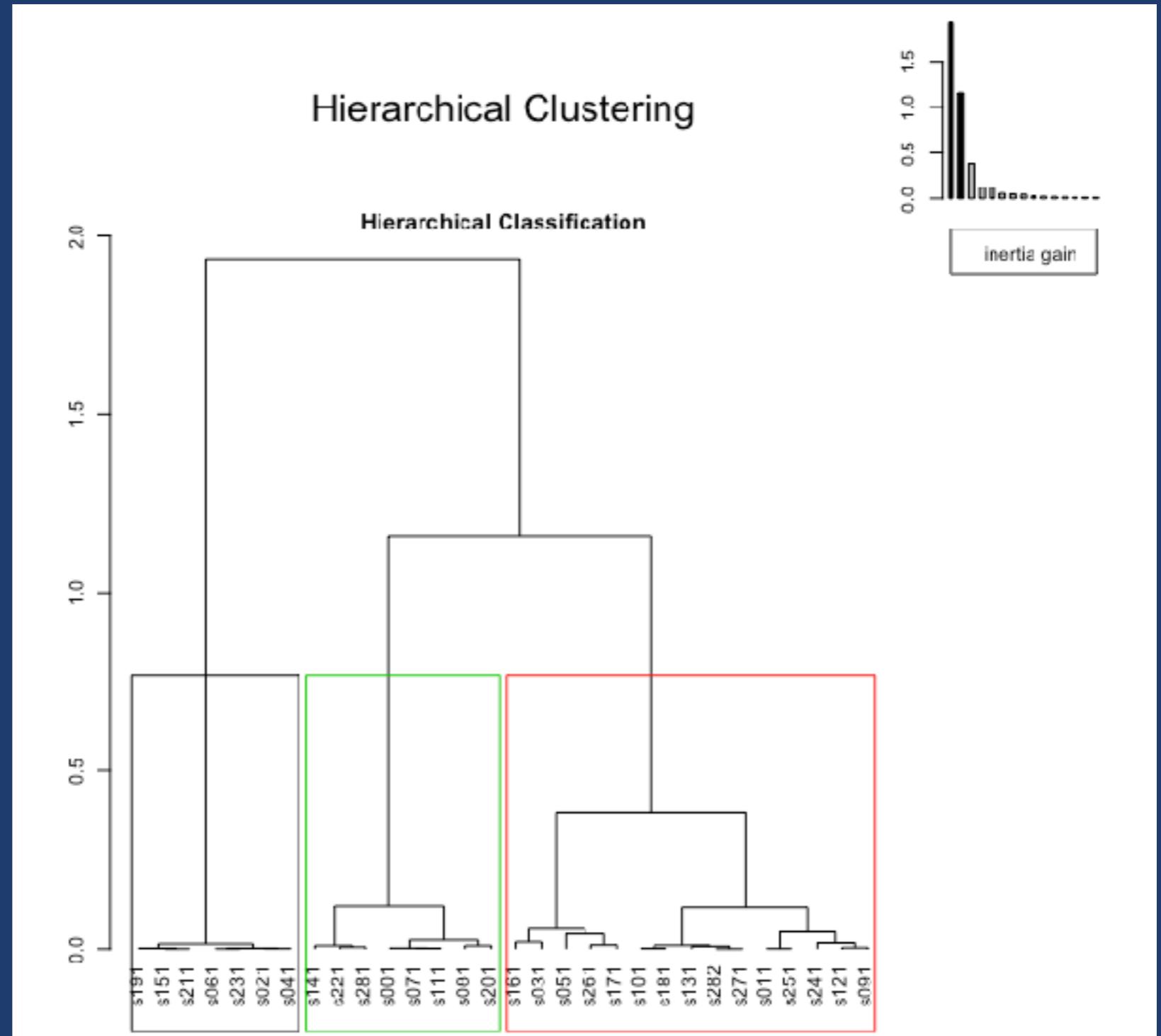
gr0 の刺激文の階層クラスタリング

- 3つのクラスターが形成されている
 - 黒 acceptables
 - 赤 questionables
 - 緑 unacceptables



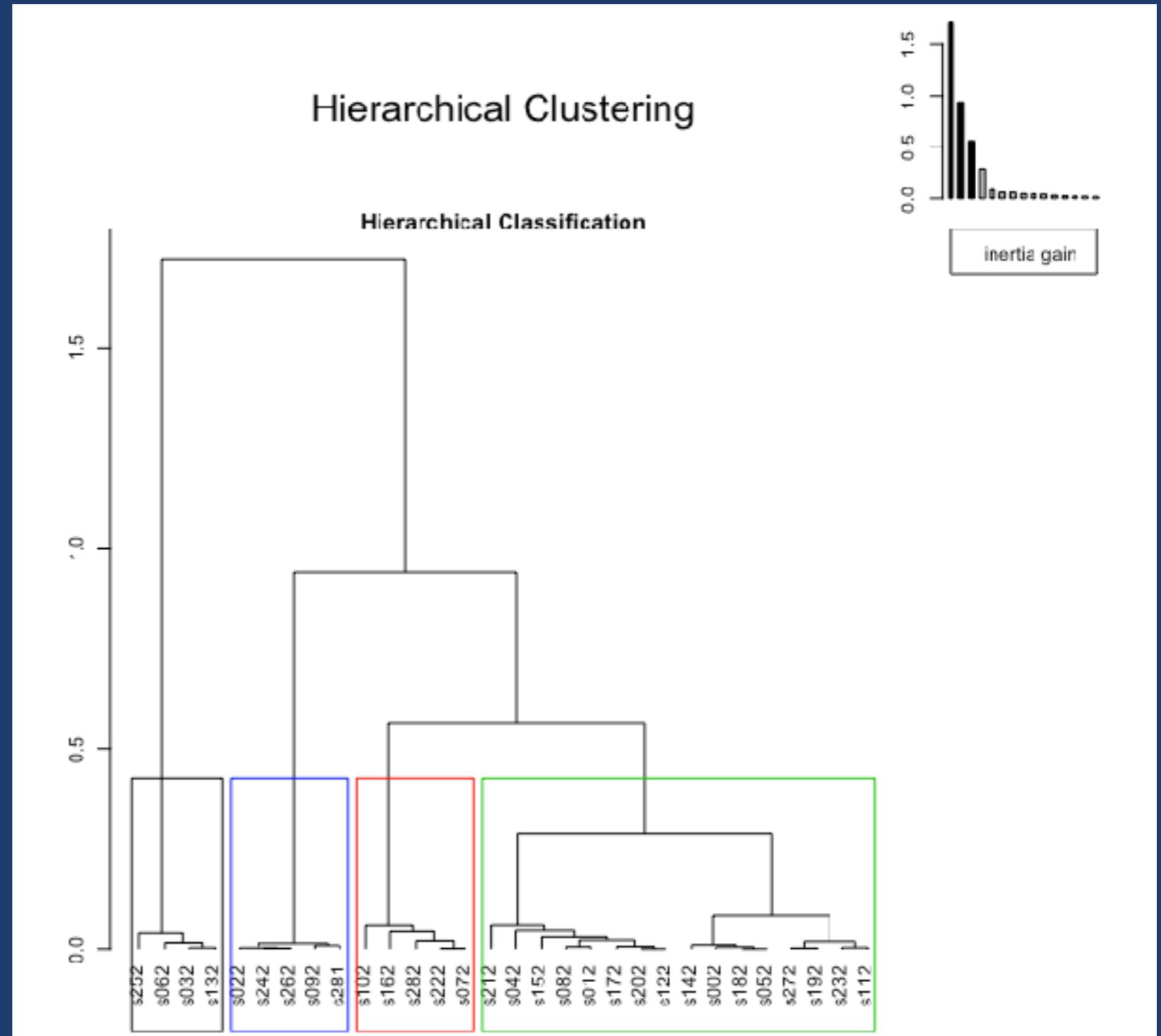
gr1 の刺激文の階層クラスタリング

- 3つのクラスターが形成されている
 - 黒 acceptables
 - 赤 questionables
 - 緑 unacceptables



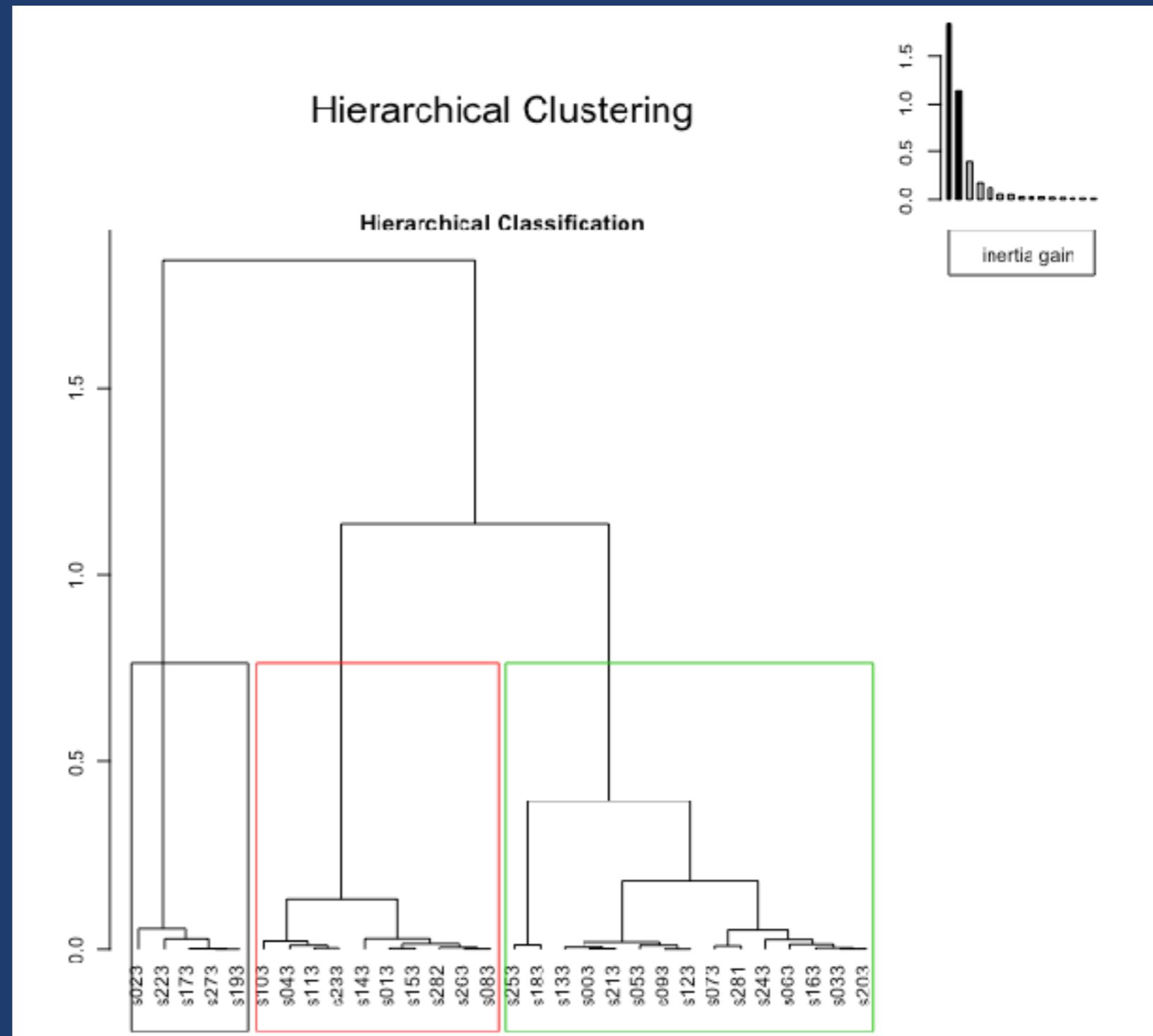
gr2 の刺激文の階層クラスタリング

- 4つのクラスターが形成されている
 - 黒 acceptables
 - 赤 questionables 1
 - 緑 questionables 2
 - 青 unacceptables



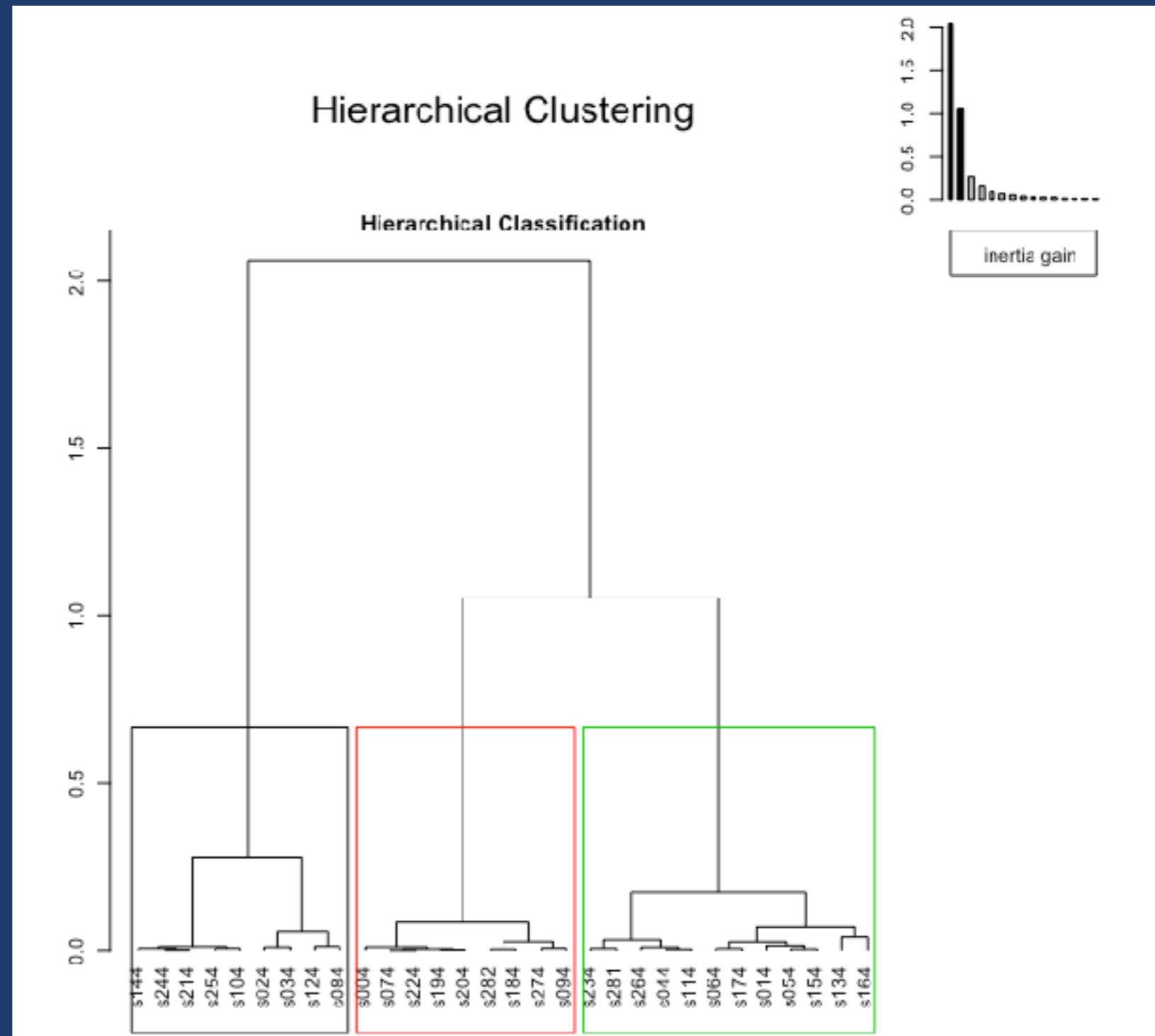
gr3 の刺激文の階層クラスタリング

- 3つのクラスターが形成されている
 - 黒 acceptables
 - 赤 questionables
 - 緑 unacceptables



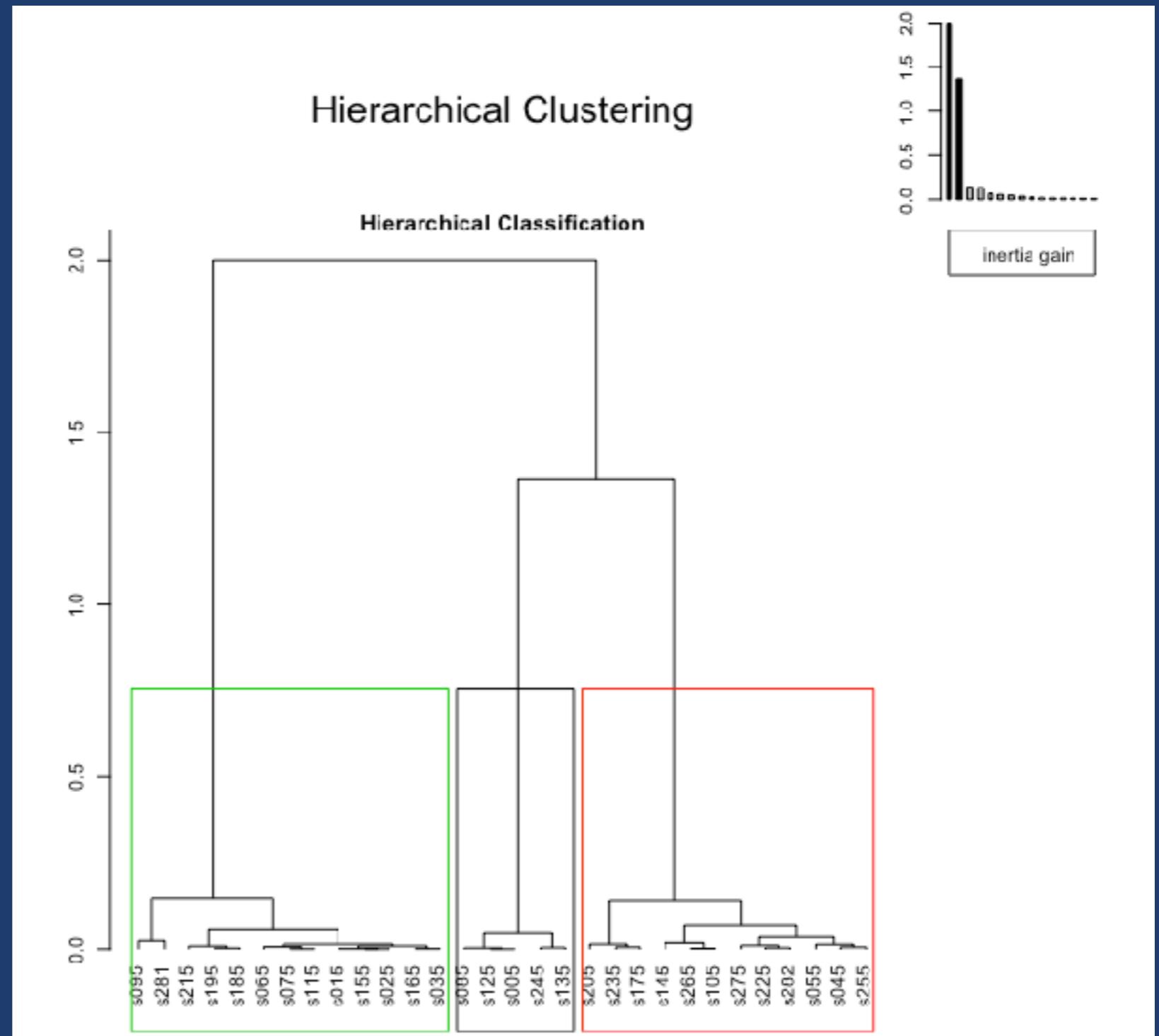
gr4 の刺激文の階層クラスタリング

- 3つのクラスターが形成されている
 - 黒 acceptables
 - 赤 questionables
 - 緑 unacceptables



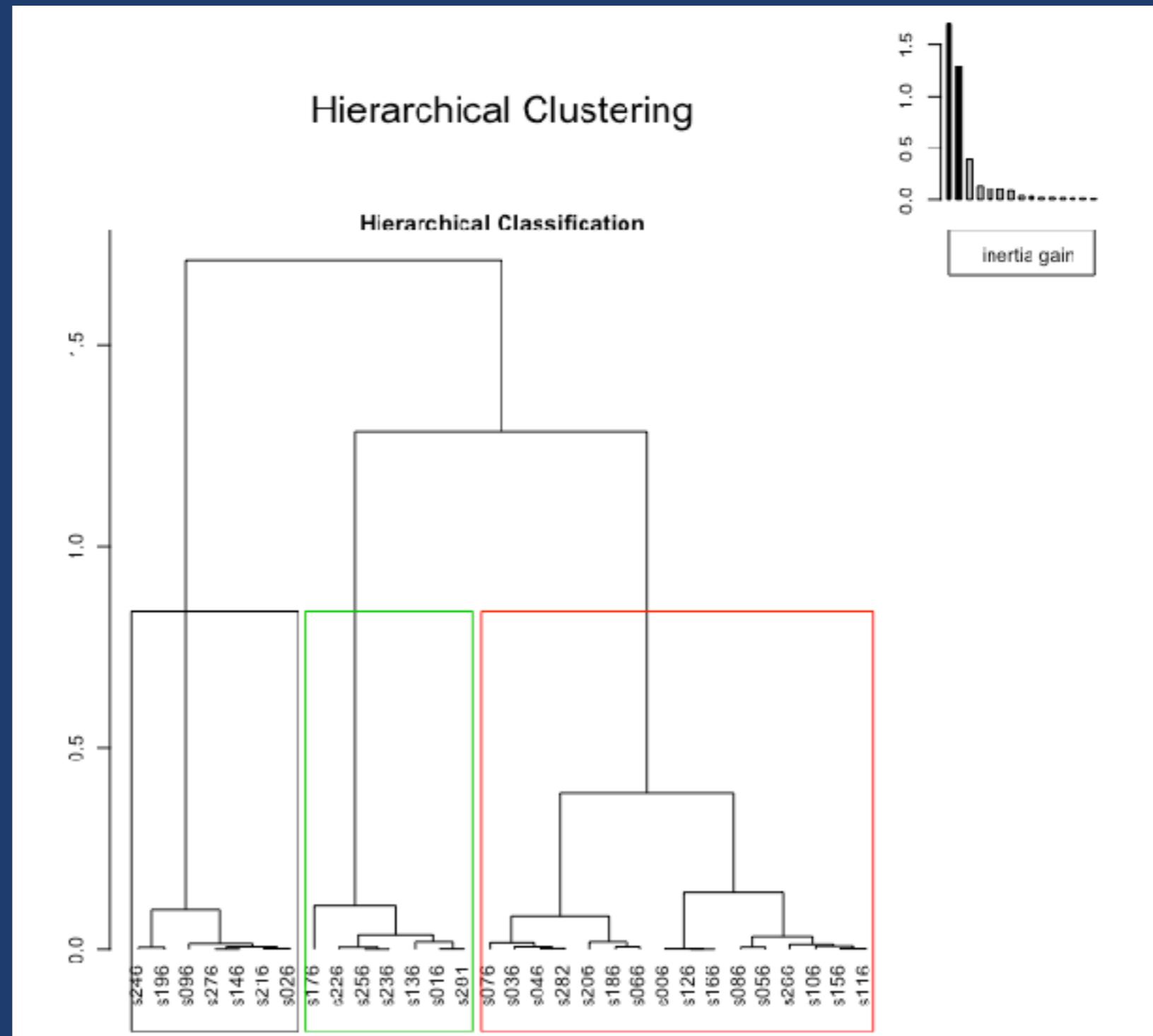
gr5 の刺激文の階層クラスタリング

- 3つのクラスターが形成されている
 - 黒 acceptables
 - 赤 questionable
 - 緑 unacceptable



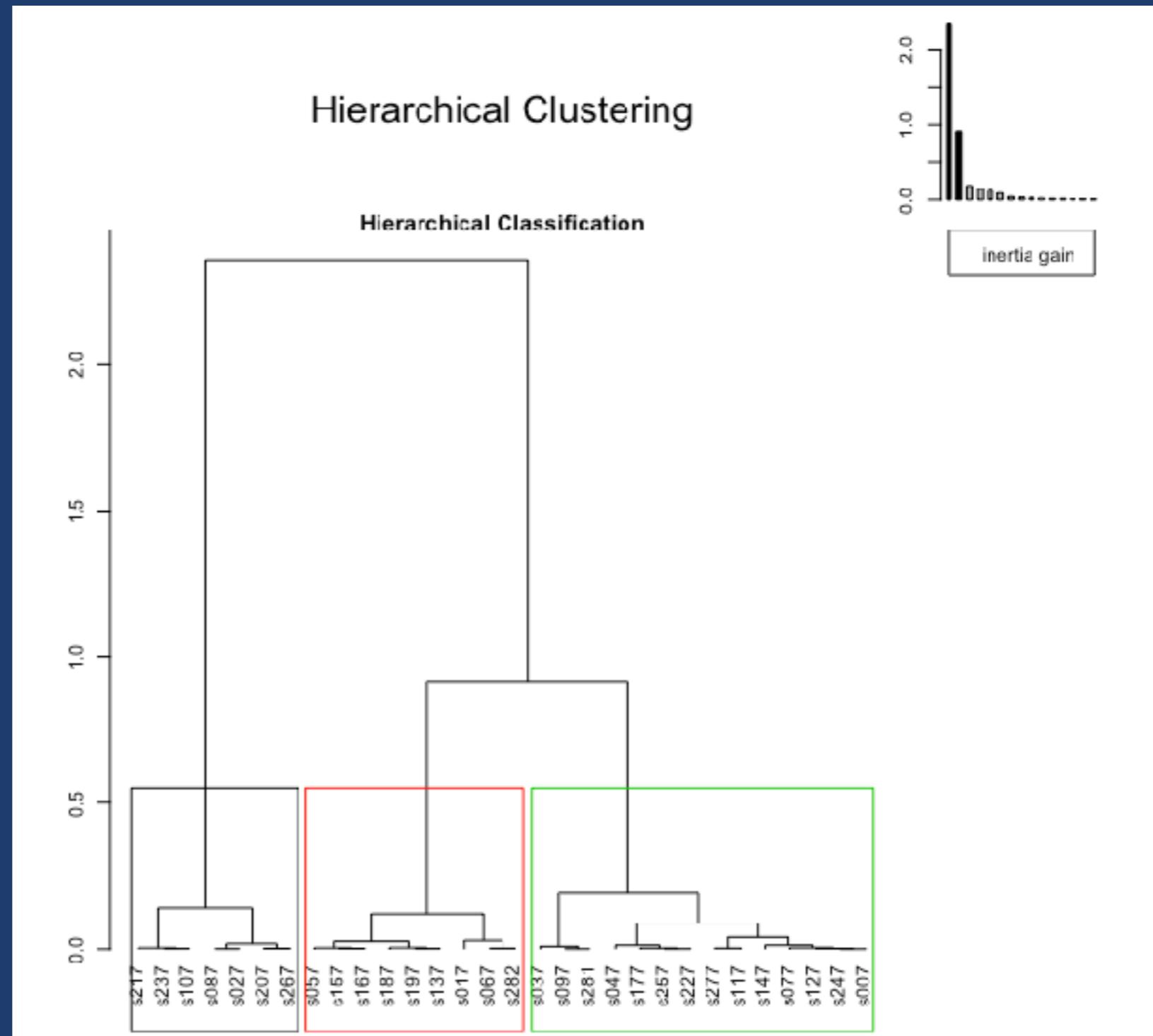
gr6 の刺激文の階層クラスタリング

- 3つのクラスターが形成されている
 - 黒 acceptables
 - 赤 questionables
 - 緑 unacceptables



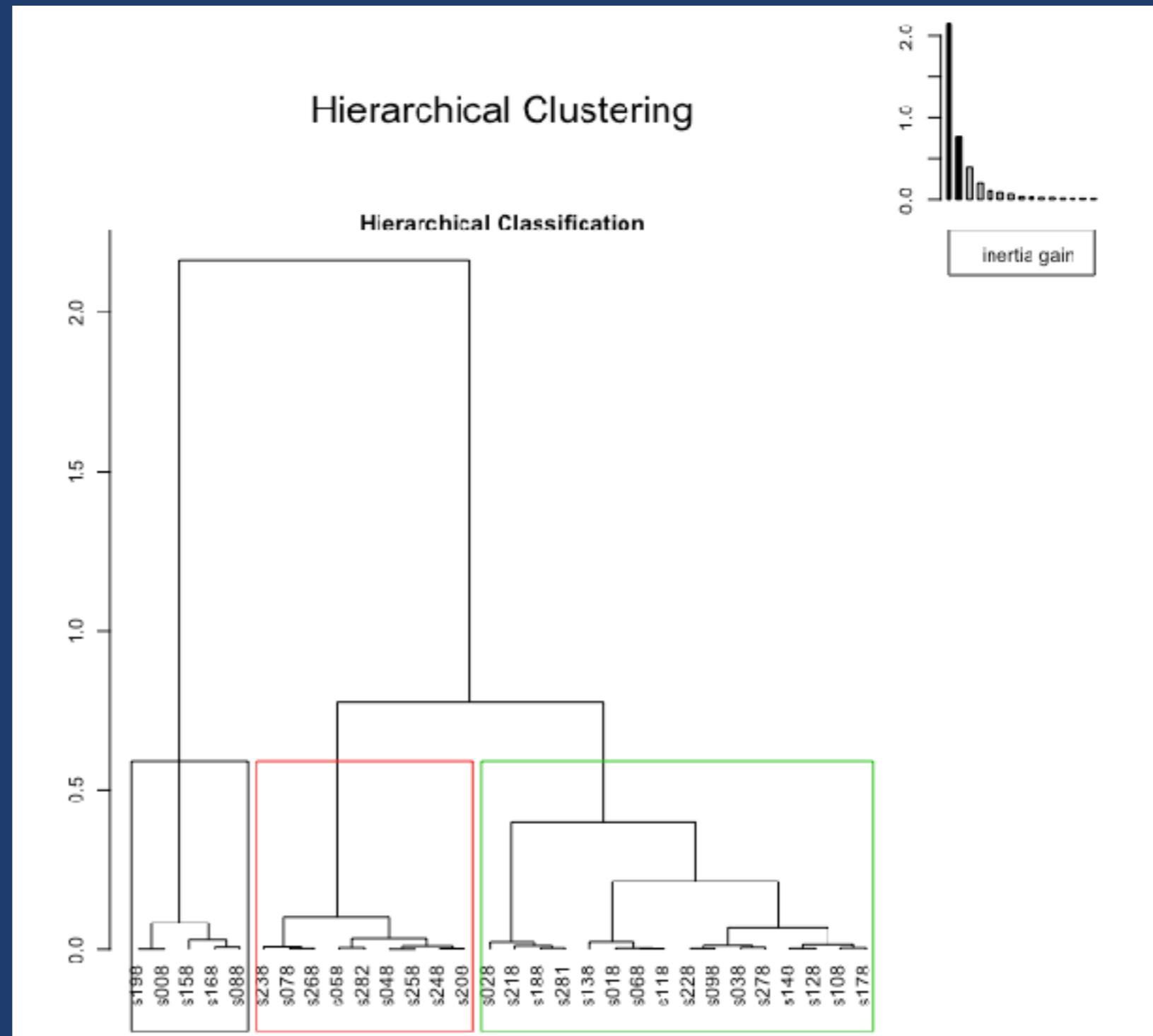
gr7 の刺激文の階層クラスタリング

- 3つのクラスターが形成されている
 - 黒 acceptables
 - 赤 questionables
 - 緑 unacceptables



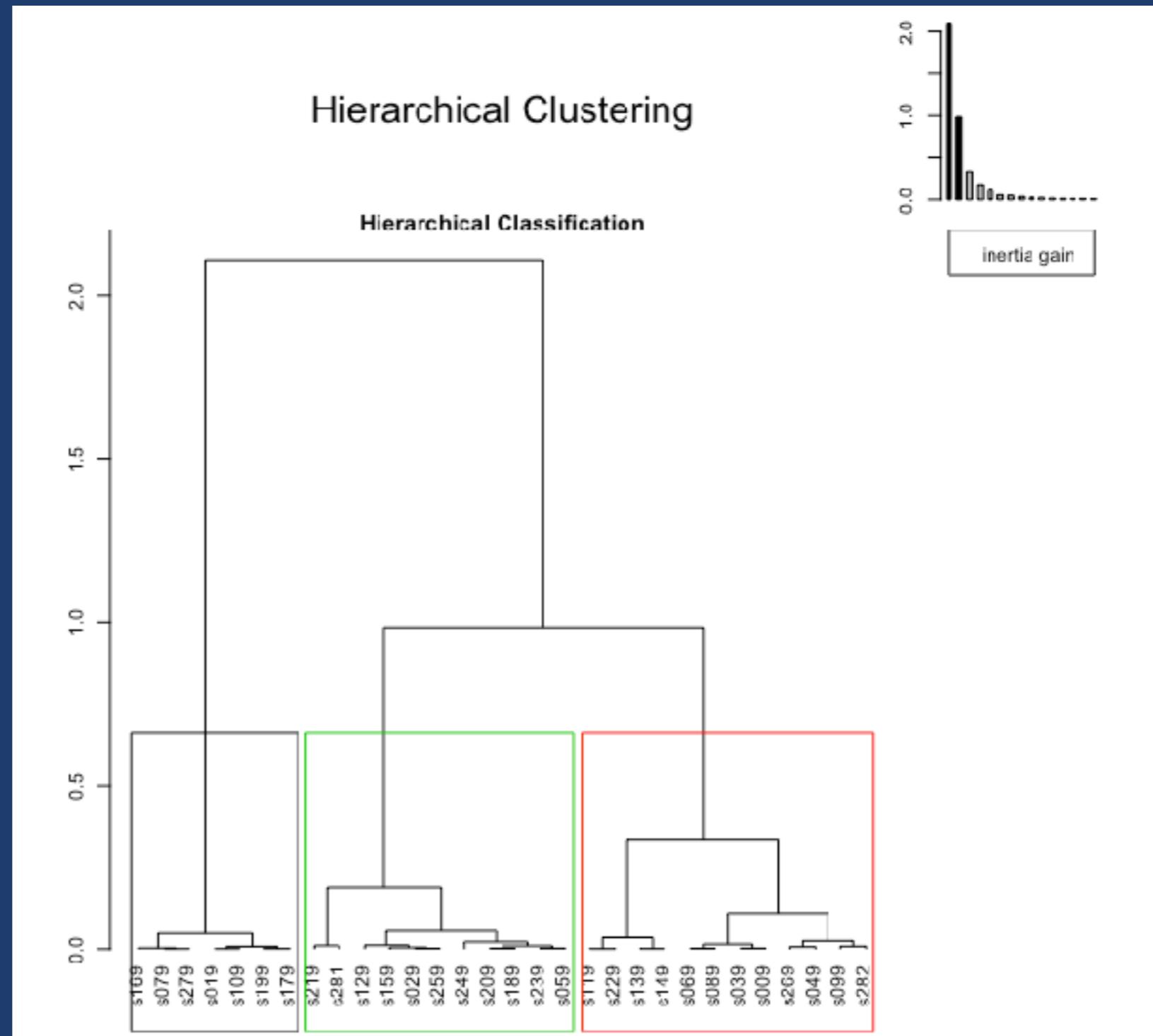
gr8 の刺激文の階層クラスタリング

- 3つのクラスターが形成されている
 - 黒 **acceptables**
 - 赤 **questionables**
 - 緑 **unacceptables**



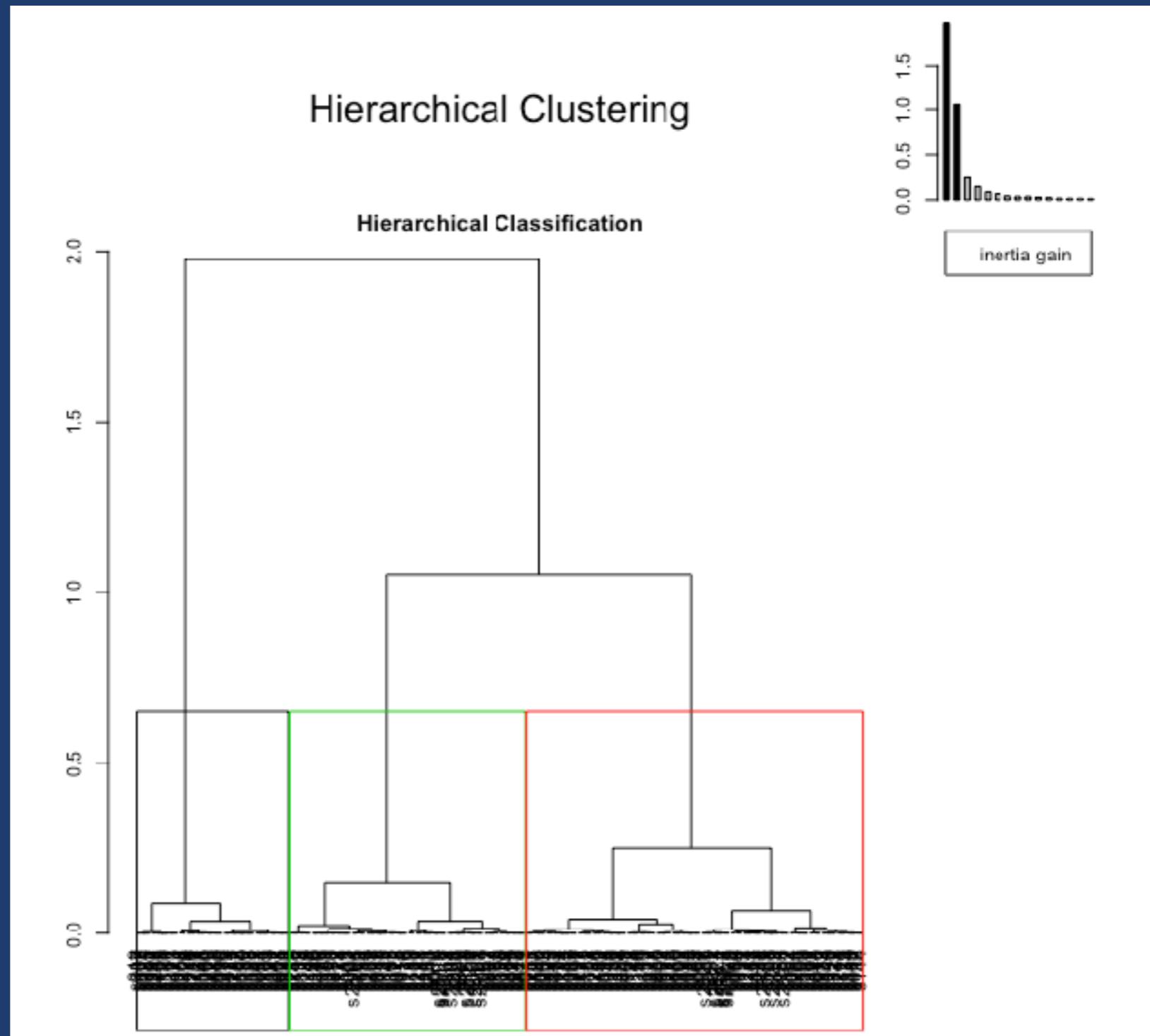
gr9 の刺激文の階層クラスタリング

- 3つのクラスターが形成されている
 - 黒 acceptables
 - 赤 questionables
 - 緑 unacceptables

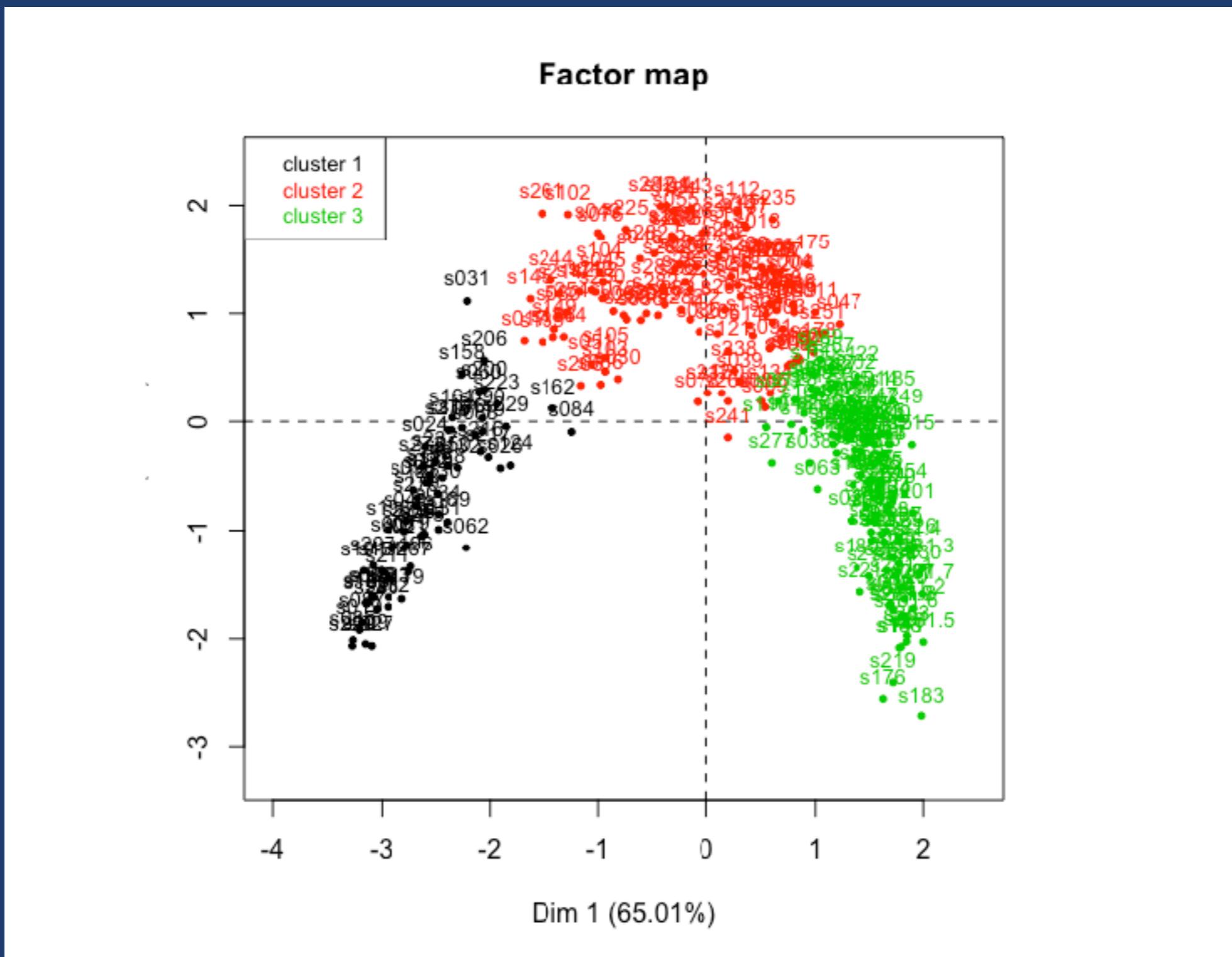


gr0-gr9 の刺激文の階層クラスタリング

- 3つのクラスターが形成されている
 - 黒 acceptables
 - 赤 questionables
 - 緑 unacceptables

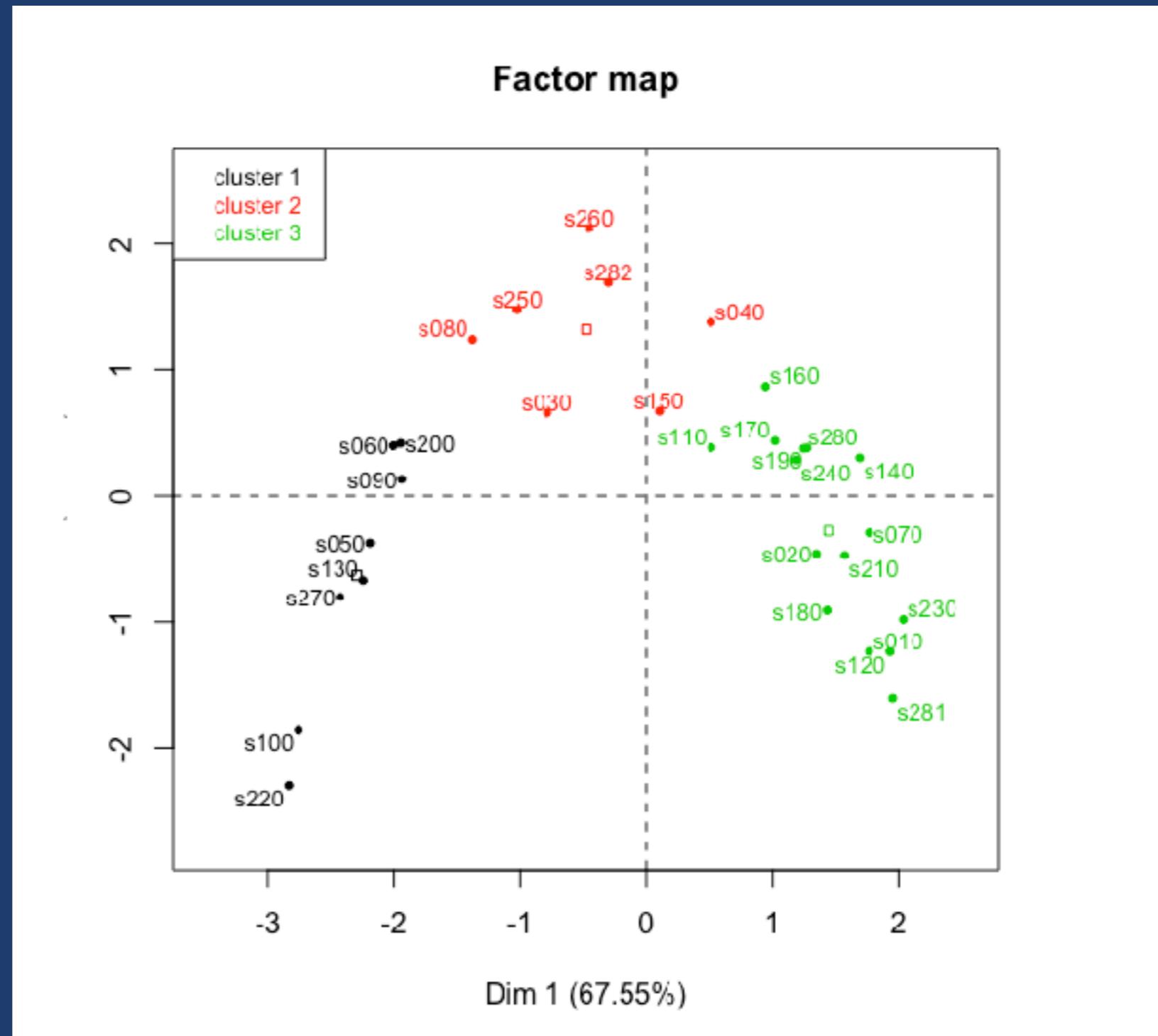


反応ポテンシャルの PCA



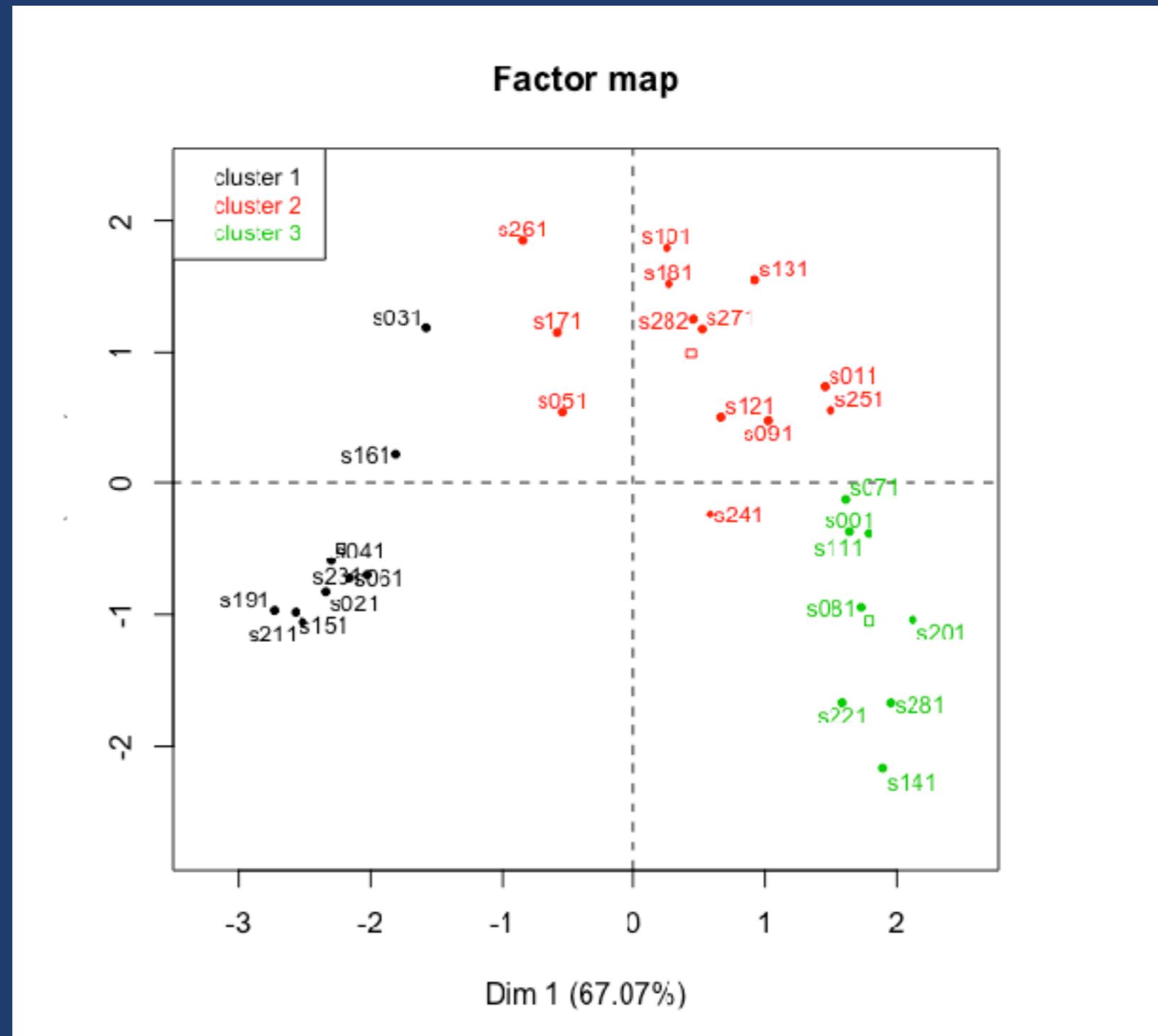
gr0 の刺激文の PCA (Factor Map)

- 3つのクラスターが形成されている
- 黒 acceptables
- 赤 questionables
- 緑 unacceptables



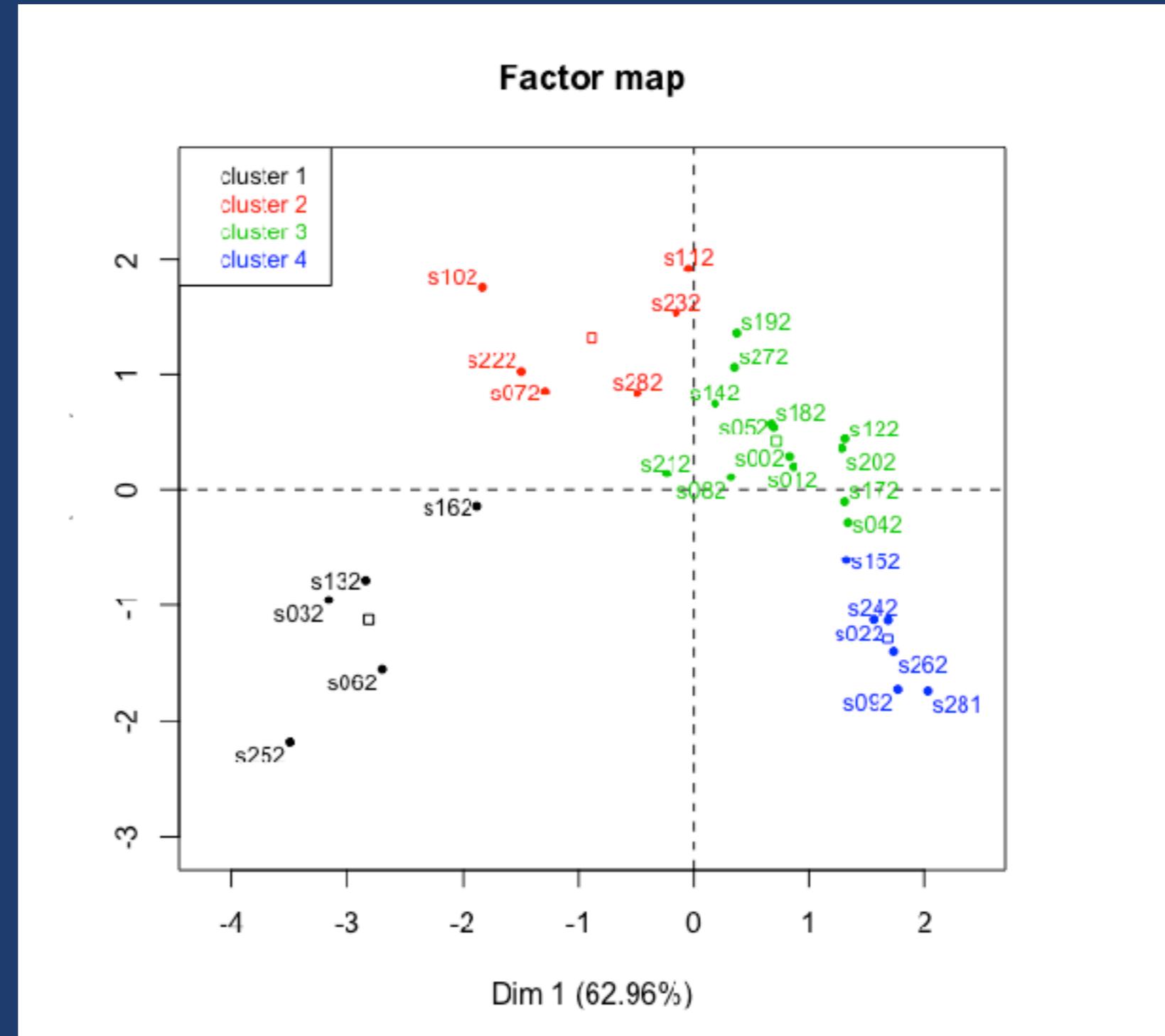
gr1 の刺激文の PCA (Factor Map)

- 3つのクラスターが形成されている
- 黒 acceptables
- 赤 questionables
- 緑 unacceptables



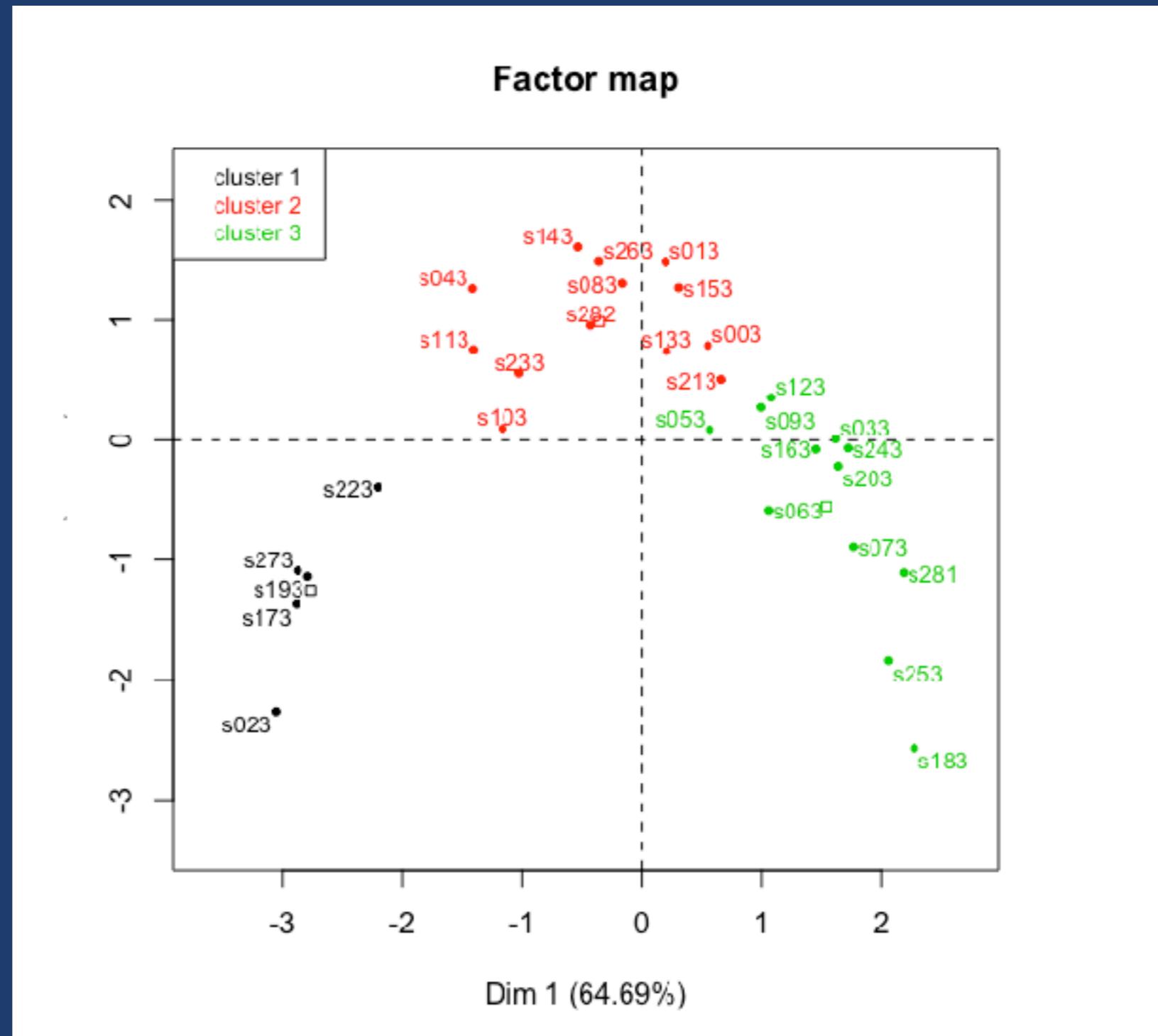
gr2 の刺激文の PCA (Factor Map)

- 4つのクラスターが形成されている
- 黒 acceptables
- 赤 weakly questionables
- 緑 strongly questionables
- 青 unacceptables



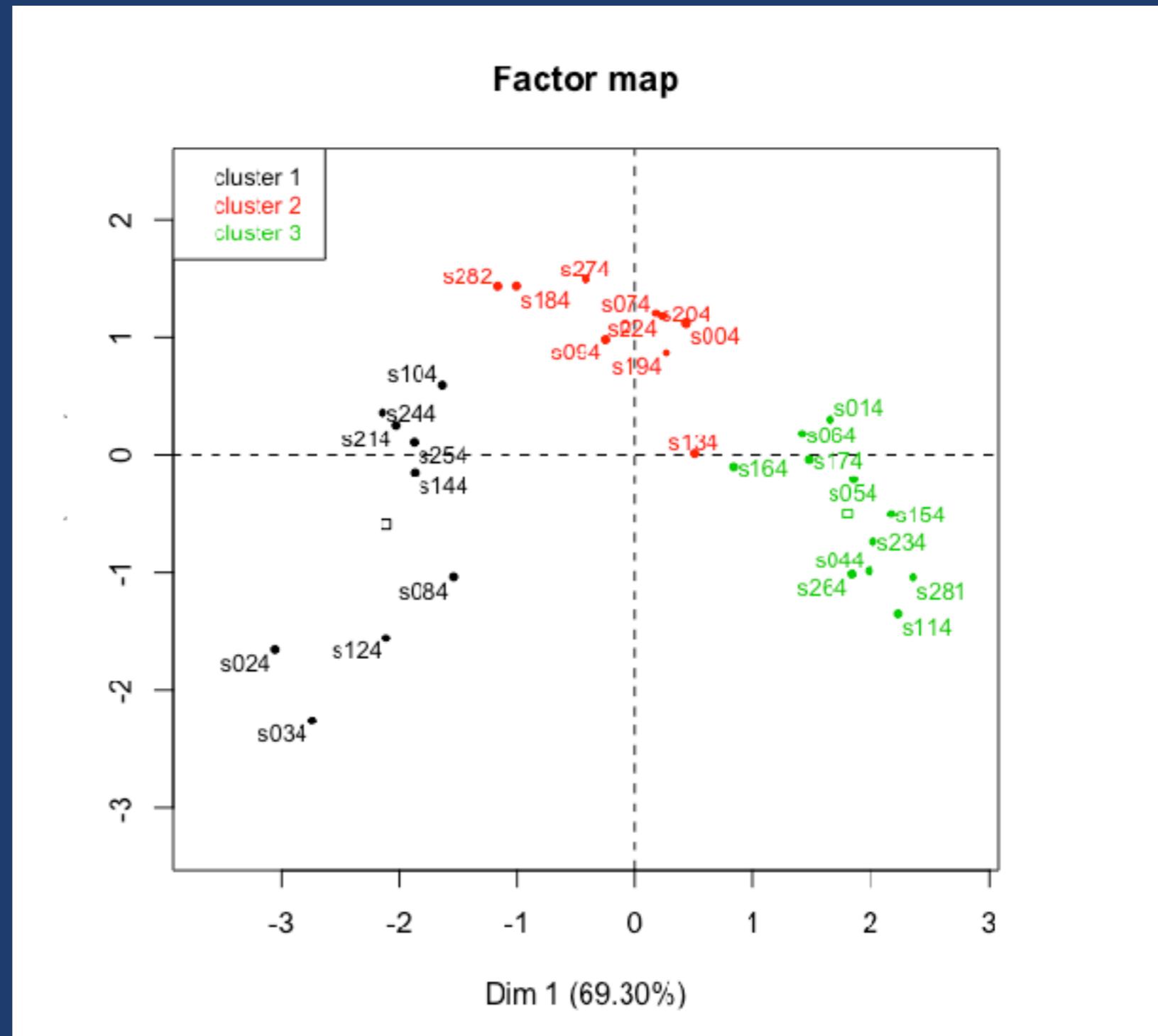
gr3 の刺激文の PCA (Factor Map)

- 3つのクラスターが形成されている
- 黒 acceptables
- 赤 questionables
- 緑 unacceptables



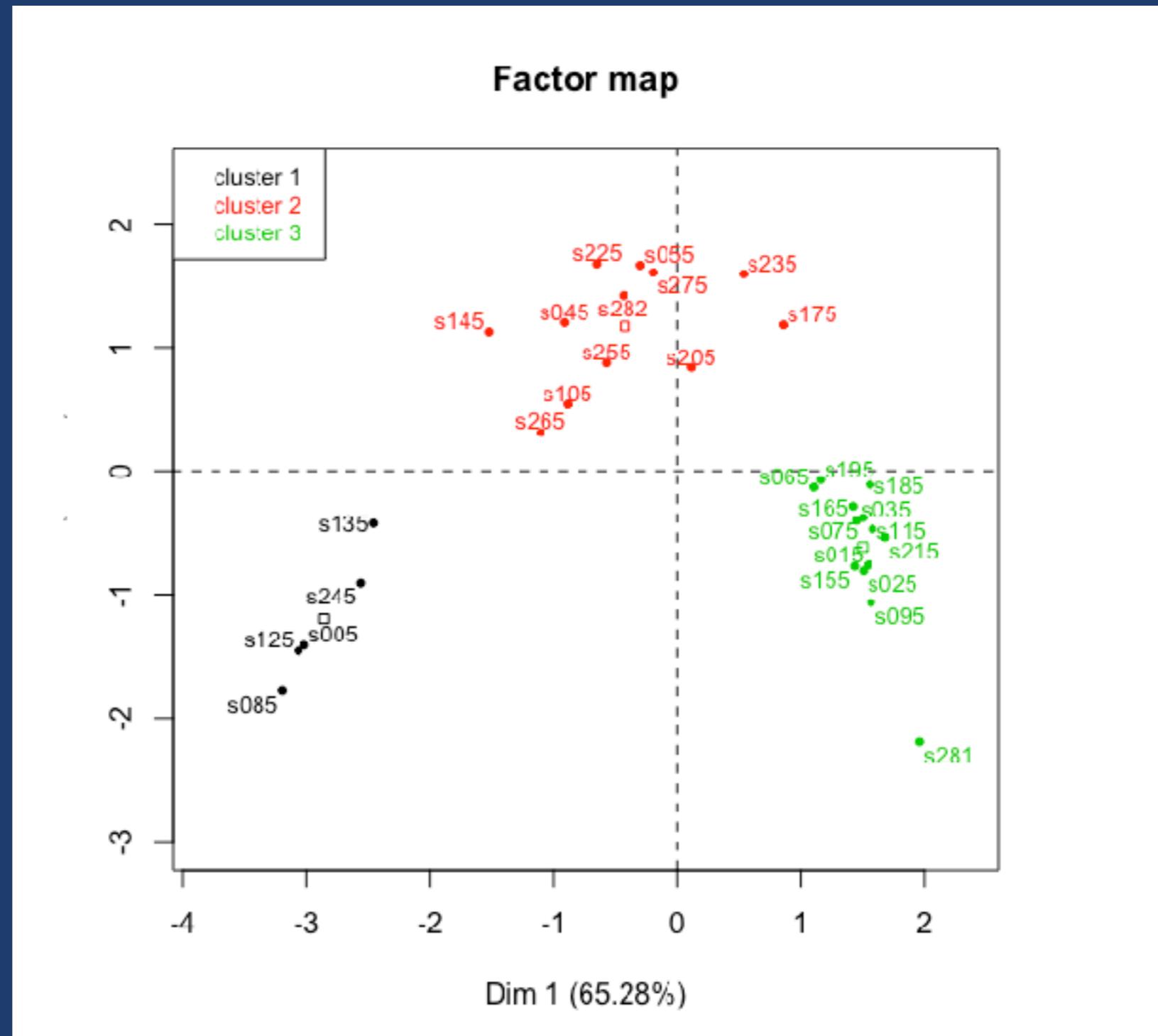
gr4 の刺激文の PCA (Factor Map)

- 3つのクラスターが形成されている
- 黒 acceptables
- 赤 questionables
- 緑 unacceptables



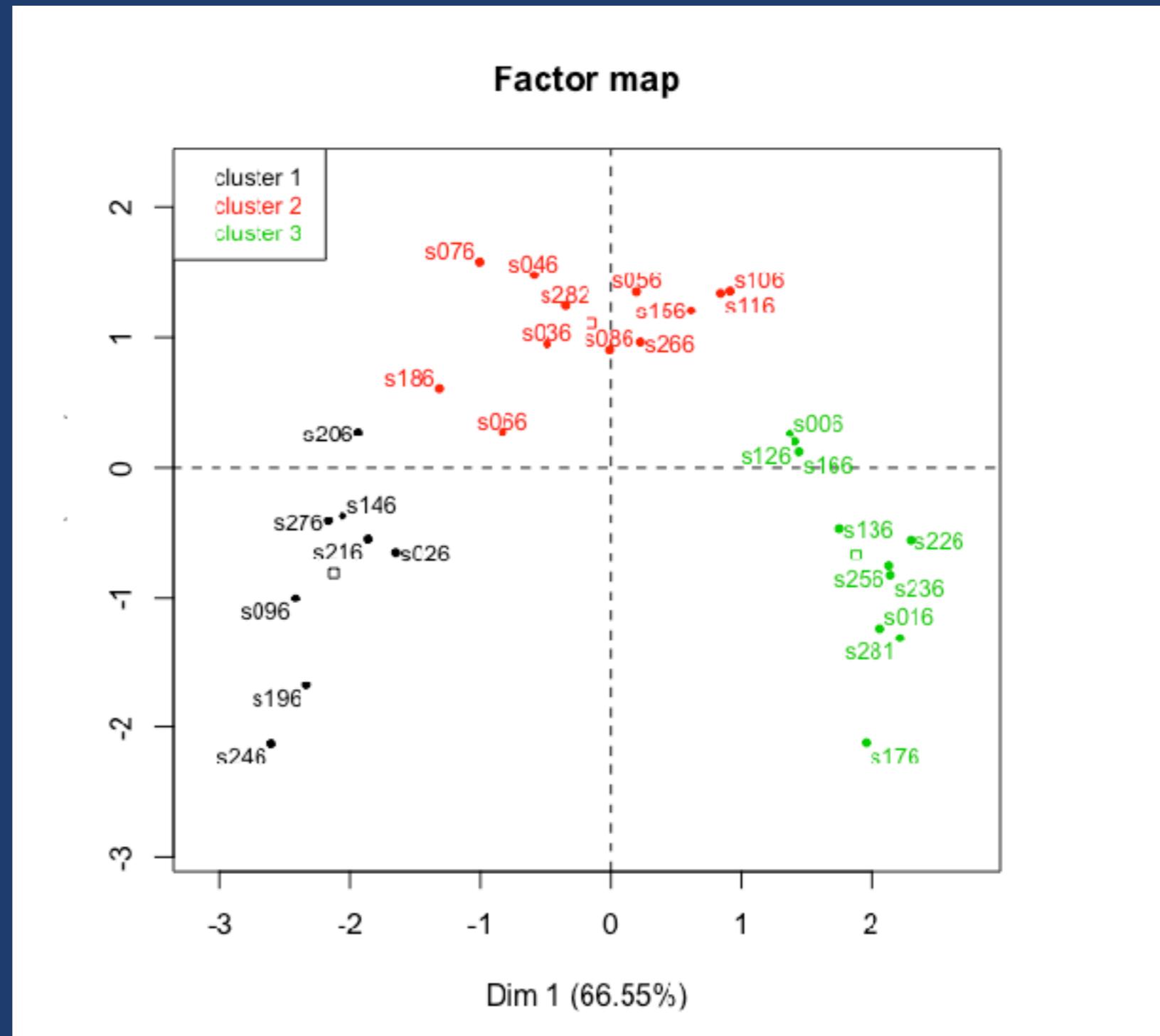
gr5 の刺激文の PCA (Factor Map)

- 3つのクラスターが形成されている
- 黒 acceptables
- 赤 questionables
- 緑 unacceptables



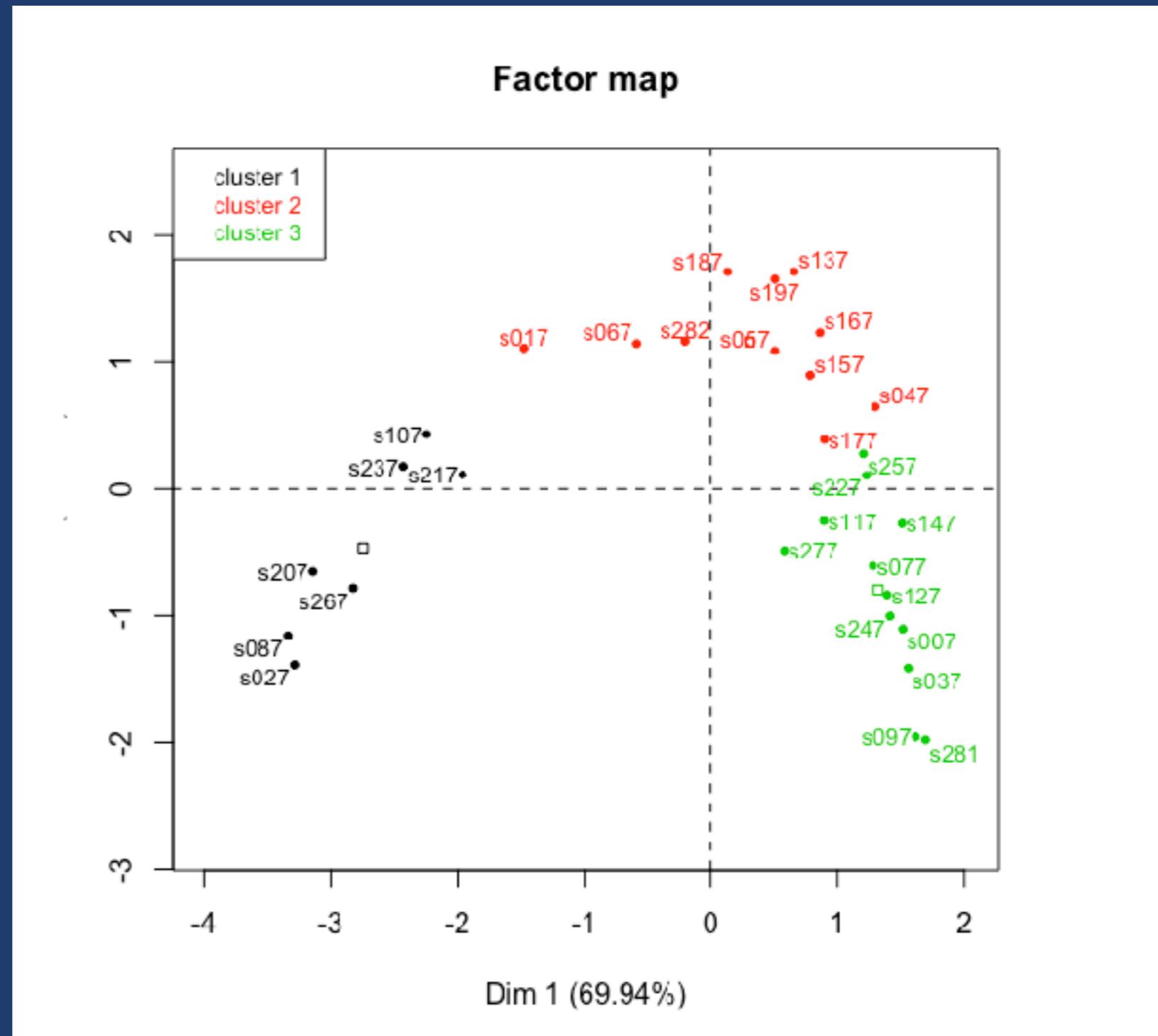
gr6 の刺激文の PCA (Factor Map)

- 3つのクラスターが形成されている
- 黒 acceptables
- 赤 questionables
- 緑 unacceptables



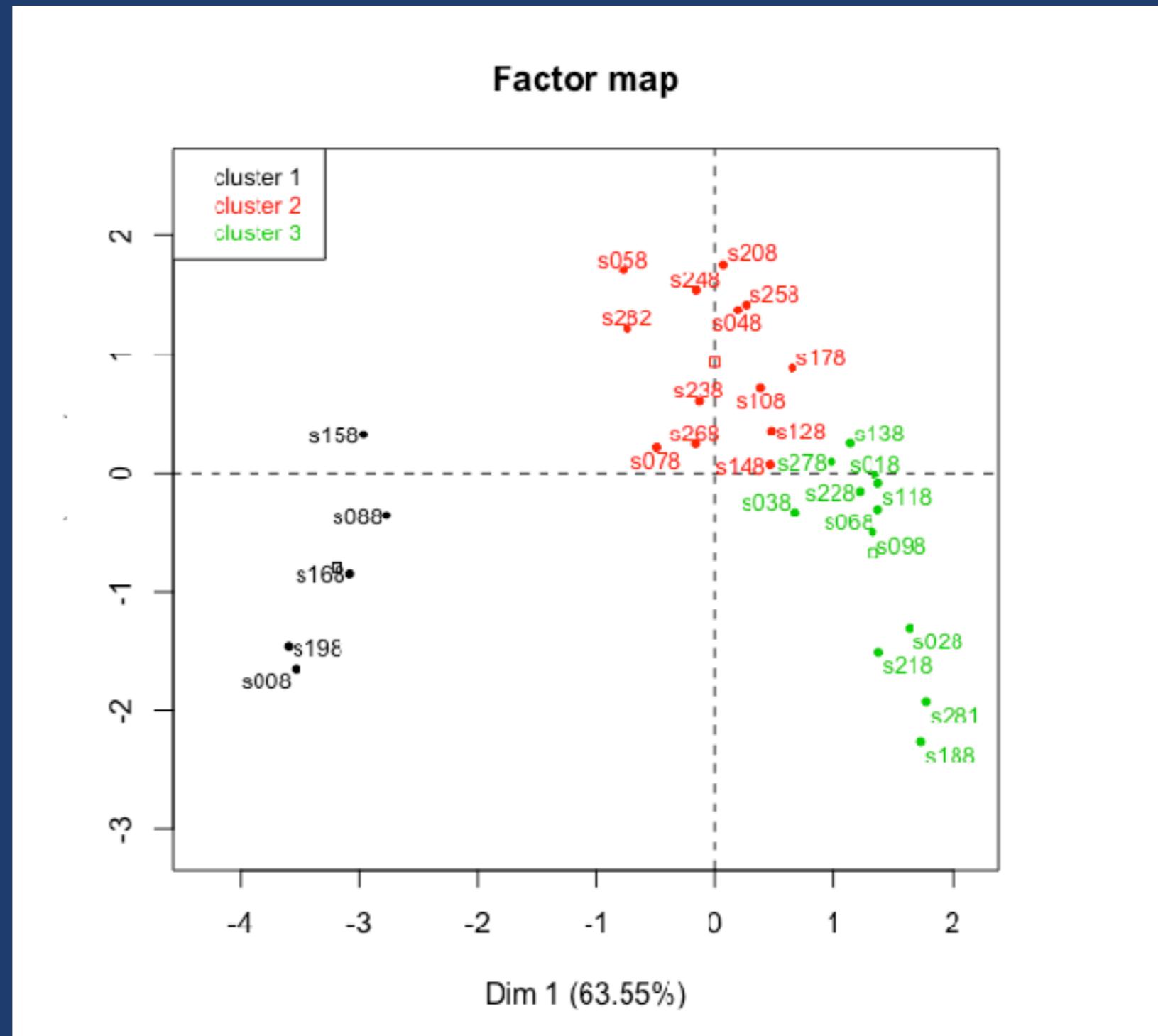
gr7 の刺激文の PCA (Factor Map)

- 3つのクラスターが形成されている
- 黒 acceptables
- 赤 questionables
- 緑 unacceptables



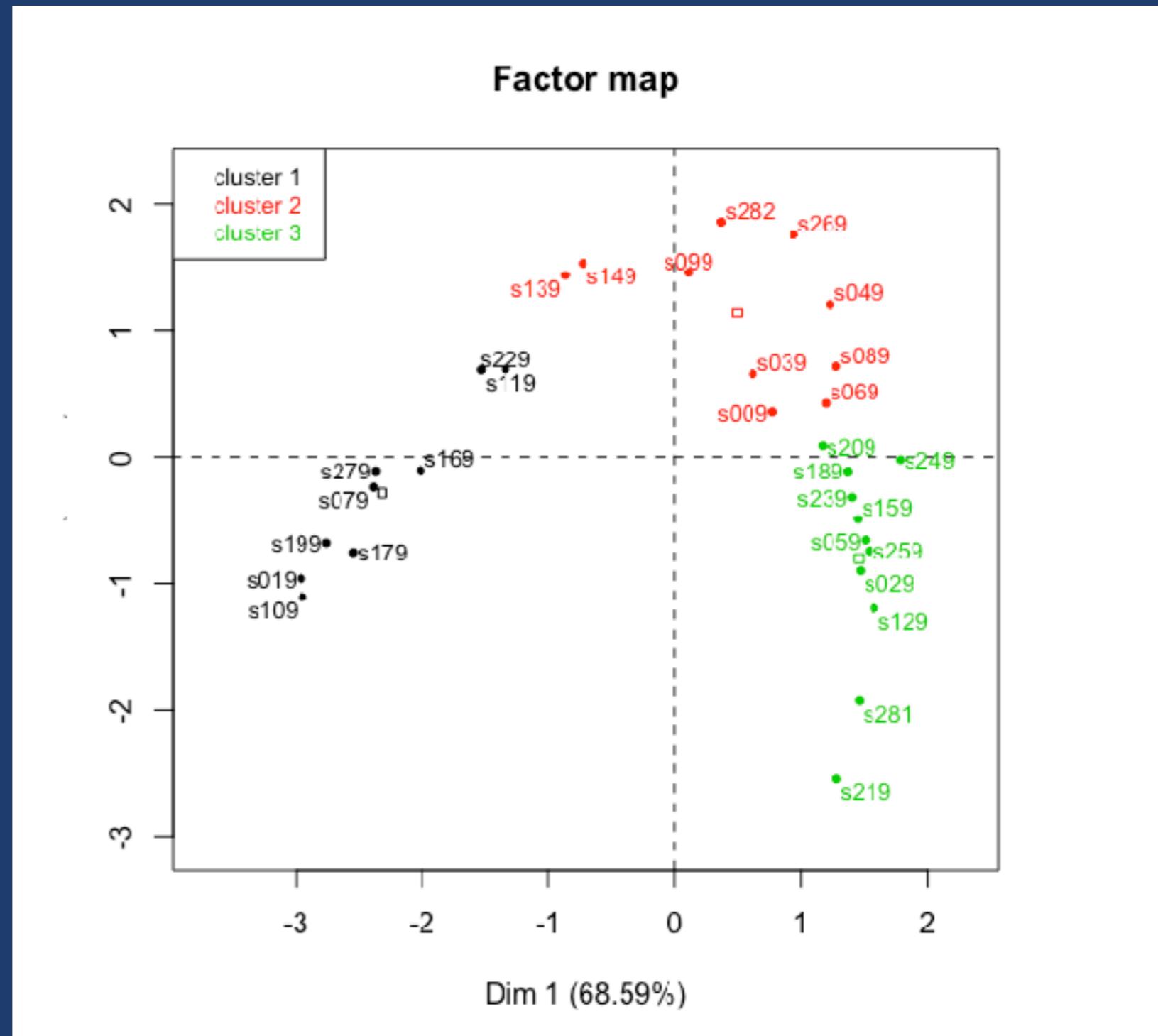
gr8 の刺激文の PCA (Factor Map)

- 3つのクラスターが形成されている
- 黒 acceptables
- 赤 questionables
- 緑 unacceptables



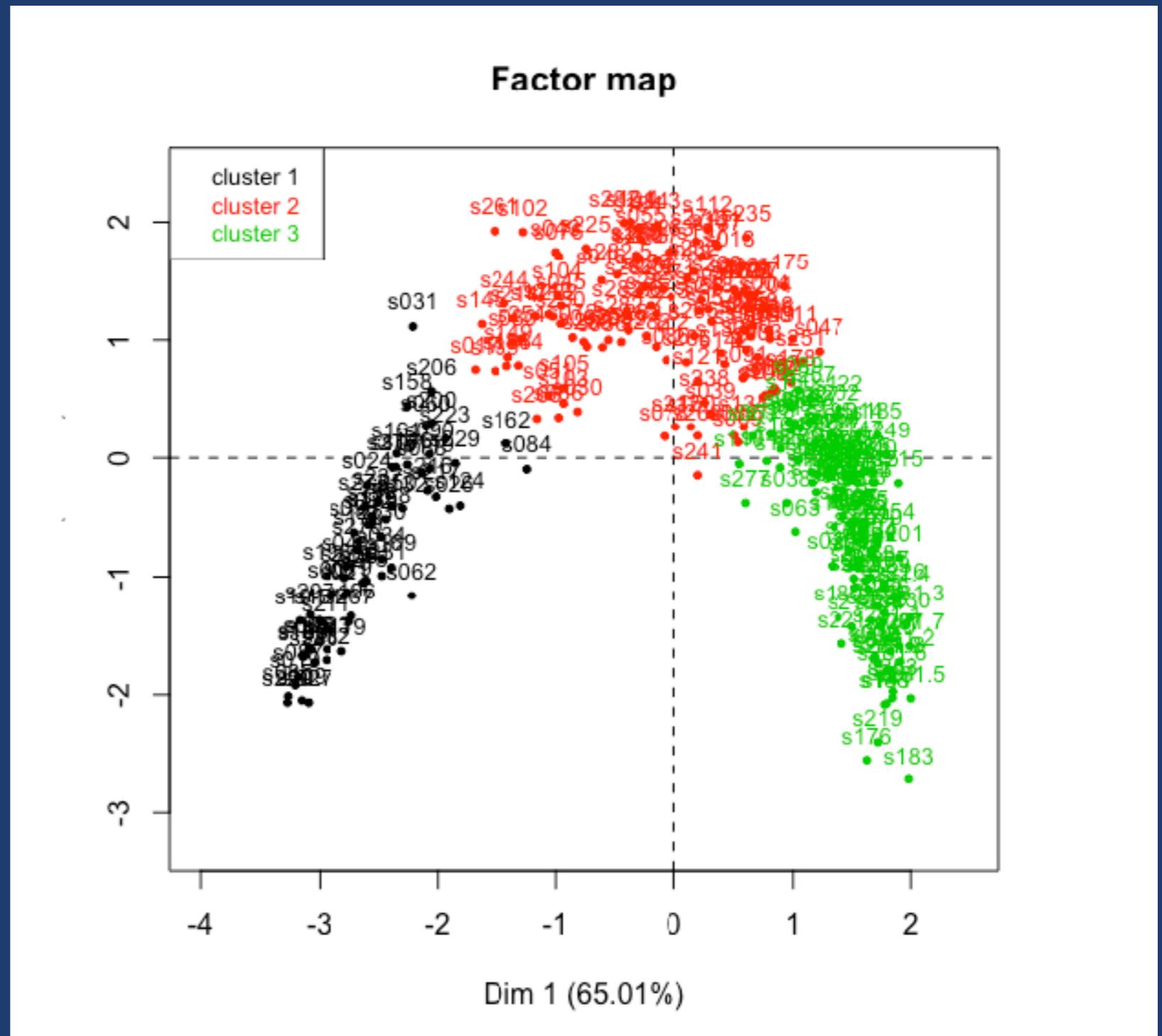
gr9 の刺激文の PCA (Factor Map)

- 3つのクラスターが形成されている
 - 黒 acceptables
 - 赤 questionables
 - 緑 unacceptables



gr0-gr9 の刺激文の PCA (Factor Map)

- 3つのクラスターが形成されている
- 黒 acceptables
- 赤 questionables
- 緑 unacceptables



興味深い点 1/2

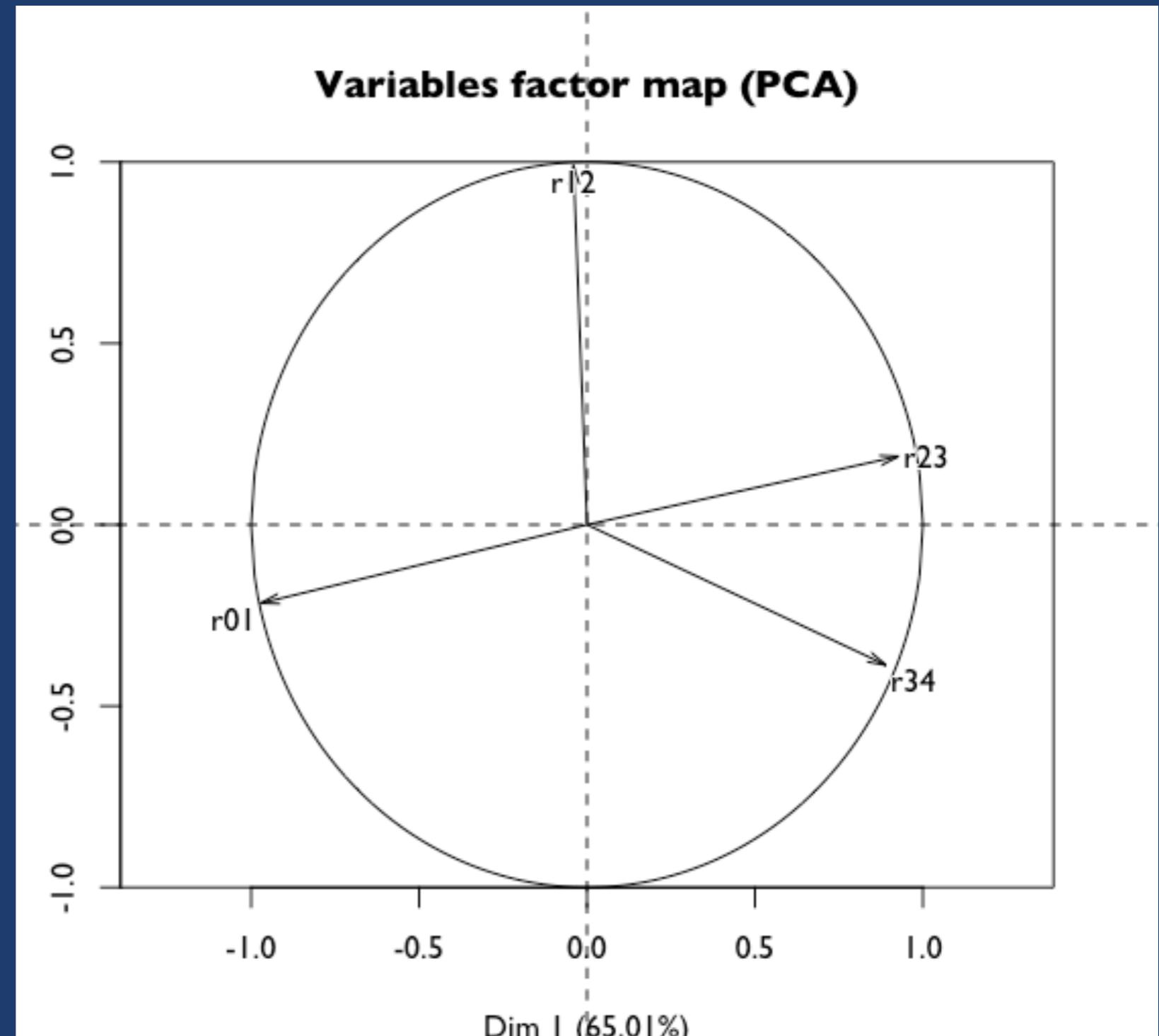
- (1) gr0-gr9 のどの刺激文の組に対しても、だいたい同じ反応の地図が得られている
- (2) それは grN の反応の地図と相同である
- (3) gr.i ($i = 0, 1, \dots, 9$) で s281.i と s282.i はだいたい同じ位置にあるが、まったく同じではない
- s281.i の位置の変異は、s282.i の位置の変異より小

興味深い点 2/2

- (1)-(3) の説明
- だいたい同じ位置にあるのは、事例の間の位相が保存されている証拠
- 位置がずれるのは文脈効果が理由

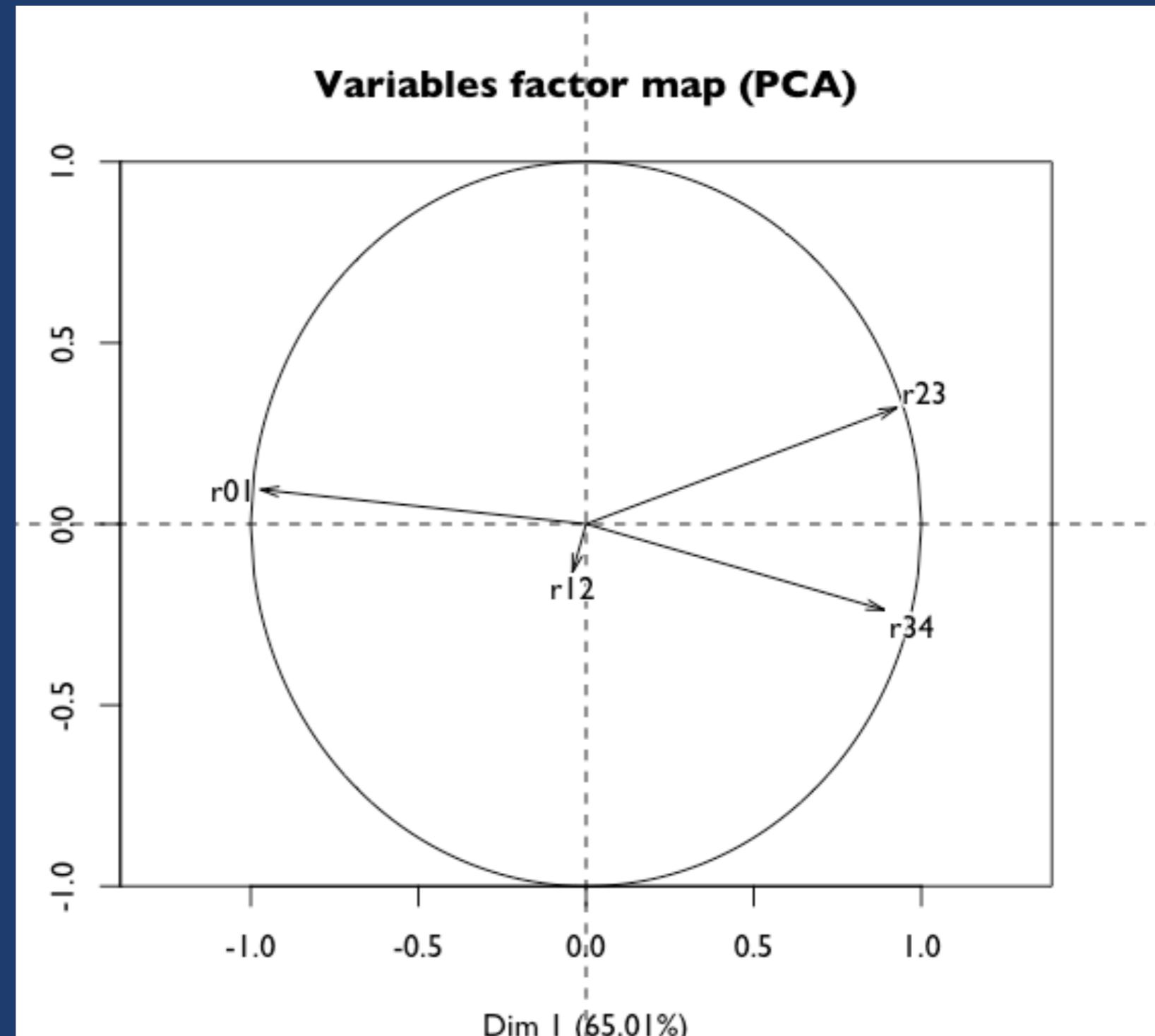
gr0-gr9 の刺激文の PCA (Factor Map) 1/3

- 説明力
 - Dim 1 と Dim 2 で反応の 95.33%を説明
- Dim 1, Dim 2に対する反応ポテンシャルの関係
 - $r[0,1]$ と $r[2,3]$ とが真逆で, Dim 1 にほぼ対応
 - それに対して $r[1,2]$ がほぼ直行で, Dim 2 にほぼ対応
 - $r[3,4]$ は $r[2,3]$ と順相関し, $r[1,2]$ と逆相関



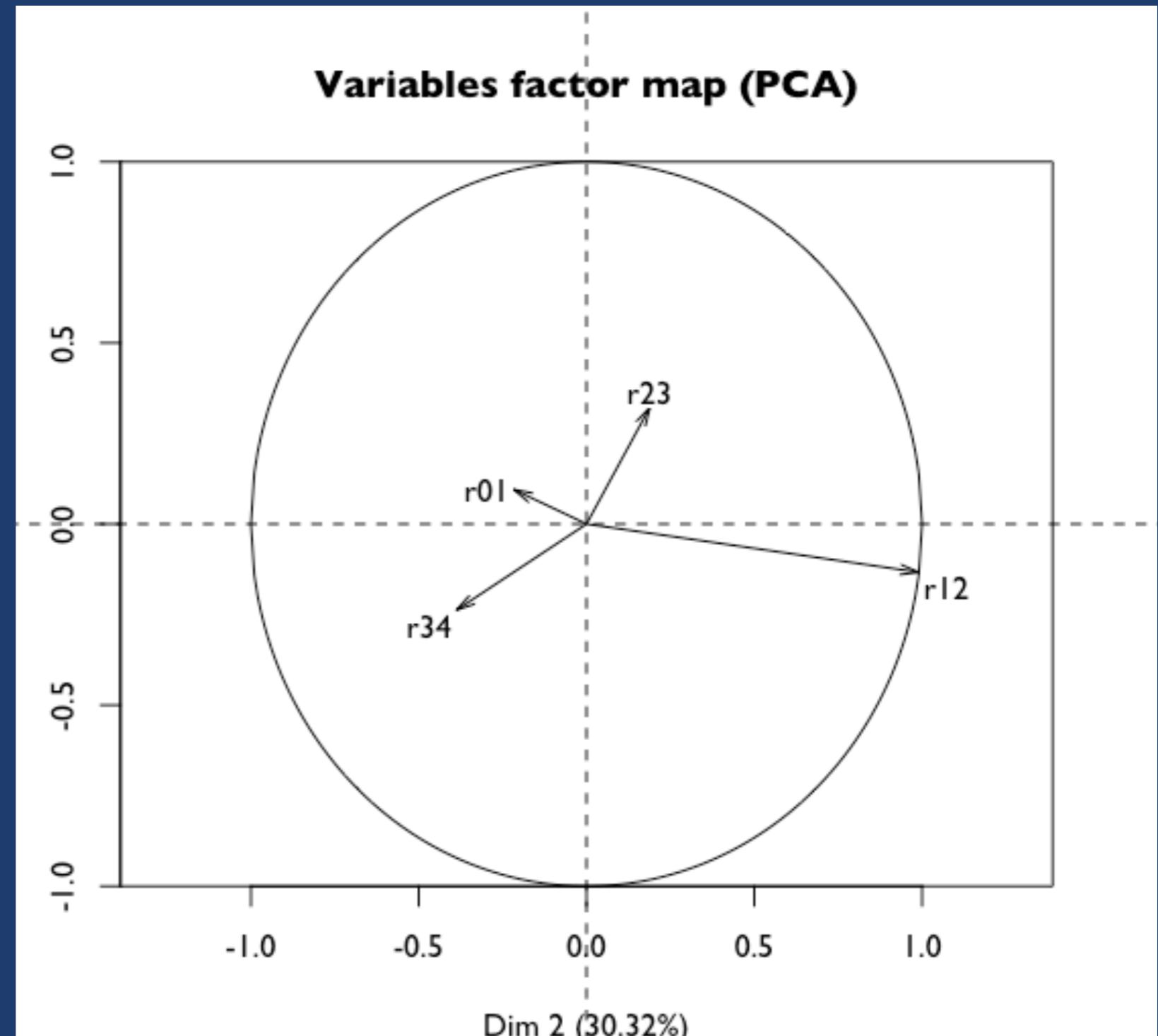
gr0-gr9 の刺激文の PCA (Factor Map) 2/3

- 説明力
 - Dim 1 と Dim 3 で反応の 69.68%を説明
- Dim 1, Dim 3 に対する反応 ポテンシャルの関係
 - $r[0,1)$ と $r[3,\infty)$ とが真逆 で Dim 1 にほぼ対応
 - $r[2,3)$ は $r[3,\infty)$ と順相関
 - それに対して $r[1,2)$ は Dim 1, Dim2 とはほぼ相関なし



gr0-gr9 の刺激文の PCA (Factor Map) 3/3

- 説明力
 - Dim 2 と Dim 3 で反応の 30.99%を説明
- Dim 2, Dim 3 に対する反応ポテンシャルの関係
 - $r[1,2]$ は D2 と強く順相関
 - $r[2,3]$ と $r[3,\infty)$ とが逆向きに Dim 3 に弱く相関
 - $r[0,1]$ は Dim 2, Dim 3 にはほとんど効果をもたない

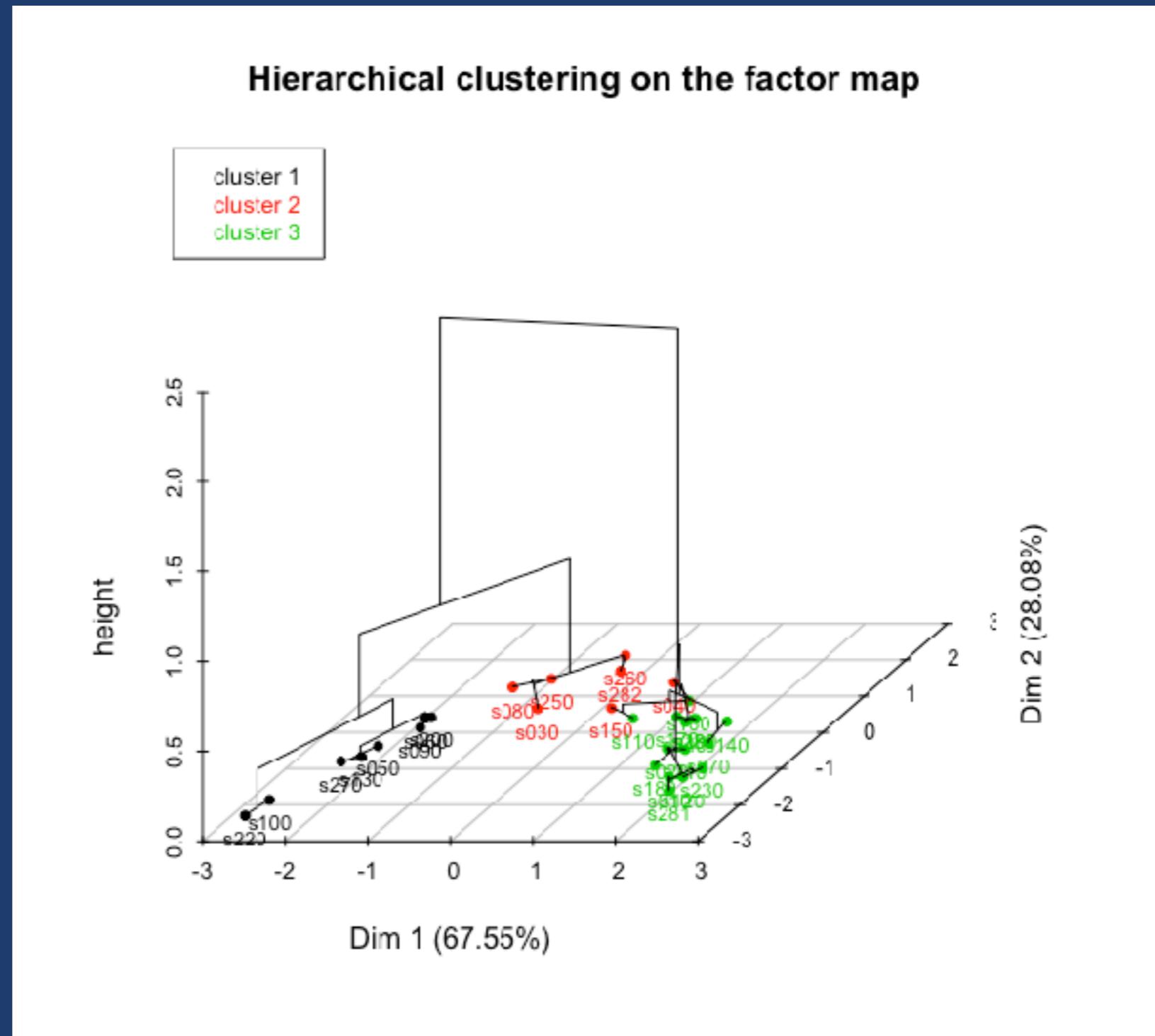


PCA の次元の解釈

- Dim 1 (65.1%) が逸脱度 (= 非容認度) に対応しているのは明らかだが, Dim 2 (30.3%), Dim 3 (4.7%) の解釈は厄介
- おそらく
 - Dim 2 は文構造の複雑さ \approx 解析処理負荷の大きさ
 - Dim 3 は理解に意外さの混じる度合い (\approx 新規性?)
- が係わっていそう

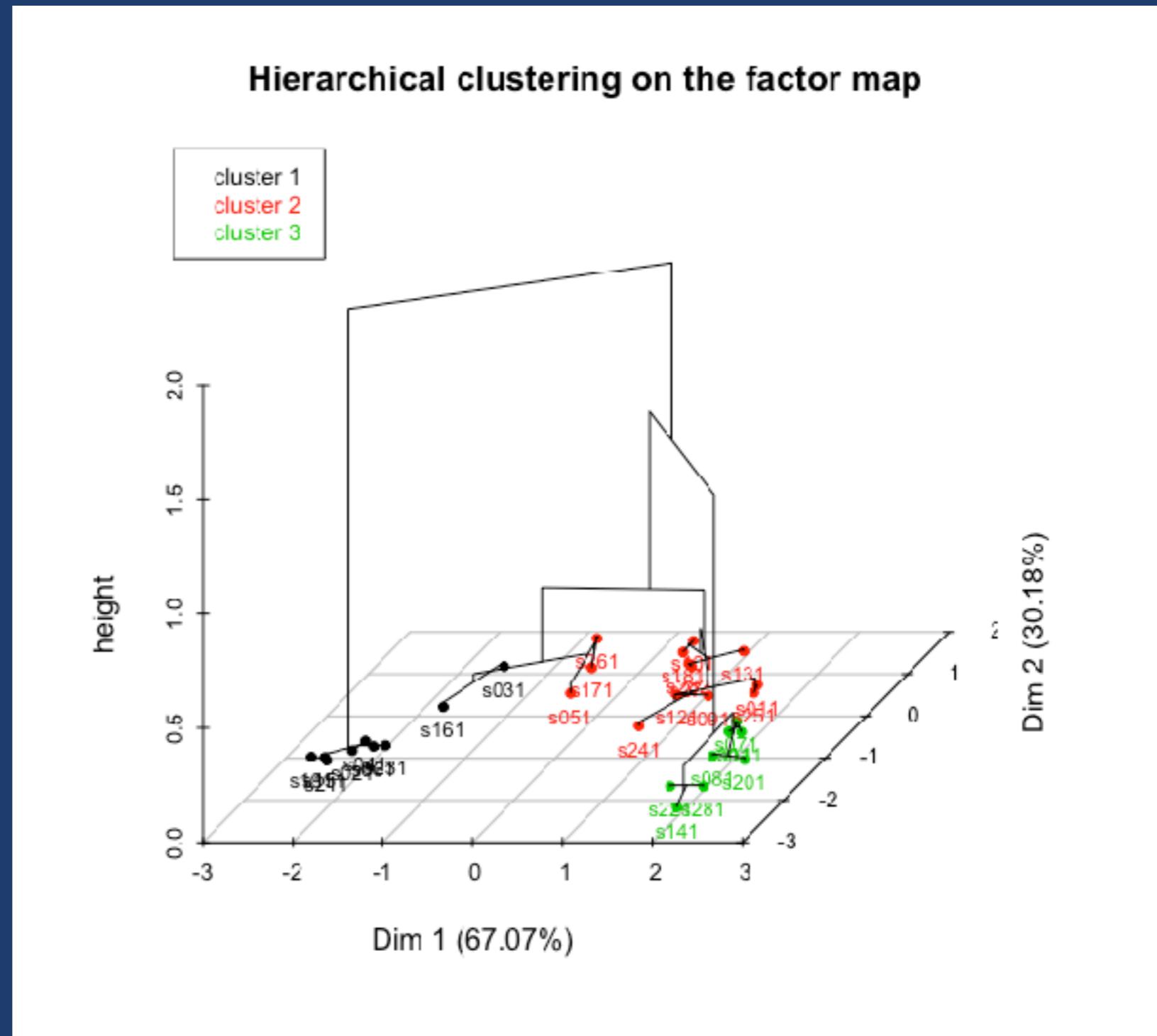
gr0 の刺激文の HC x PCA

- 3つのクラスターの階層化
 - 黒 acceptables
 - 赤 questionables
 - 緑 unacceptables



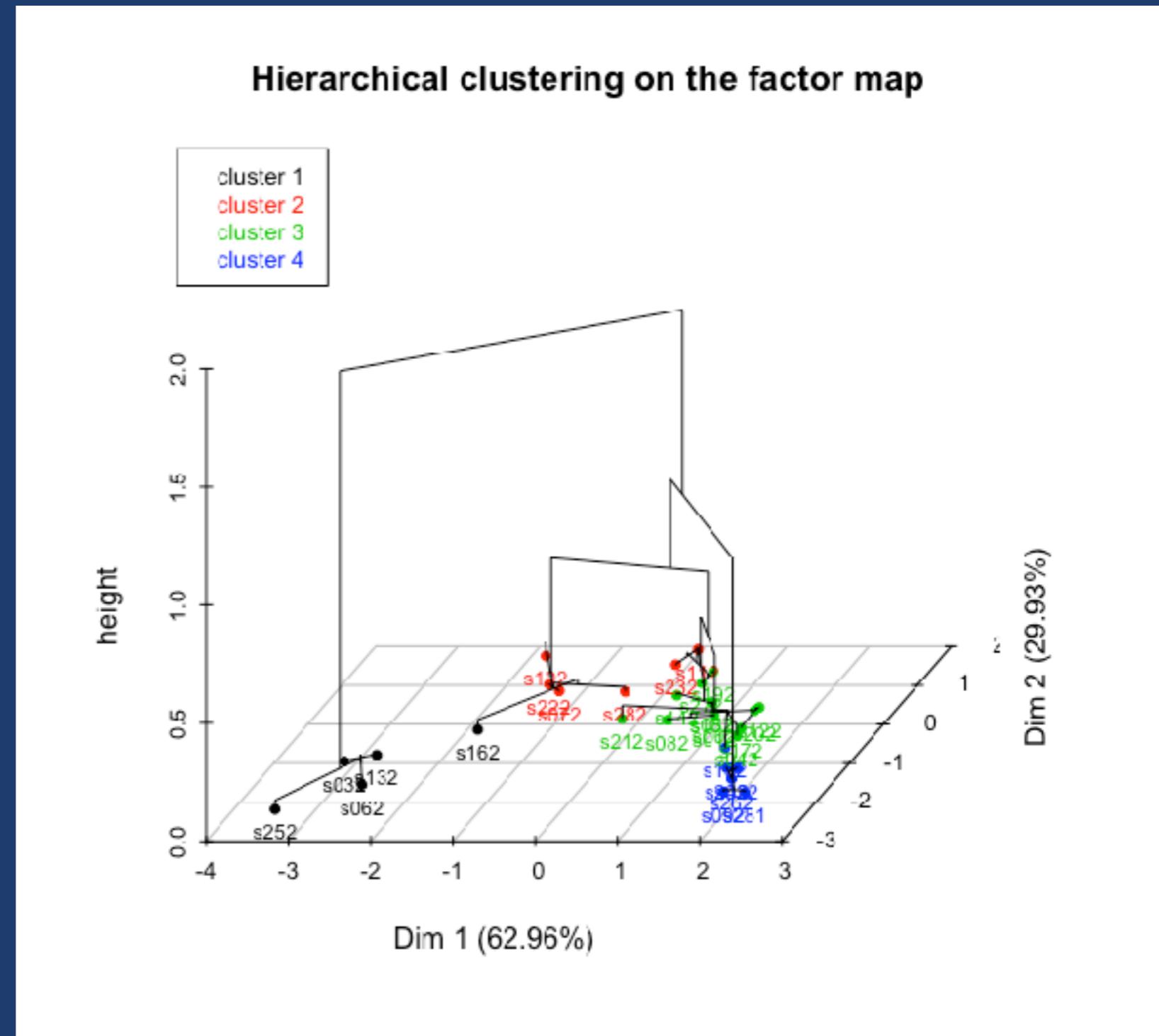
gr1 の刺激文の HC x PCA

- 3つのクラスターの階層化
 - 黒 acceptables
 - 赤 questionable
 - 緑 unacceptable



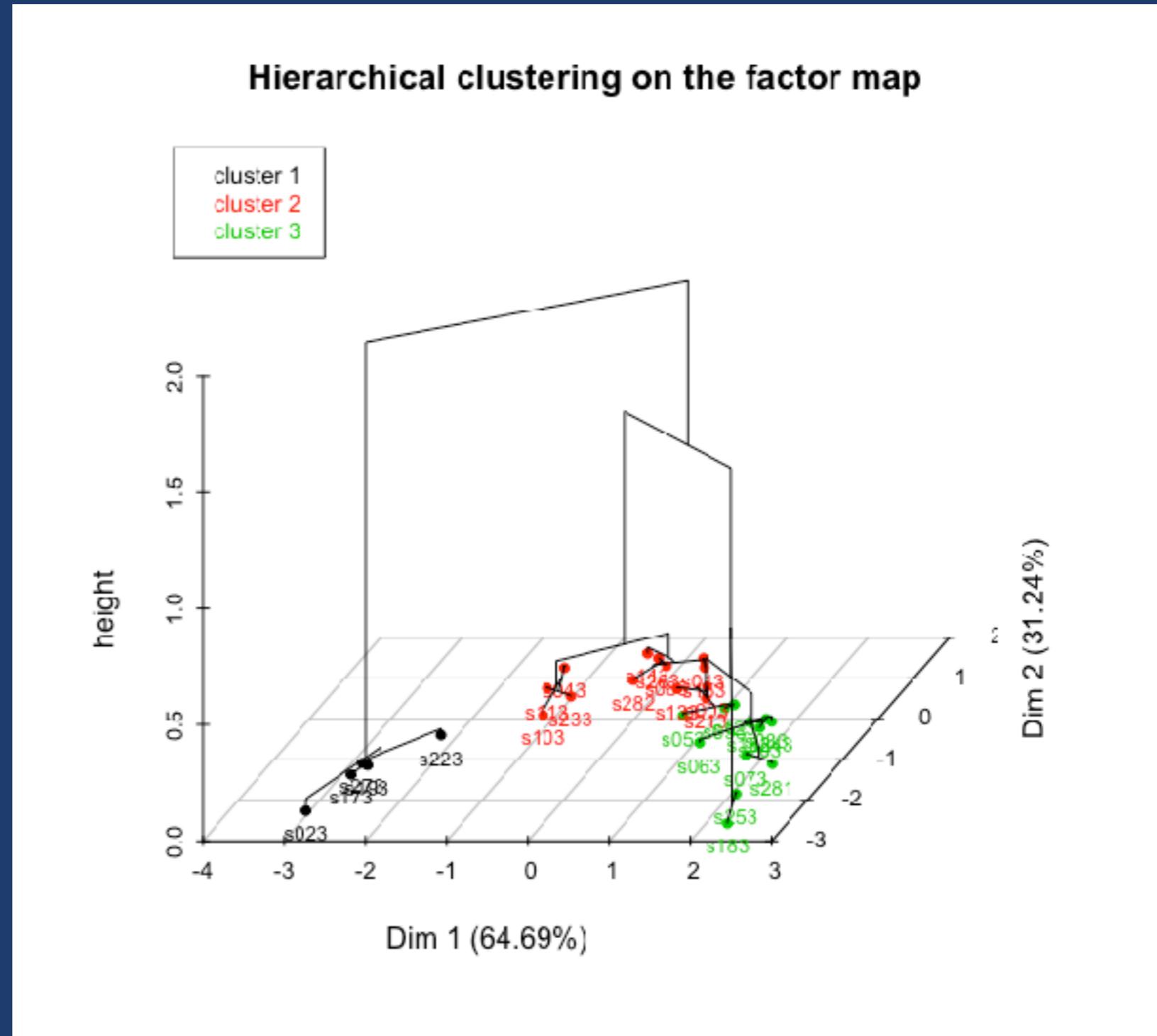
gr2 の刺激文の HC x PCA

- 3つのクラスターの階層化
 - 黒 acceptables
 - 赤 questionables 1
 - 緑 questionables 2
 - 青 unacceptables



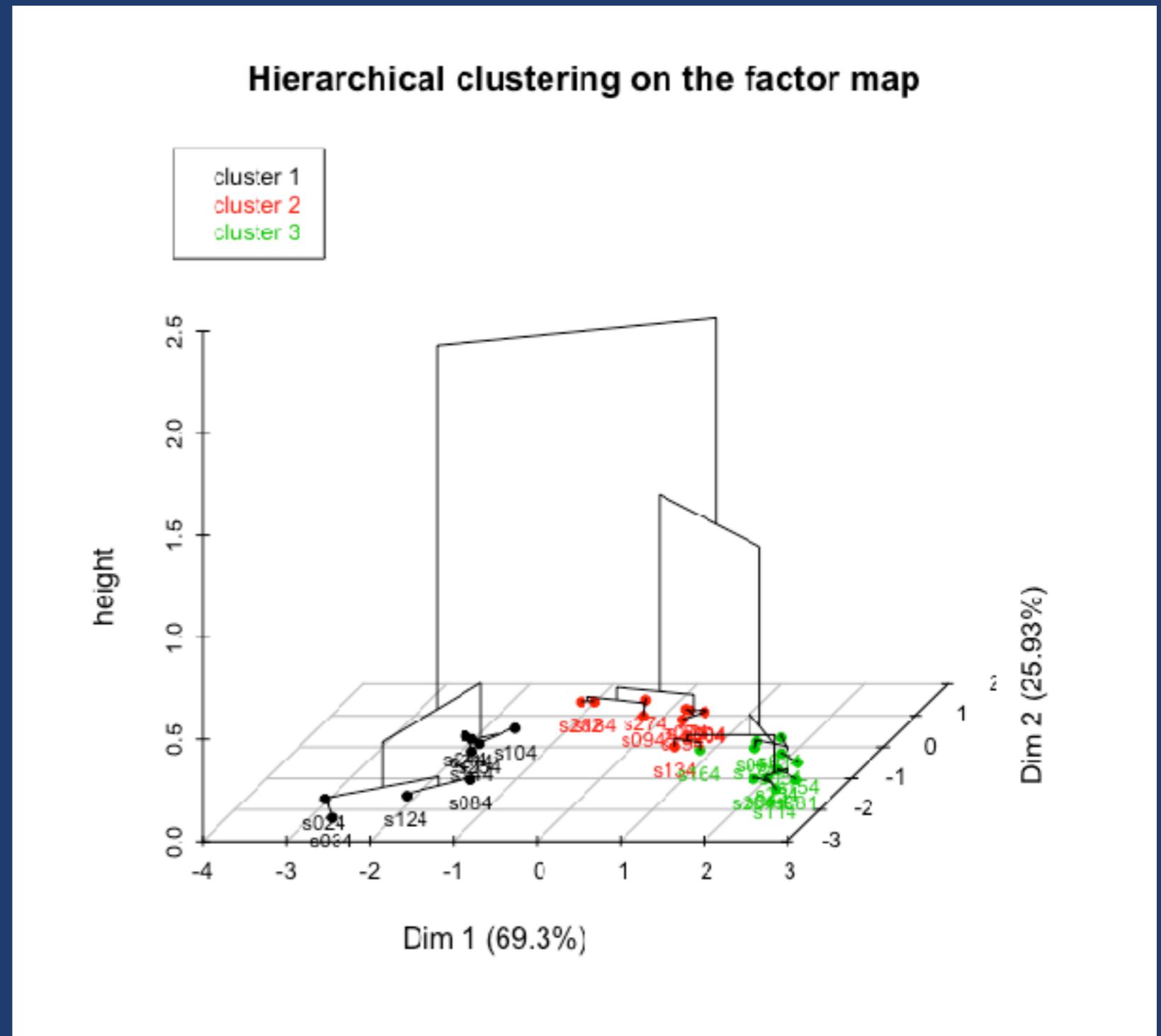
gr3 の刺激文の HC x PCA

- 3つのクラスターの階層化
 - 黒 acceptables
 - 赤 questionables
 - 緑 unacceptables



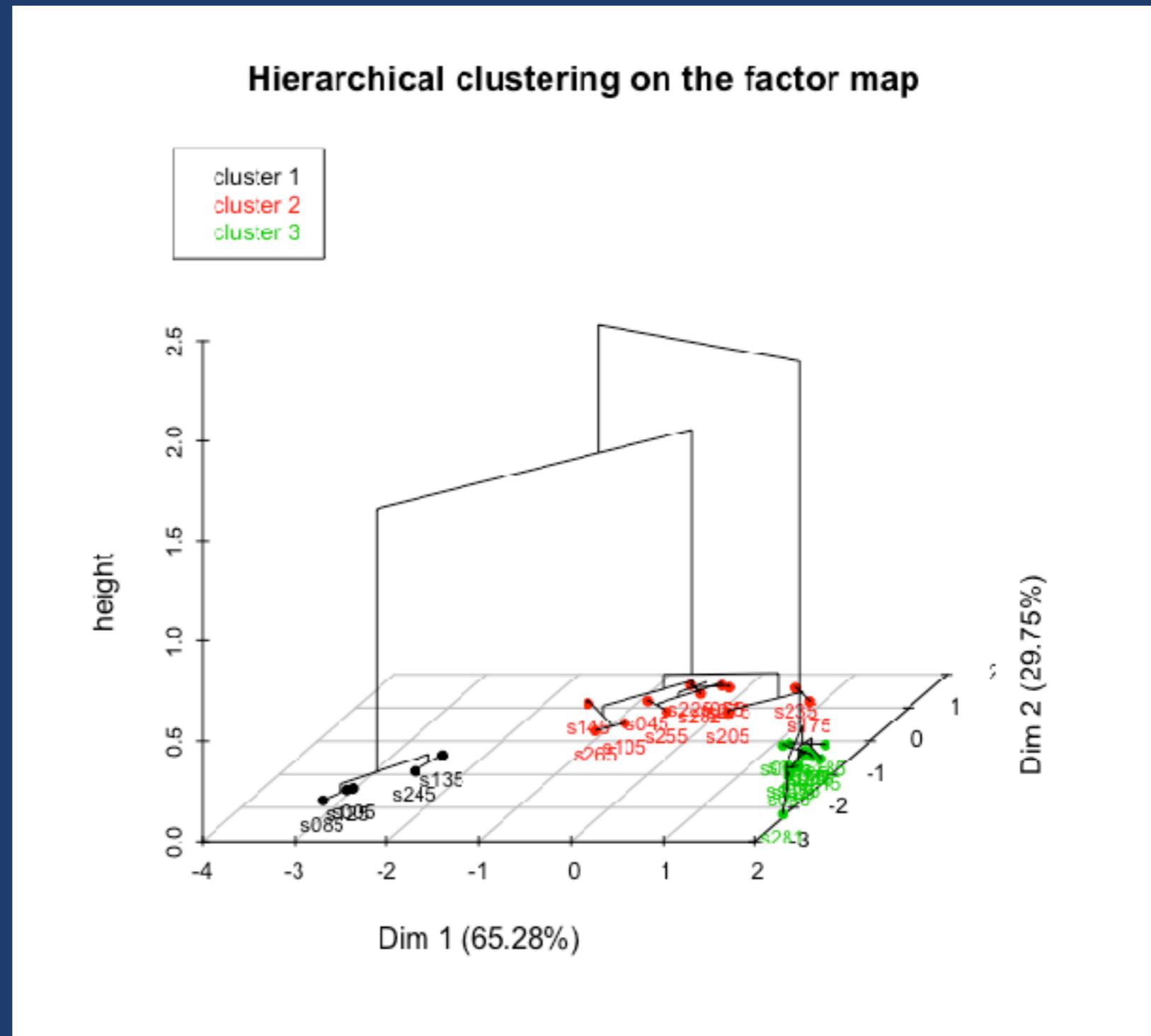
gr4 の刺激文の HC x PCA

- 3つのクラスターの階層化
 - 黒 acceptables
 - 赤 questionables
 - 緑 unacceptables



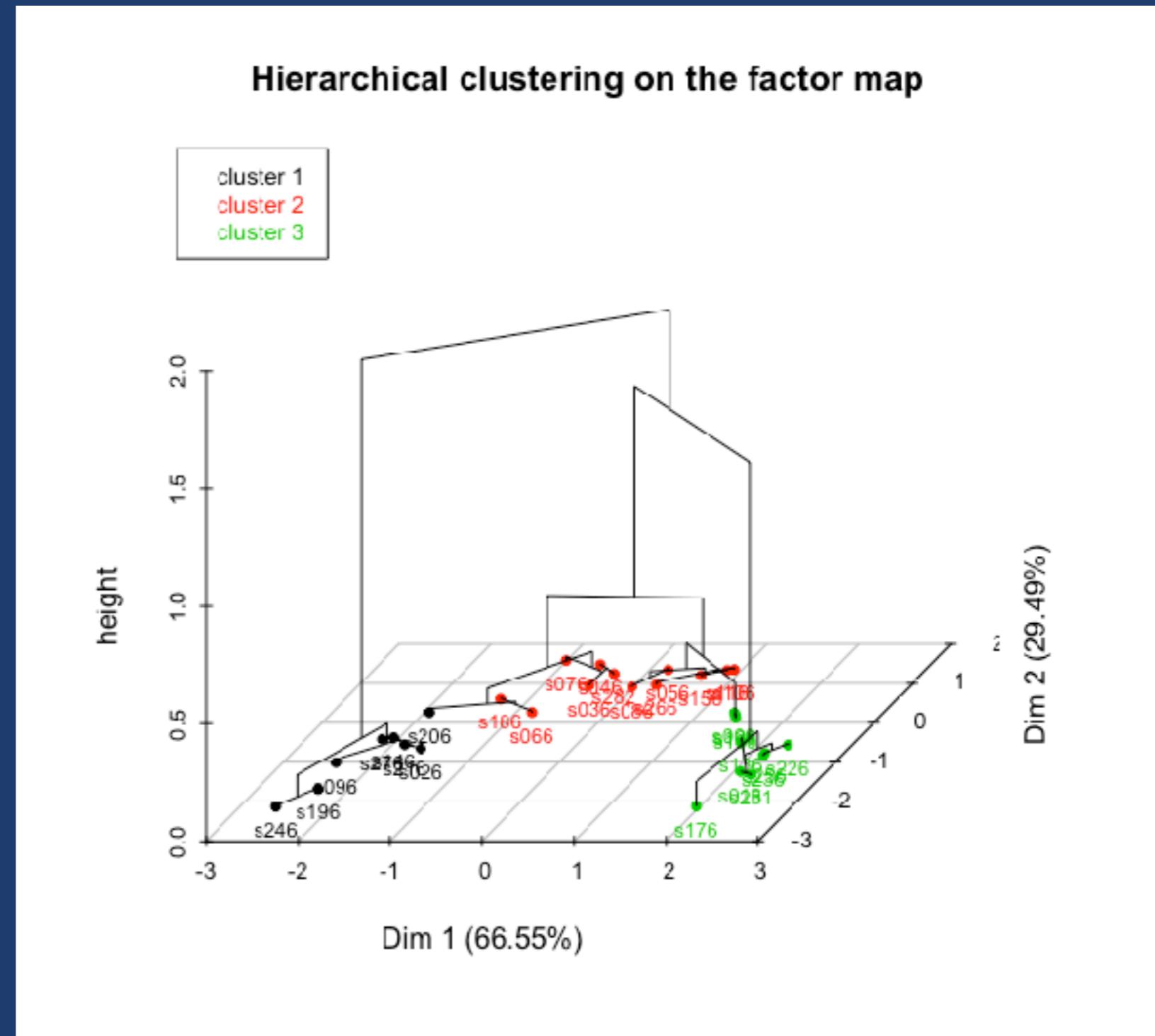
gr5 の刺激文の HC x PCA

- 3つのクラスターの階層化
 - 黒 acceptables
 - 赤 questionables
 - 緑 unacceptables



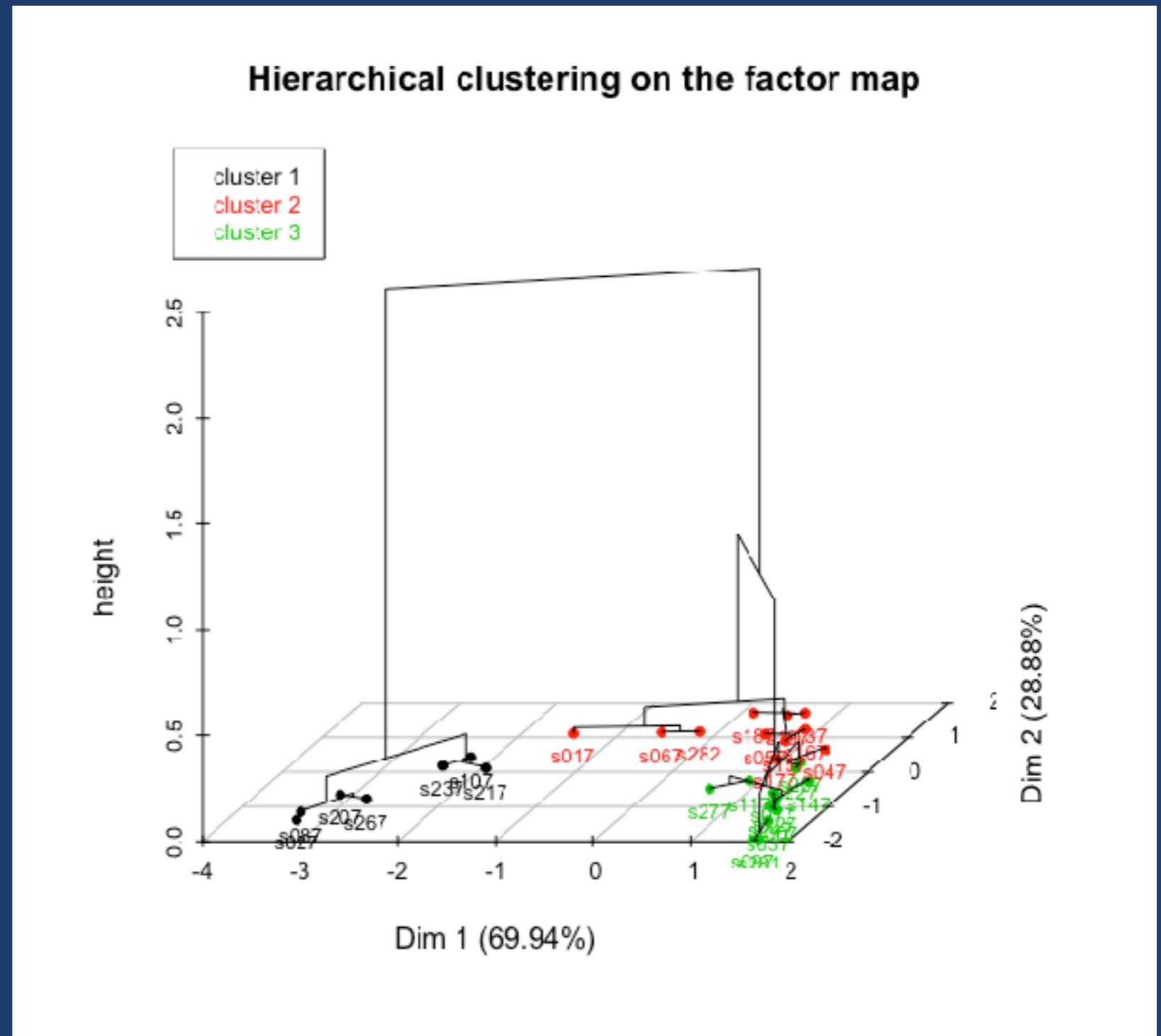
gr6 の刺激文の HC x PCA

- 3つのクラスターの階層化
 - 黒 acceptables
 - 赤 questionables
 - 二つの下位クラスターあり
 - 緑 unacceptables



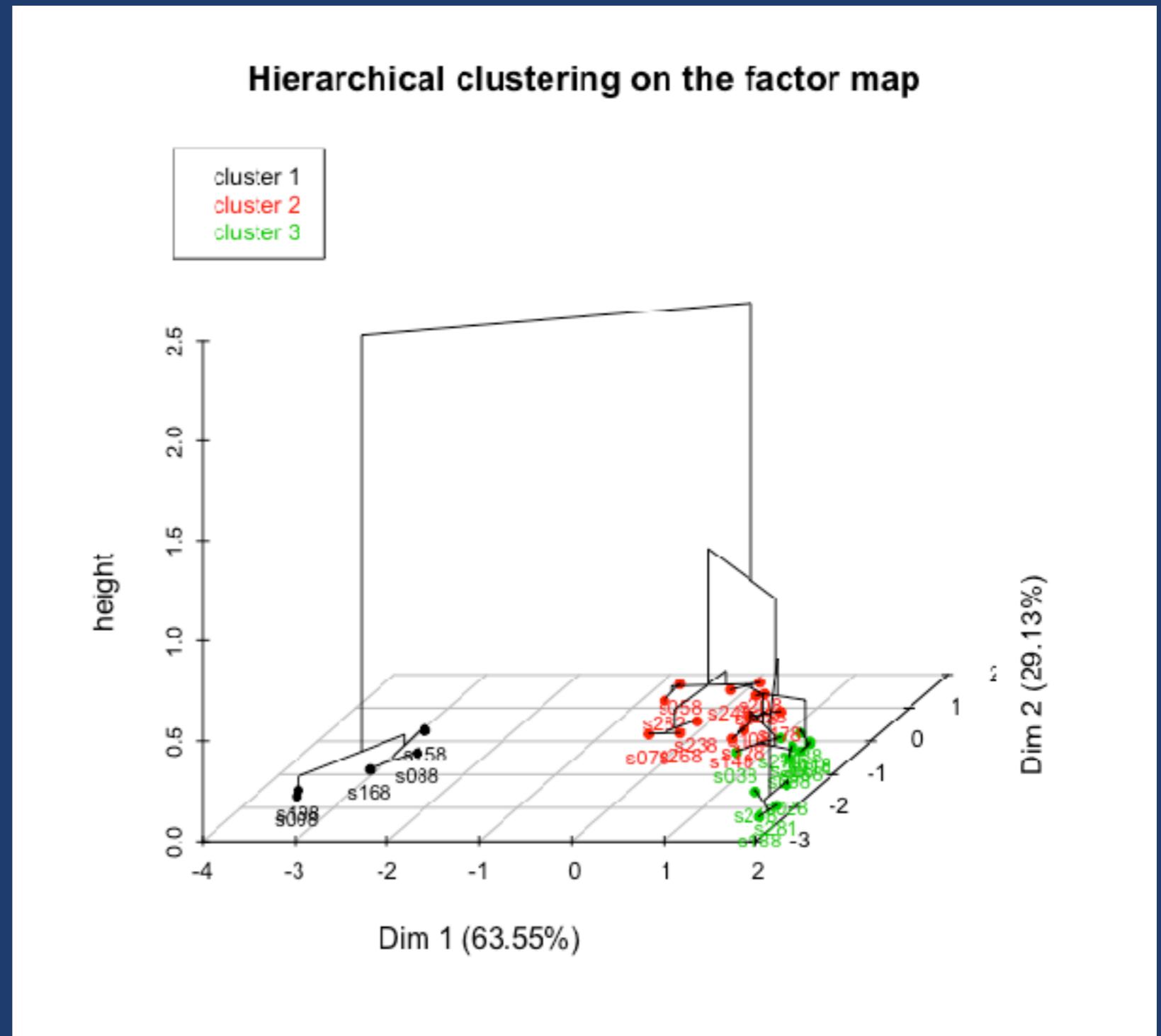
gr7 の刺激文の HC x PCA

- 3つのクラスターの階層化
 - 黒 acceptables
 - 赤 questionables
 - 緑 unacceptables



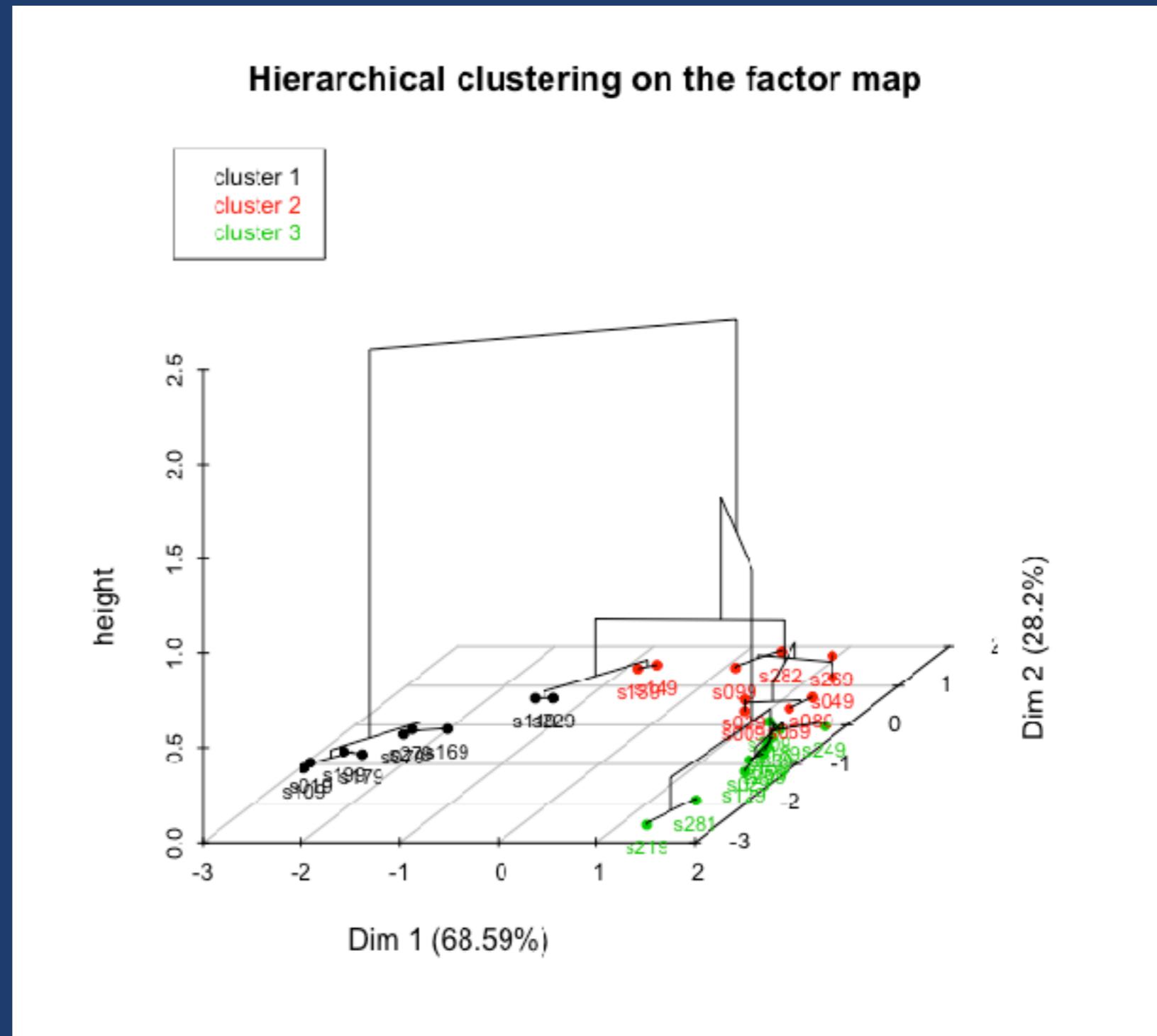
gr8 の刺激文の HC x PCA

- 3つのクラスターの階層化
 - 黒 acceptables
 - 赤 questionables
 - 緑 unacceptables



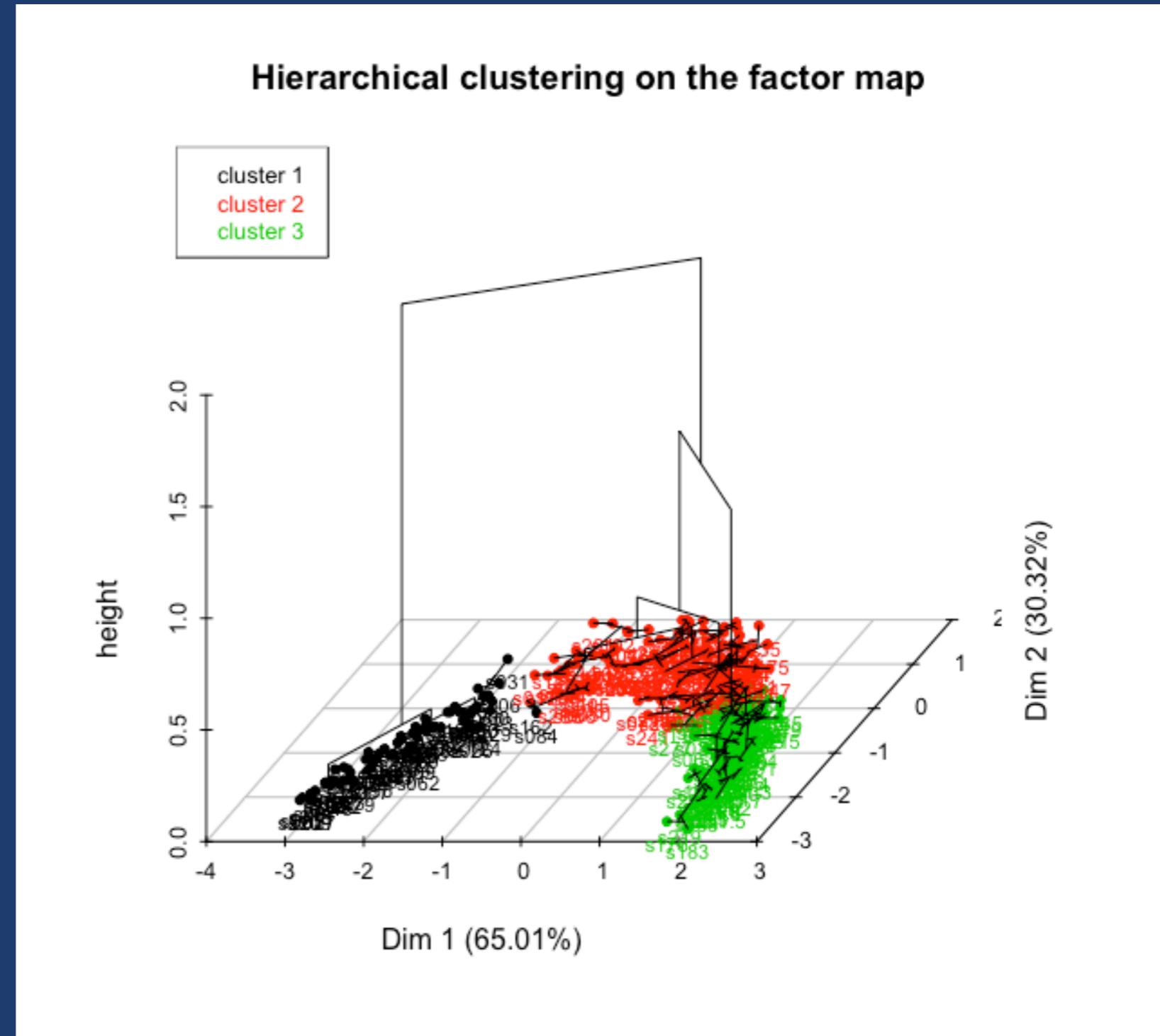
gr9 の刺激文の HC x PCA

- 3つのクラスターの階層化
 - 黒 acceptables
 - 赤 questionables
 - 緑 unacceptables



gr0-gr9 の刺激文の HC x PCA

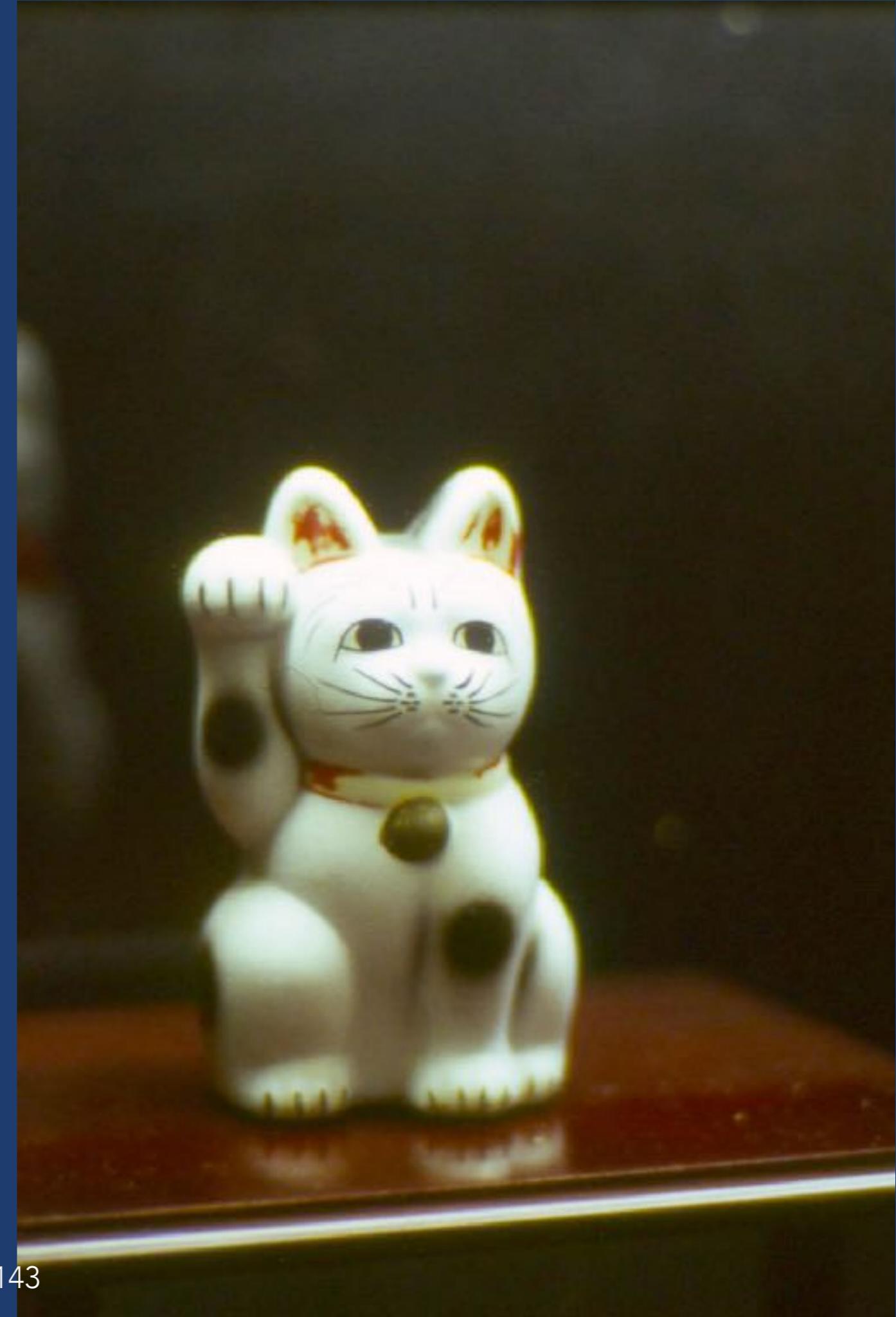
- 大きな対比は
 - 黒 acceptables
 - {赤 questionables, 緑 unacceptables}
- の間にあり, Dim 1 で表現されている
- 疑問
 - Dim 2 < 0 で 2 < Dim 1 < 1 の領域が空白なのはなぜか？



結果

- (1) 大きな対比は
 - 黒 acceptables vs {赤 questionables, 緑 unacceptables}
 - questionables は unacceptables に近い
- 次の問題
 - この配置から容認性判断が正確に何であるかが言えるか？

容認性判断をモ
デル化する

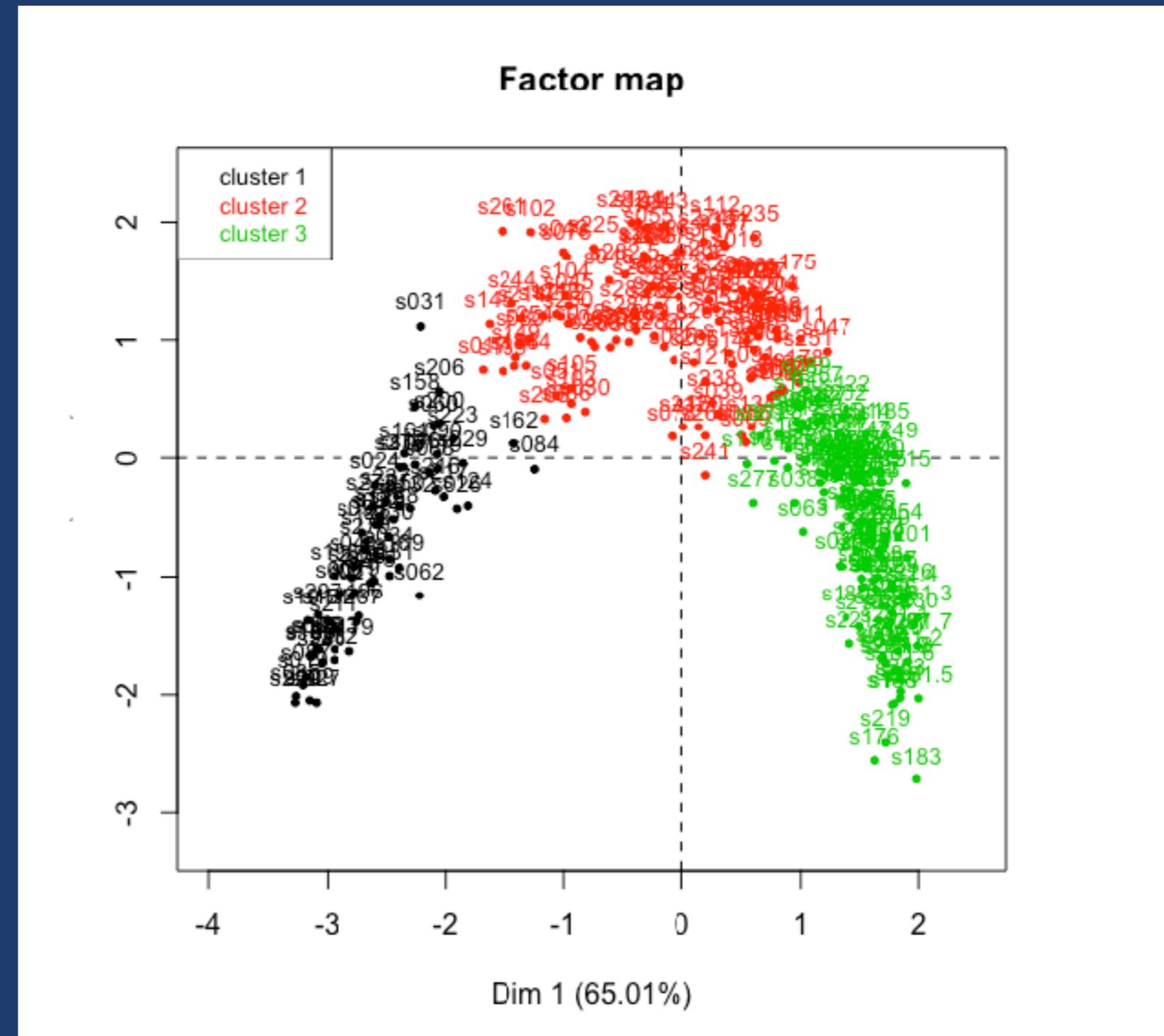


課題の定式化 1/2

- ・次の目標は容認性判断のモデル化
- ・具体的に言うと
 - ・説明変数: `prob[0, 1], prob[1, 2], prob[2, 3], prob[3, ∞)`
 - ・を使って
 - ・目的変数: `A(acceptable), UNA(acceptable)`
 - ・を記述するモデルを構築する事

課題の定式化 2/2

- 課題は、282事例を
 - 事例クラスター1（黒色の）を A(acceptables) か
 - 事例クラスター3（緑色の）UNA(acceptables) か
 - 必要なら、その他 X に
 - 対応させる事
 - 事例クラスター2（赤色）をどう扱うかが鍵



半教師あり分類でモデル化

- 半教師あり分類の手法の一つである **Semi-supervised Local Fisher Discriminat Analysis SELF(A)** (Sugiyama et al. 2010) を使って分類をモデル化
- 教師なし分類だとラベルが認定できないが、教師あり分類だとバイアスがかかる

試した教師条件 1/2

- $P = \{p_{01}, p_{12}, p_{12}, p_{34}\}$ で
 - $p_{02} (= p_{01}+p_{12}) > 0.6$ なら A,
そうでないなら UNA
- 条件 A(acceptability)0
 - $p_{02} (= p_{01}+p_{12}) > 0.5$ なら A,
そうでないなら UNA
- 条件 A(acceptability)0r
 - $p_{02} (= p_{01}+p_{12}) \geq 0.5$ なら A, そうでないなら UNA
- 条件 A(acceptability)1
 - $p_{02} (= p_{01}+p_{12}) > 0.6$ なら A,
そうでないなら UNA
- 条件 A(acceptability)2
 - $p_{02} (= p_{01}+p_{12}) > p_{13} (= p_{12}+p_{13}) > p_{24} (= p_{23}+p_{3\infty})$ なら A, そうでないなら UNA
- 以上は単純な二値分類 = カテゴリカルな判断 (categorical judgment) のモデル化

試した教師条件 2/2

- 条件 A(acceptability)3

- $p_{02} > 0.5$ なら A, $p_{24} > 0.5$ なら UNA, 残り X

- 条件 A(acceptability)5

- p_{02}, p_{13}, p_{24} の最大値が p_{02} なら A, p_{24} なら UNA, 残り X

- 条件 A(acceptability)4

- $p_{02} > 0.6$ なら A, $p_{24} > 0.6$ なら UNA, 残り X

- 条件 A(acceptability)6

- p_{01} が最大なら A, p_{34} が最大なら UNA, 残り X

- 条件 A(acceptability)4r

- $p_{02} > 0.7$ なら A, $p_{24} > 0.7$ なら UNA, 残り X

- 条件 A(acceptability)6r

- p_{01} が最大なら A, p_{34} が最大なら UNA, その他の時, $p_{12} < p_{23}$ なら X, そうでないなら Y

- 条件 A(acceptability)4r2

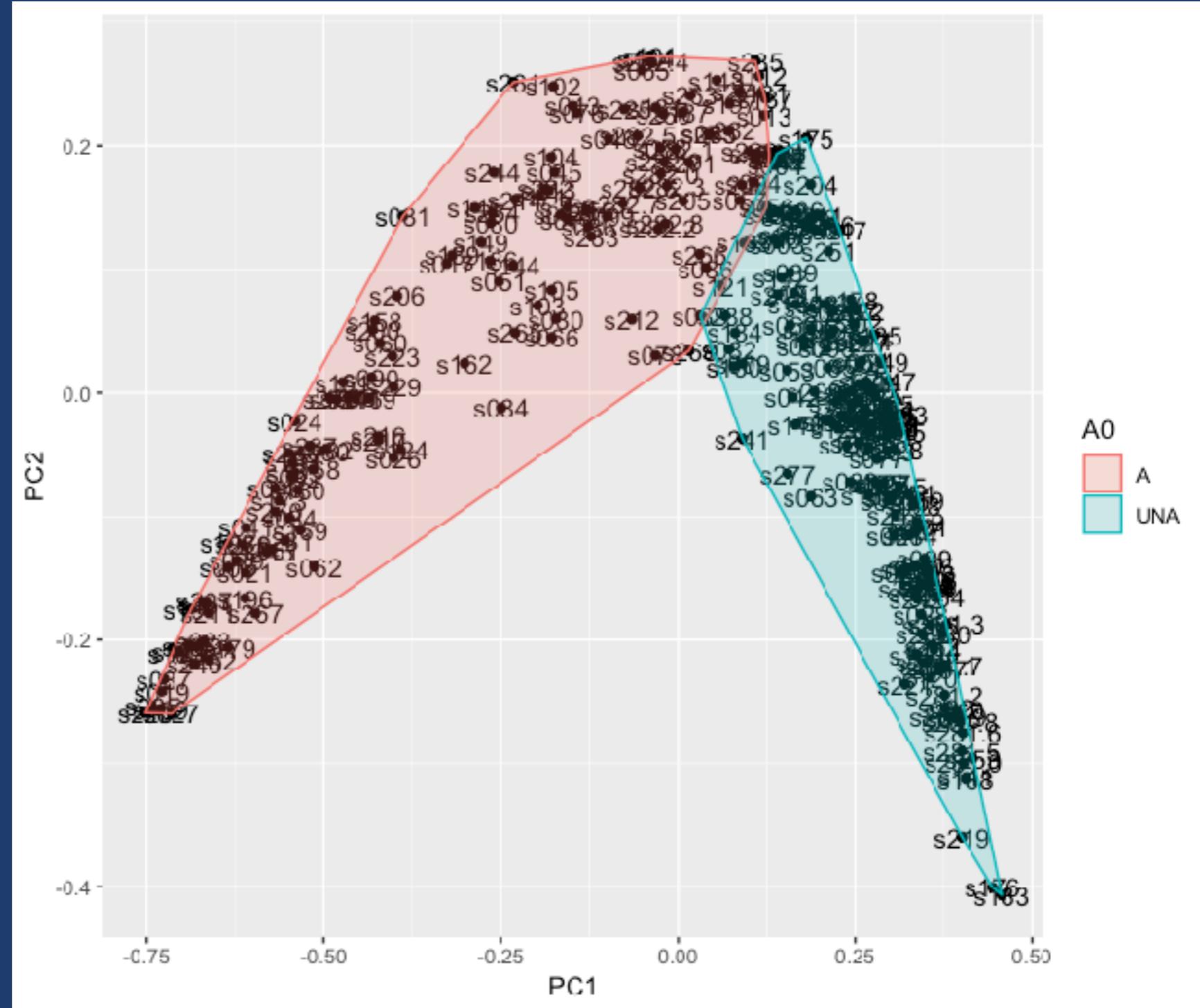
- $p_{02} > 0.8$ なら A, $p_{24} > 0.6$ なら UNA, 残り X

- 以上は中間段階ありの評価のモデル化

SELFA ($r=3$)

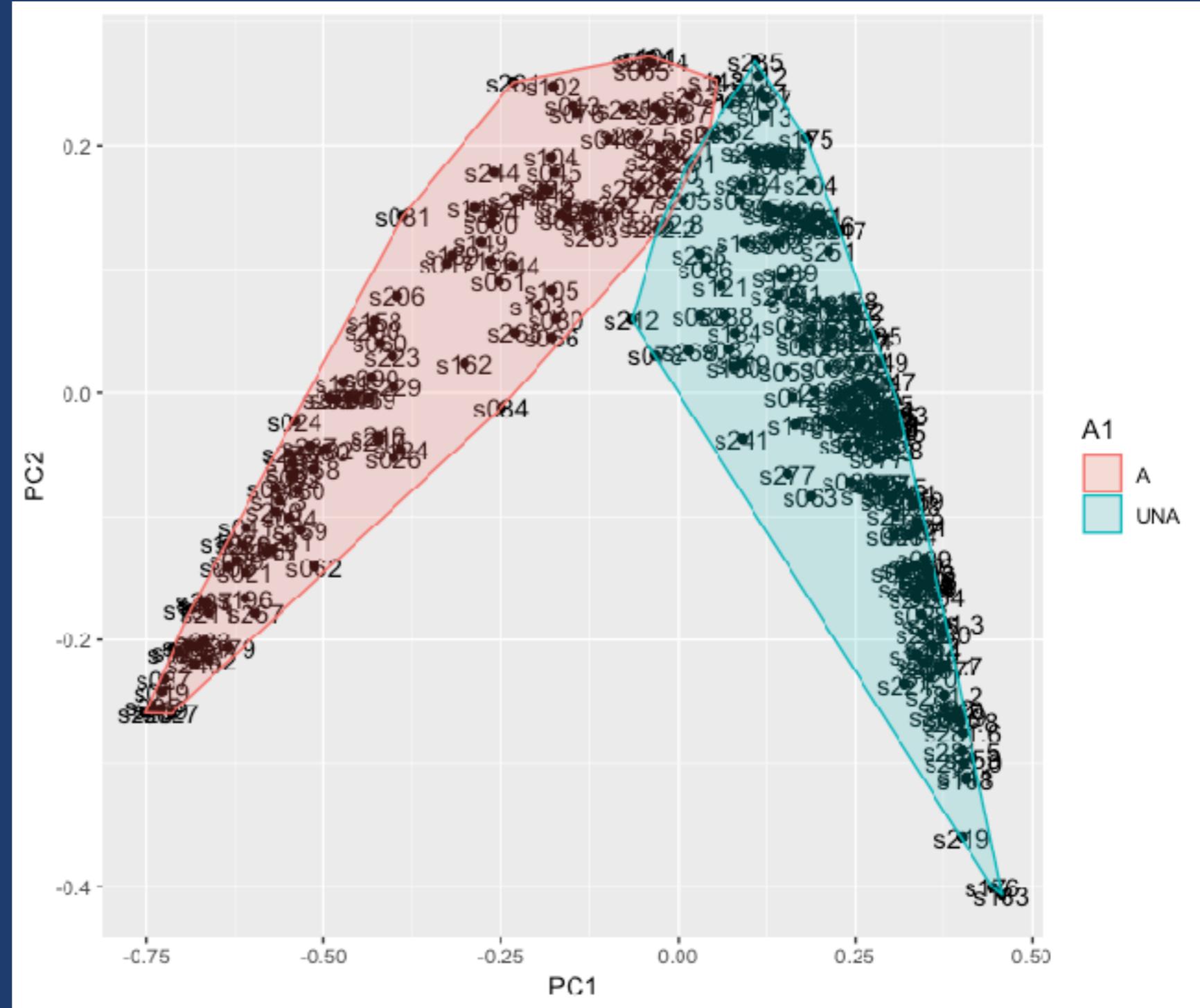
A0-A0

- A/UNA の2値分類
 - if $p_{01} > 0.5$: A;
 - else: UNA



SELFA ($r=3$) A0-A1

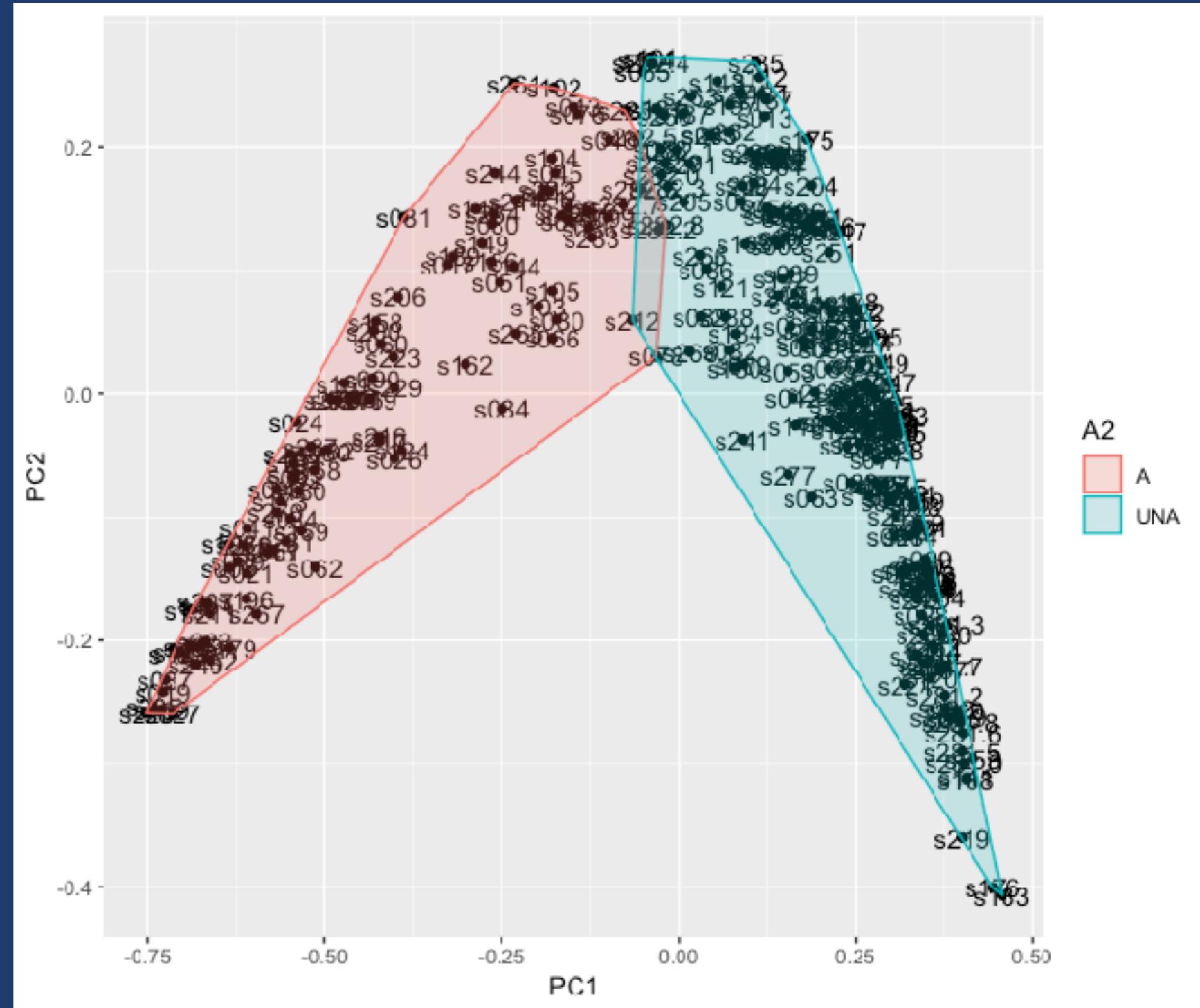
- A/UNA の2値分類
 - if $p_{01} > 0.6$: A;
 - else: UNA



SELFA ($r=3$)

A0-A2

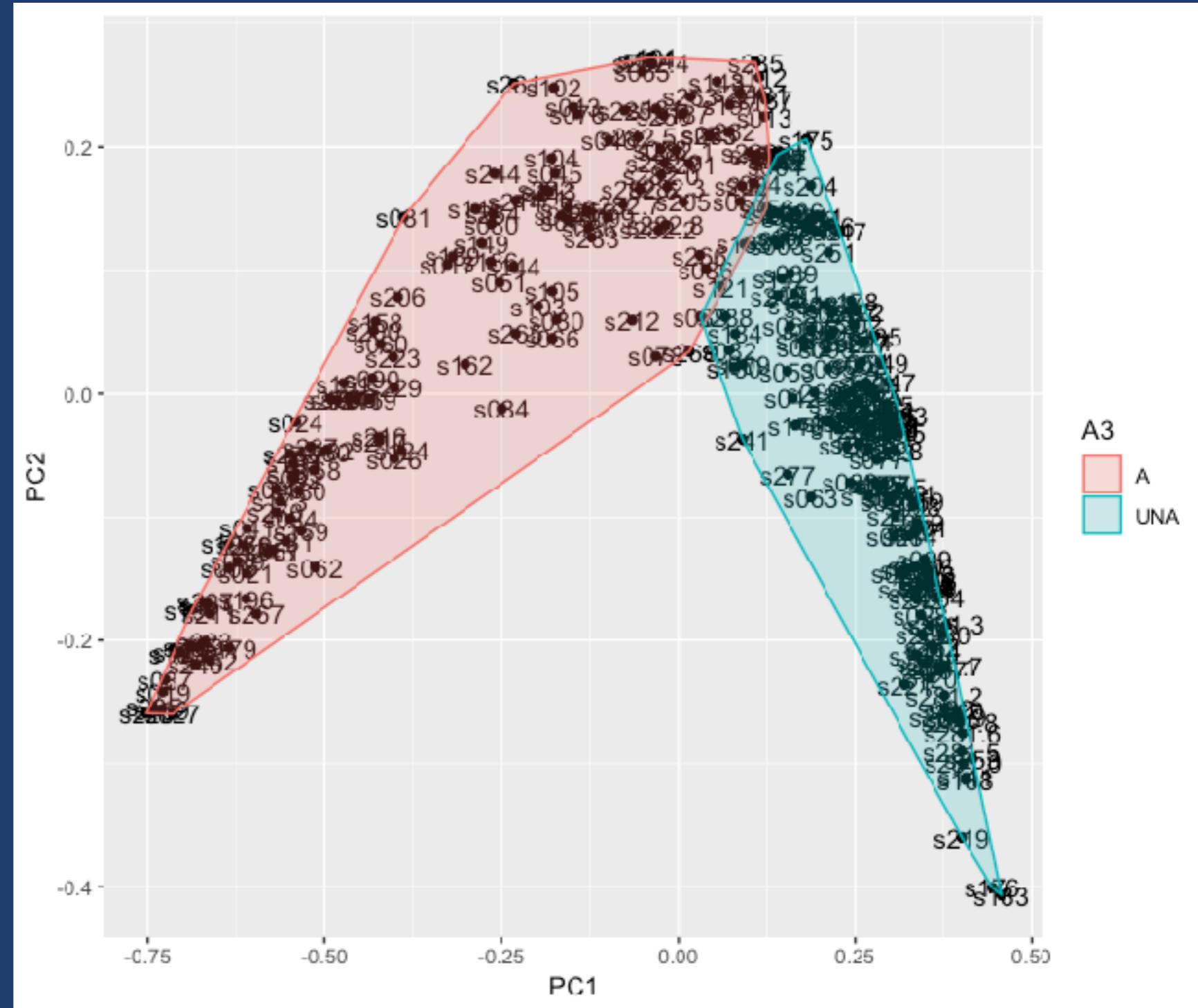
- A/UNA の2値分類
 - if $p_{01} + p_{12} > p_{12}$
+ $p_{12} > p_{12} + p_{34}$:
A;
 - else: UNA



SELFA ($r=3$)

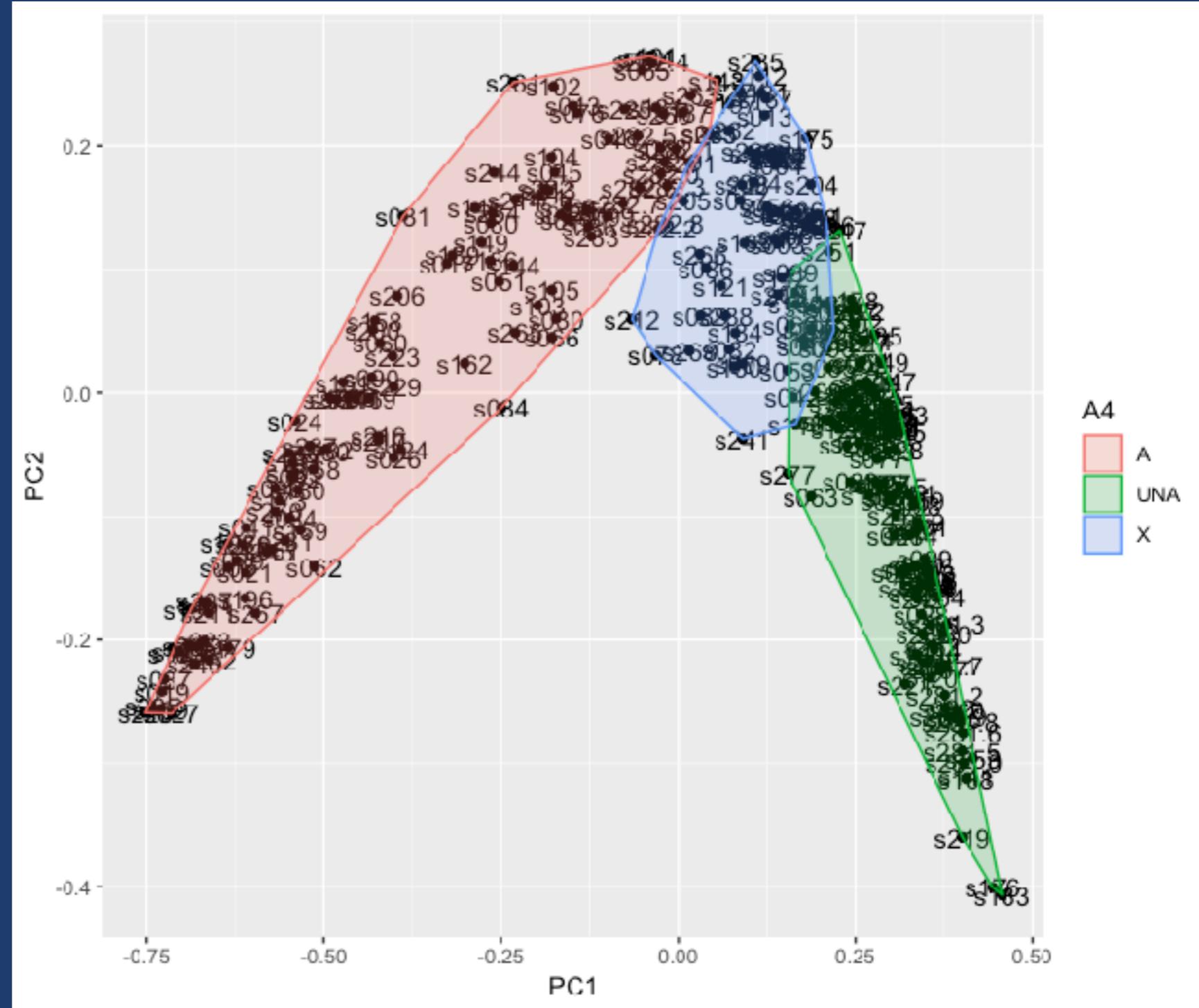
A0-A3

- A/UNA の3値分類
- if $p_{01} + p_{12} > 0.5$: A;
- else if: $p_{12} + p_{34} > 0.5$: UNA
- else: X
- 注意
 - 2値分類の A0 と変わらない



SELFA (r=3) A0-A4

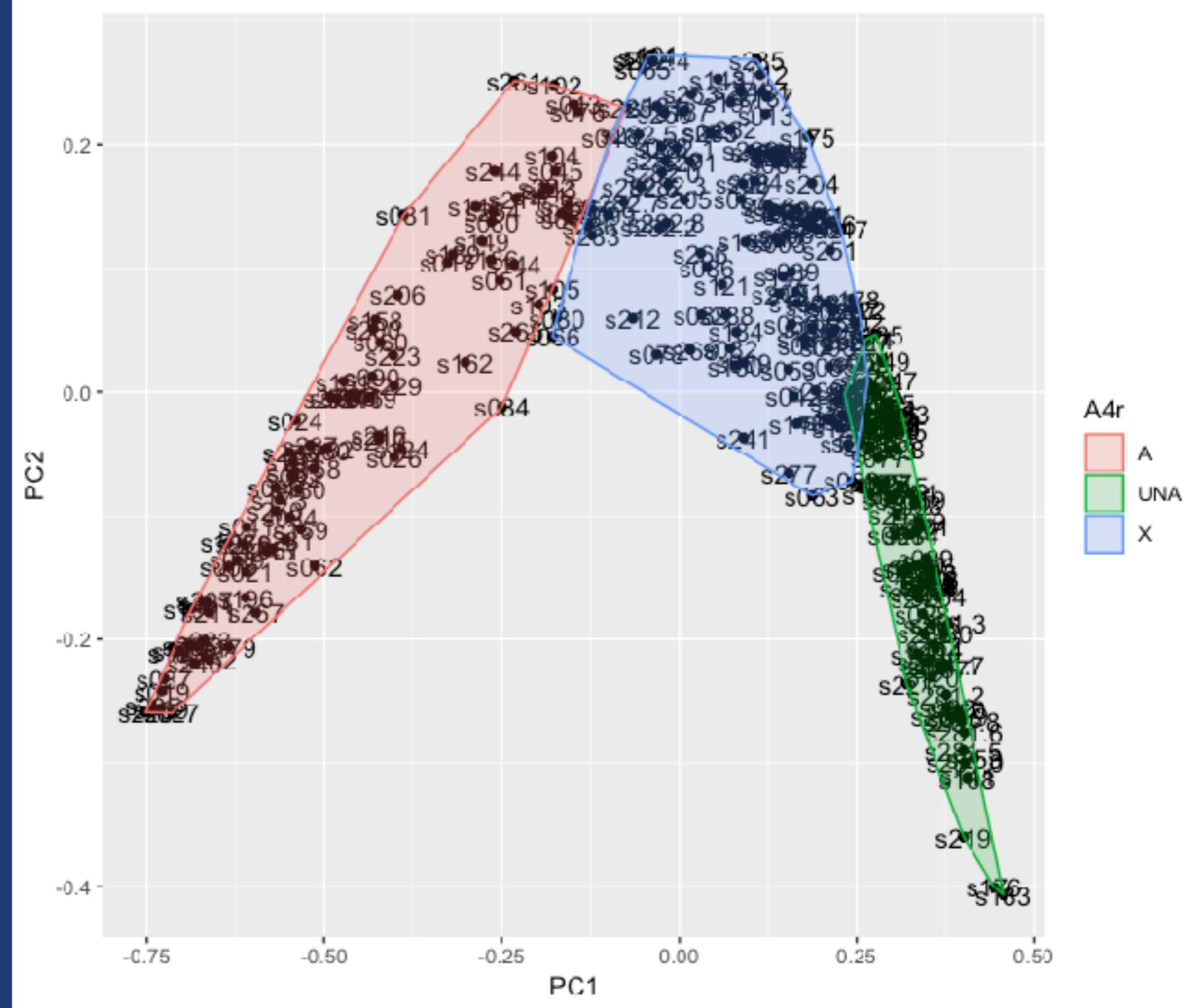
- A/X/UNA の3値分類
- if $p_{01} + p_{12} > 0.6$: A;
- else if: $p_{12} + p_{34} > 0.6$: UNA
- else: X



SELFA (r=3)

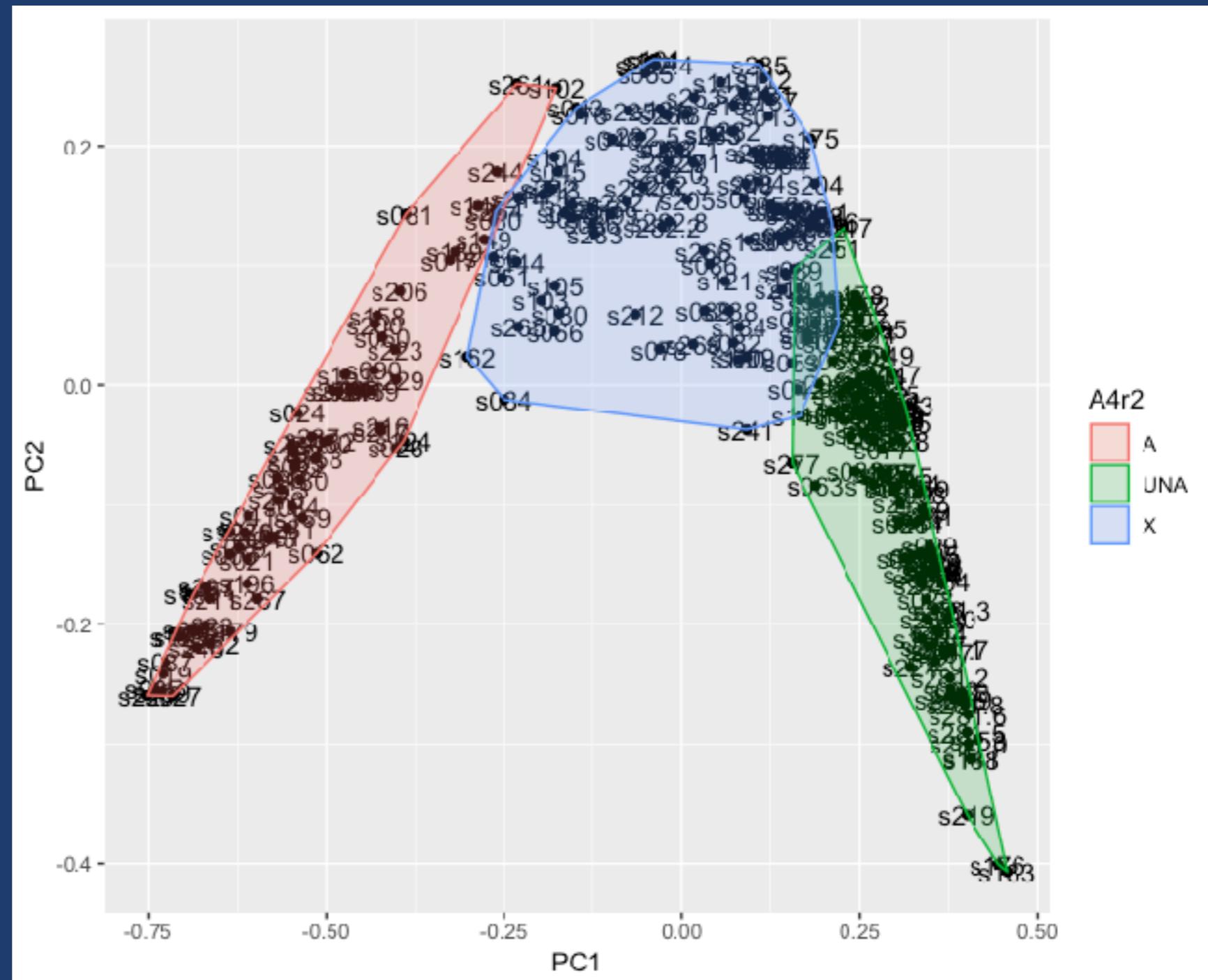
A0-A4r

- A/X/UNA の3値分類
 - if $p_{01} + p_{12} > 0.7$: A;
 - else if: $p_{12} + p_{34} > 0.7$: UNA
 - else: X



SELFA (r=3) A0-A4r2

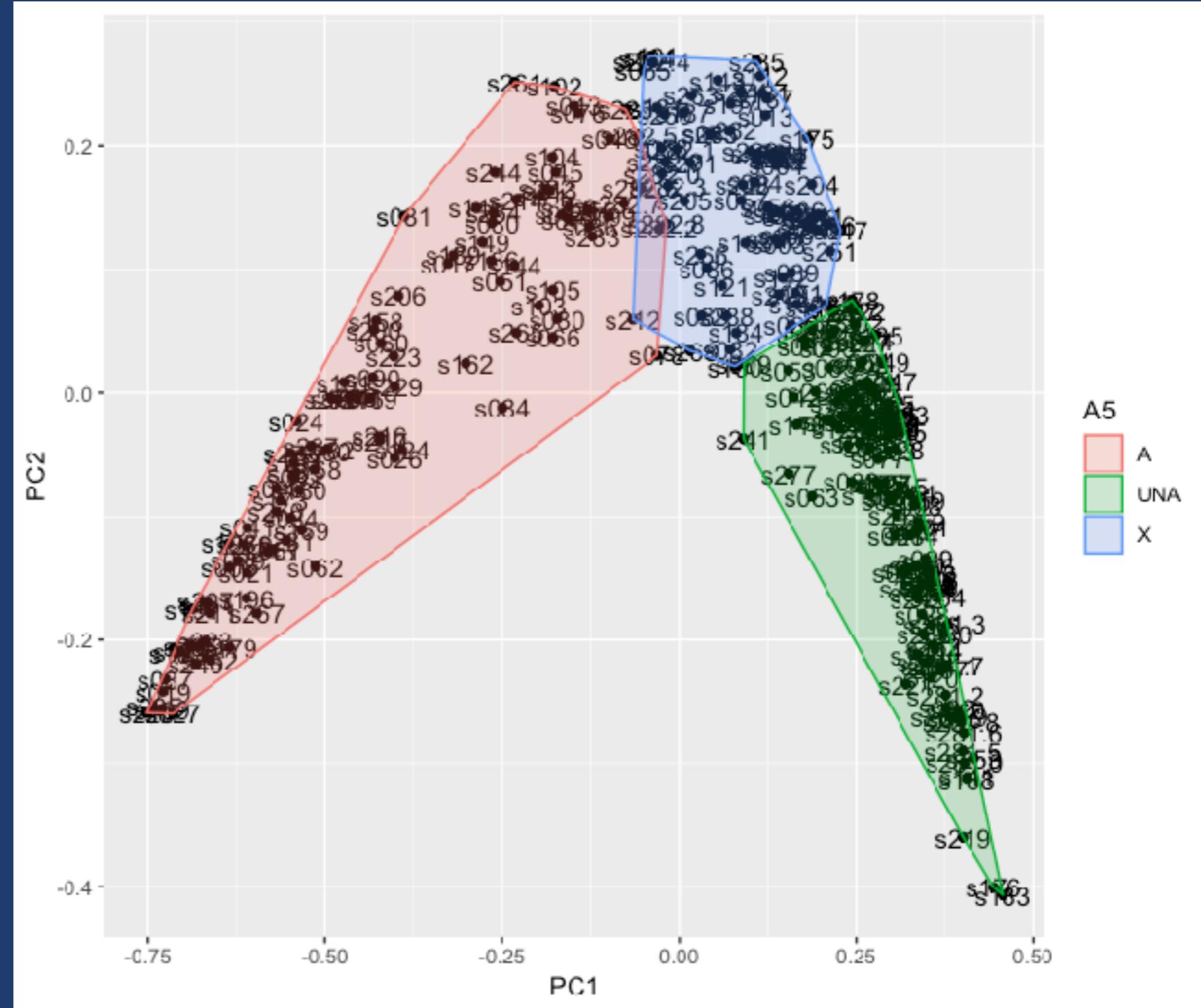
- A/X/UNA の3値分類
- if $p_{01} + p_{12} > 0.8$: A;
- else if: $p_{12} + p_{34} > 0.6$: UNA
- else: X



SELFA ($r=3$)

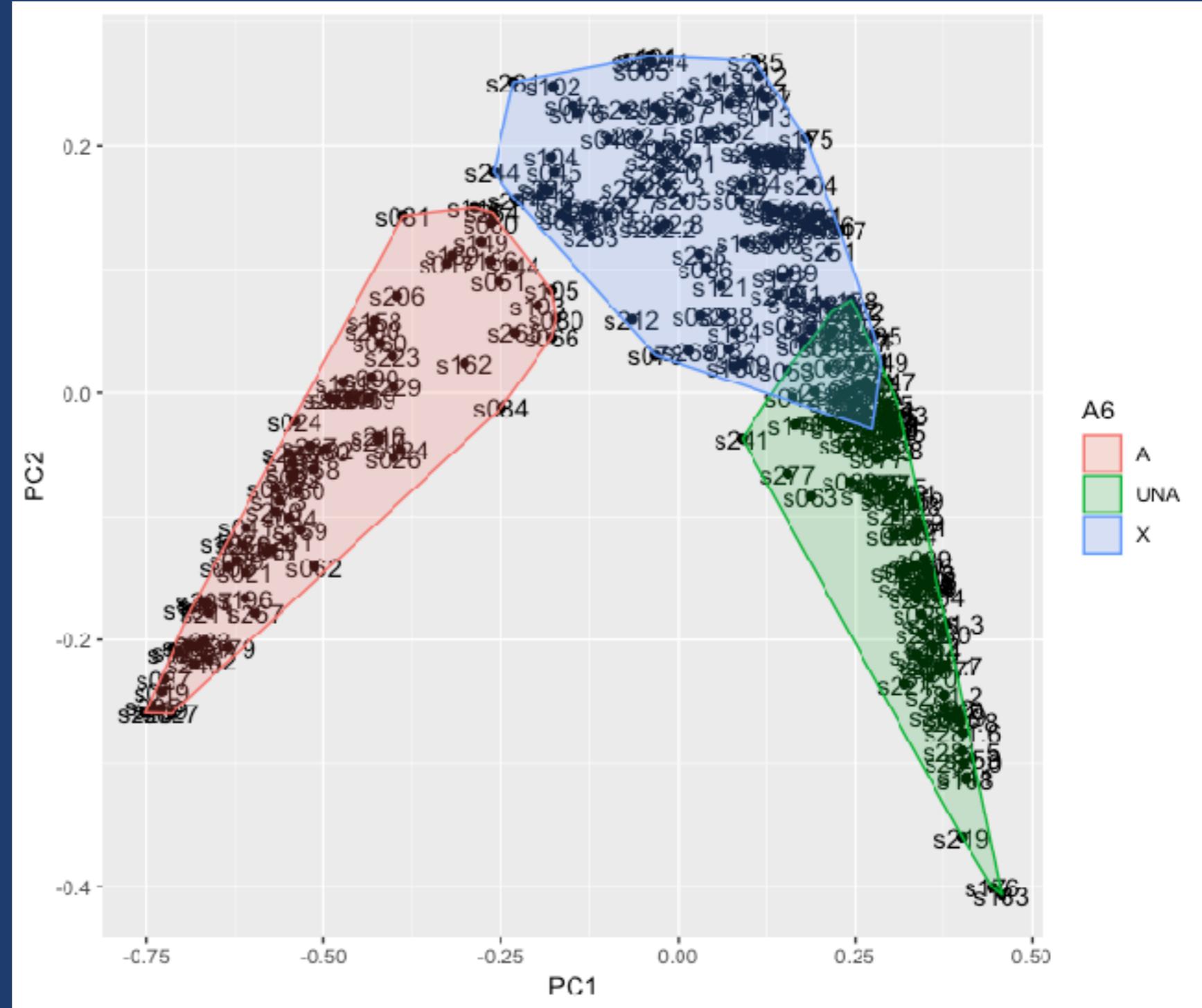
A0-A5

- A/X/UNA の3値分類
 - if $p_{02} == \max$: A;
 - else if $p_{12} == \max$: UNA
 - else: X



SELFA (r=3) A0-A6

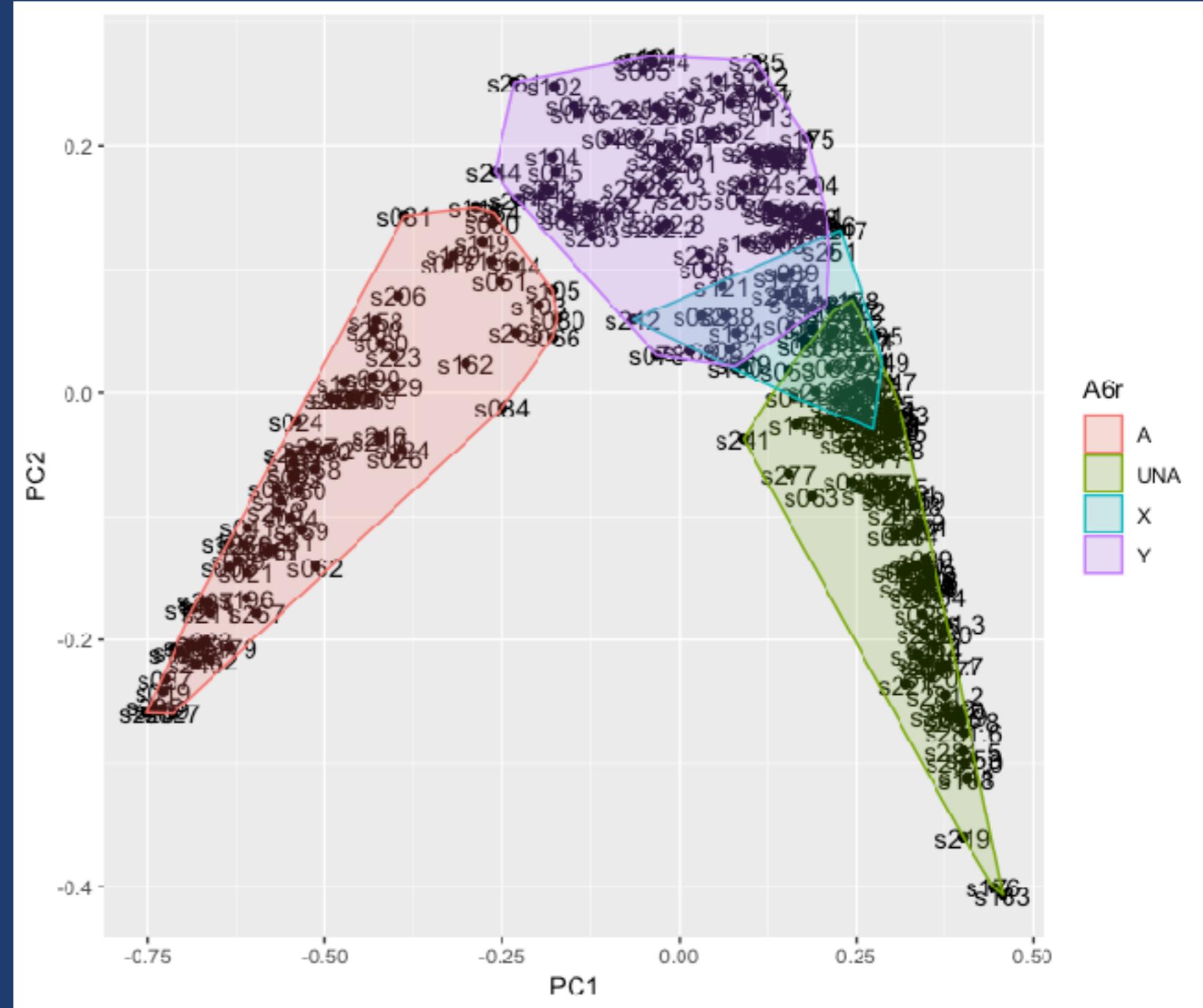
- A/X/UNA の3値分類
 - if $p_{01} == \max$: A;
 - else if: $p_{34} == \max$: UNA;
 - else X



SELFA ($r=3$)

A0-A6r

- A/X/UNA の4値分類
 - if $p_{01} == \max$: A;
 - else if: $p_{34} == \max$: UNA;
 - else
 - if $p_{12} < p_{12}$: X
 - else: Y



評価

- 知見 1
 - A0, A0r, A1, A2, A3 までは cluster 2 (赤色)が認識されないので, 不適
- 知見 2
 - A4, A5 は cluster 2 (赤色)に相当する領域が小さすぎるので, 不適
- 知見 3
 - A4r, A4r2, A6 (=A6r) は3つのクラスターに程々に対応している
- 知見 4
 - cluster 1 と cluster 2 の境界の形を見ると, A6 (=A6r)の方が良い対応 (帰属を100%説明)
 - A6r なら gr2 の PCA でクラスターが4つに分れた事も記述可能

まとめと今後の展開



結論

- 結論
 - 応答ポテンシャル $p[0,1], p[1,2], p[2,3], p[3,\infty)$ を使った抽象化) を使えば、ヒトの示す容認度評定の多様性をそれなりにうまく表現できる
- 提案
 - 基準A6 (か A6r) を使えば容認度評定がそれなりの精度で表わせるのでは？
 - 容認性判断は容認度評定の下位クラスなので、これで容認性判断のモデル化に成功した事になるのでは？

今後の課題

- 分析の点で
 - 第一部のデータと比較する
 - 実施済み
 - 層別解析を実施して個々の属性の影響を見る
 - 実は実施済み
 - 個々の文で容認度の決めている、層別化できない本質要因を特定
 - データ構築の点で
 - 変異の数を増やす
 - 時制や時相の変異を加える
 - 現状は過去形のみ
 - 動詞ごとの変異を増やす
 - 一次元化の仮定をはずして評定課題を実施する
 - 0, 1, 2, 3 は一次元化
 - 反応時間データを取って、それと対応させる

References

- Daniel Adler, Duncan Murdoch et al (2018). *rgl: 3D visualization using OpenGL.* R package version 0.99.16. <https://CRAN.R-project.org/package=rgl>
- Sebastien Le, Julie Josse, Francois Husson (2008). *FactoMineR: An R package for multivariate analysis.* *J. of Statistical Software*, 25(1), 1-18. [10.18637/jss.v025.i01](https://doi.org/10.18637/jss.v025.i01)
- Masaru Sugiyama, Tsuyoshi Ide, Shin'ichi Nakajima, and Jun Sese (2010). *Semi-supervised local Fisher discriminant analysis for dimensionality reduction.* *Machine Learning* 78(1-2): 35–61.
- Yuan Tang, and Wenxuan Li (2016). *lfda: An R Package for Local Fisher Discriminant Analysis and Visualization.* [arXiv:1612.09219](https://arxiv.org/abs/1612.09219)
- January Weiner (2017). *pca3d: Three dimensional PCA plots.* R package version 0.10. <https://CRAN.R-project.org/package=pca3d>

Thank you for your
attention

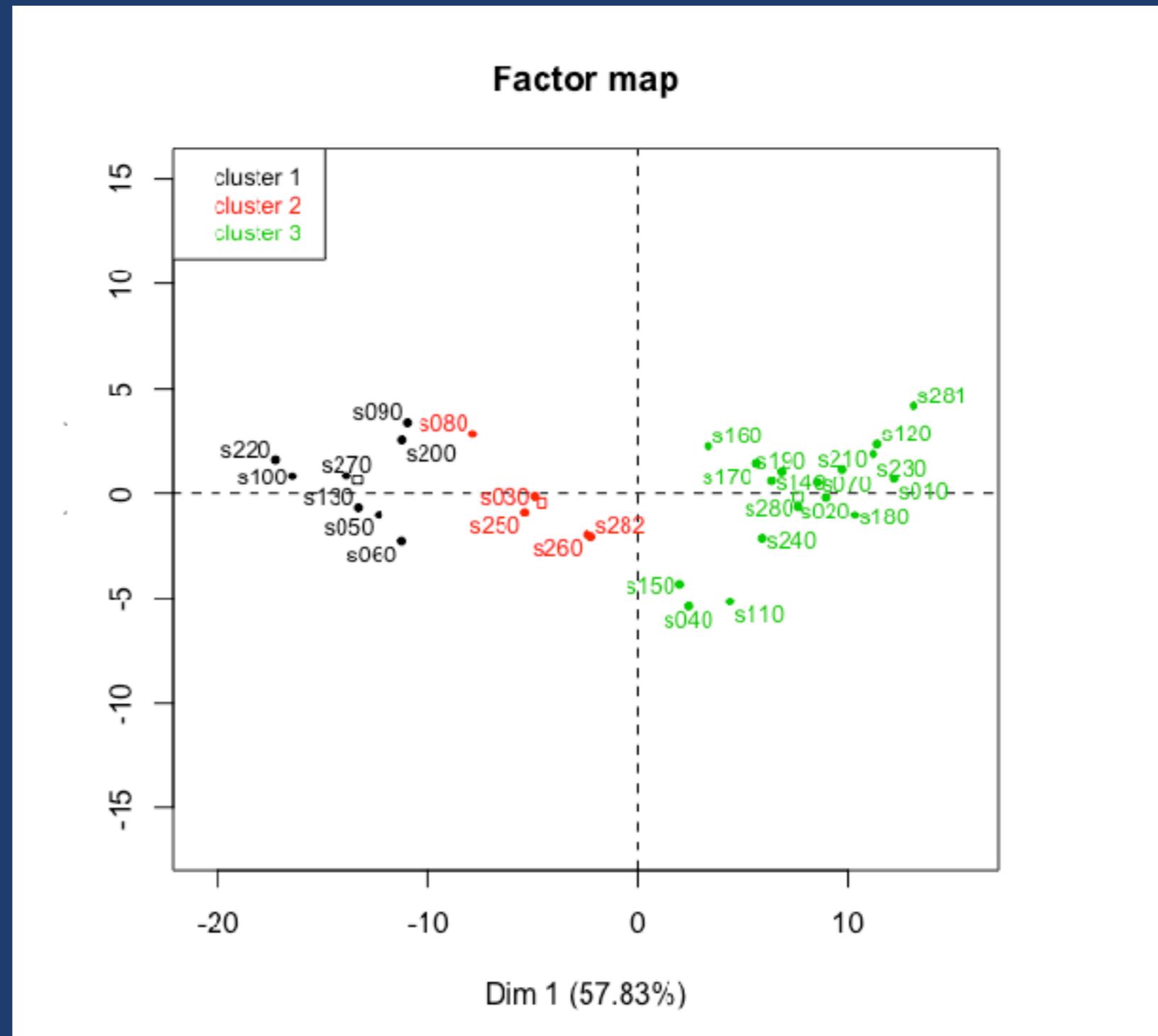
Appendix 1

参考データ：
normalization
なしで分析すると



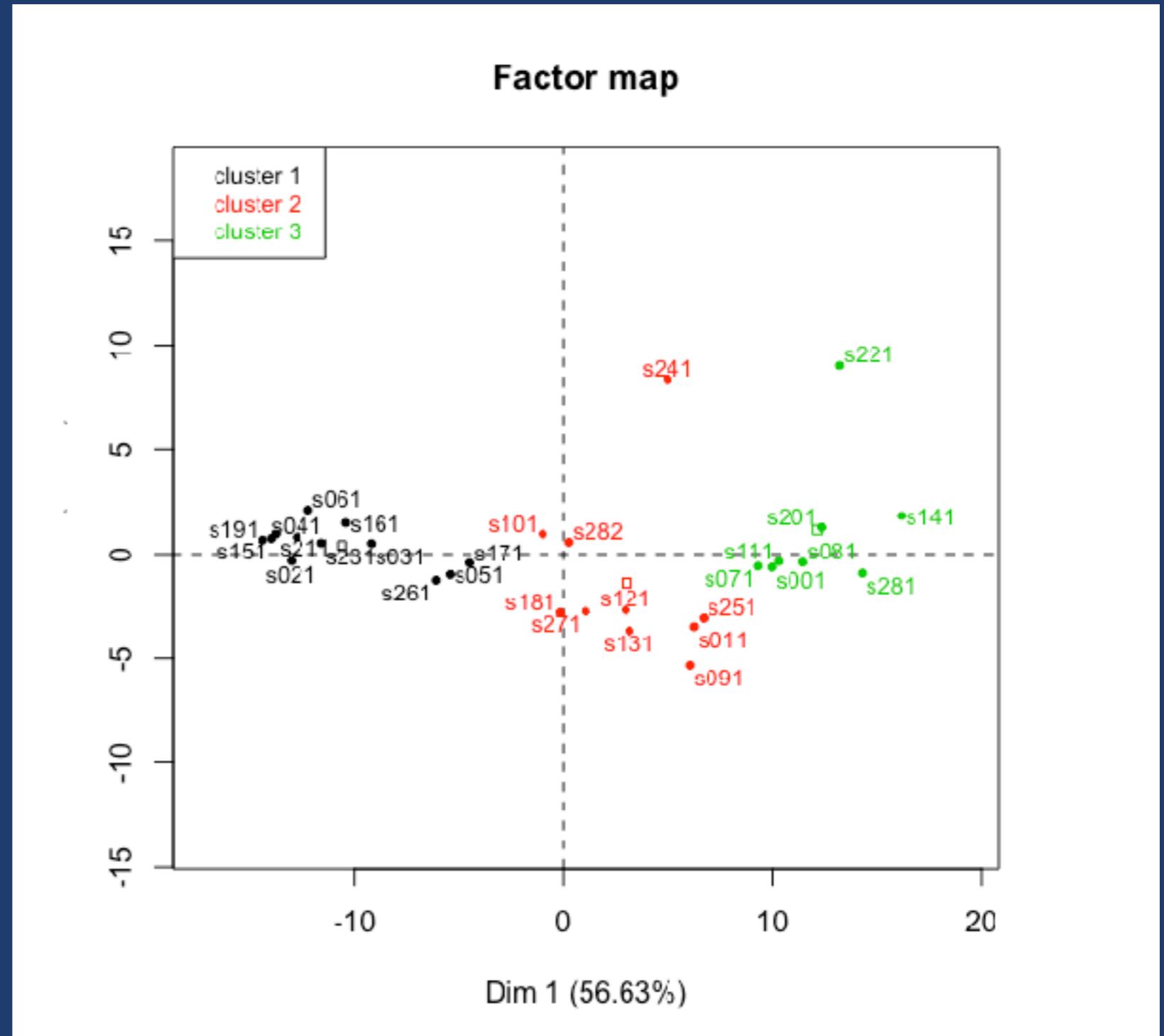
gr0 の刺激文の標準化なしの PCA (Factor Map)

- 3つのクラスターが形成されている
- 黒 acceptables
- 赤 questionables
- 緑 unacceptables



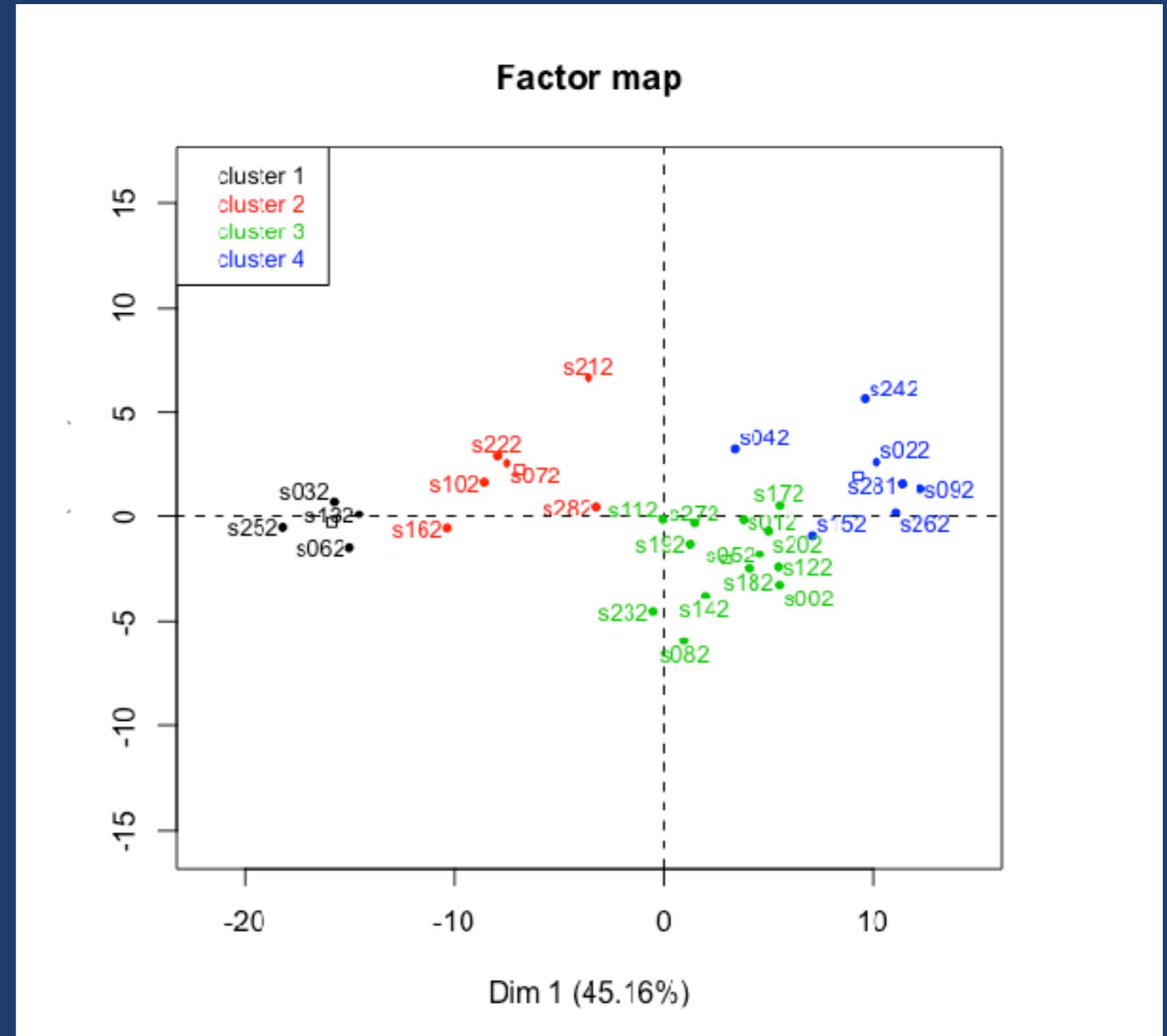
gr1 の刺激文の標準化なしの PCA (Factor Map)

- 3つのクラスターが形成されている
 - 黒 acceptables
 - 赤 questionables
 - 緑 unacceptables



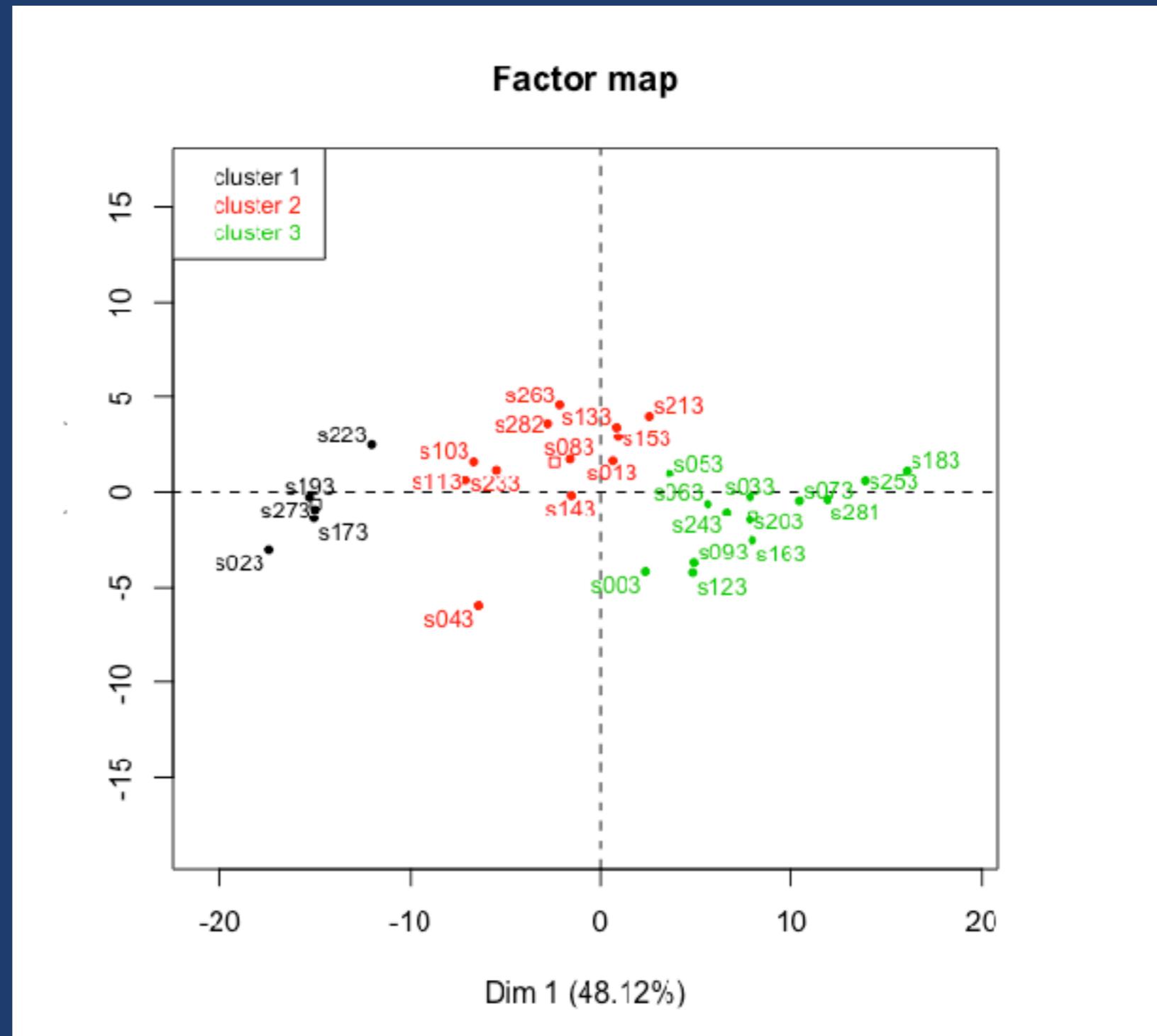
gr2 の刺激文の標準化なしの PCA (Factor Map)

- 4つのクラスターが形成されている
- 黒 acceptables
- 赤 weakly questionables
- 緑 strongly questionables
- 青 unacceptables



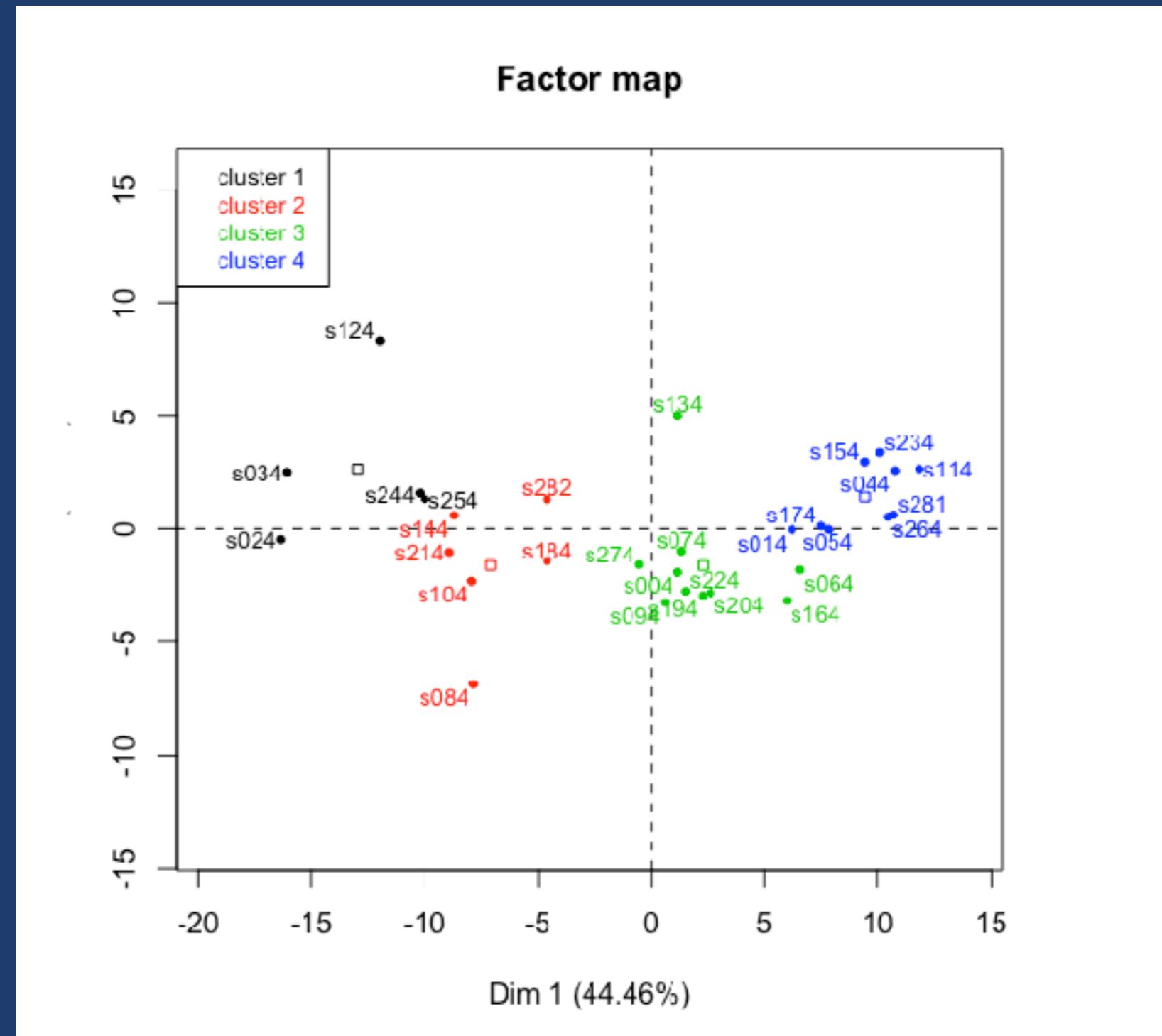
gr3 の刺激文の標準化なしの PCA (Factor Map)

- 3つのクラスターが形成されている
- 黒 acceptables
- 赤 questionables
- 緑 unacceptables



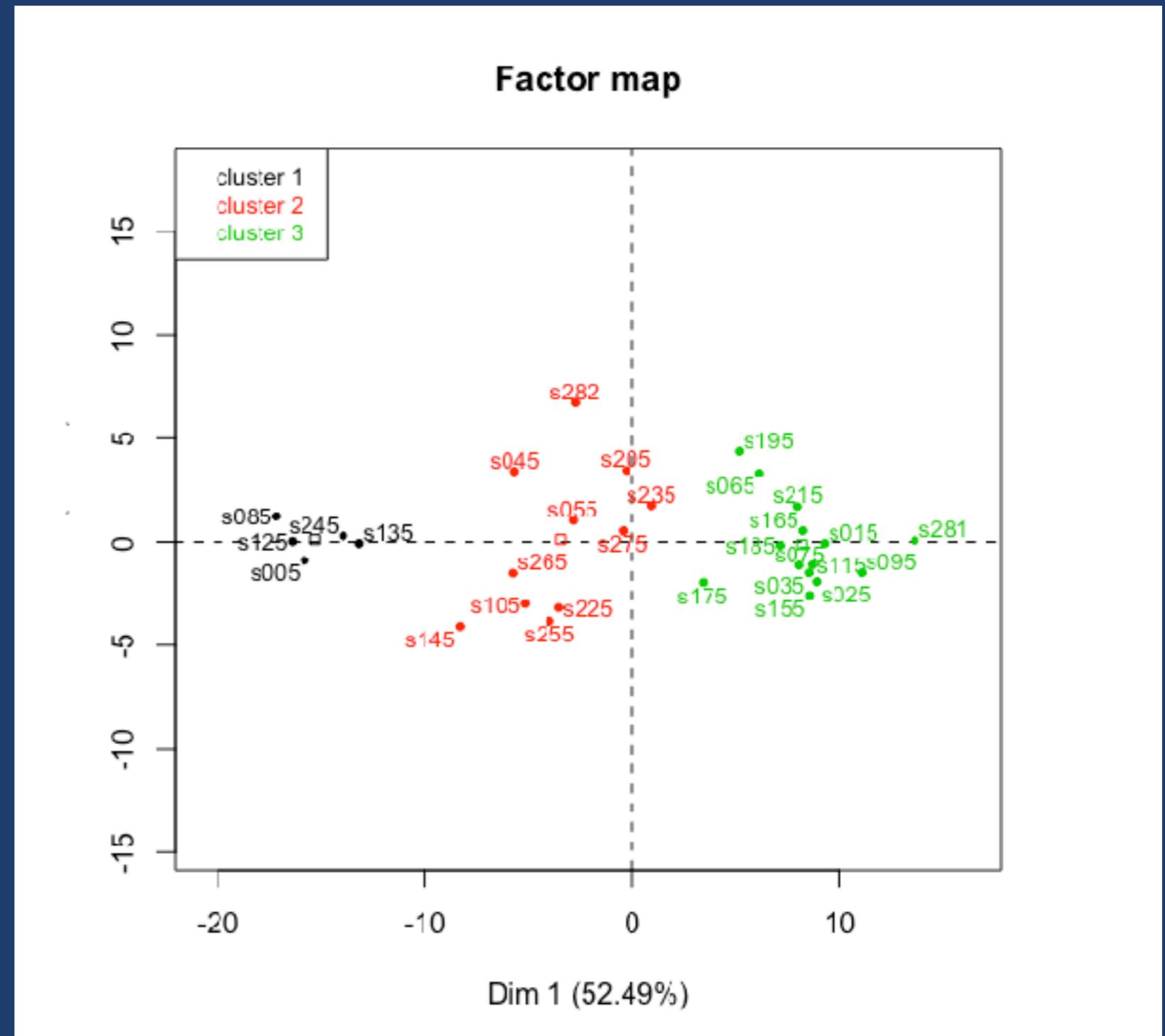
gr4 の刺激文の標準化なしの PCA (Factor Map)

- 3つのクラスターが形成されている
- 黒 acceptables
- 赤 questionables
- 緑 unacceptables



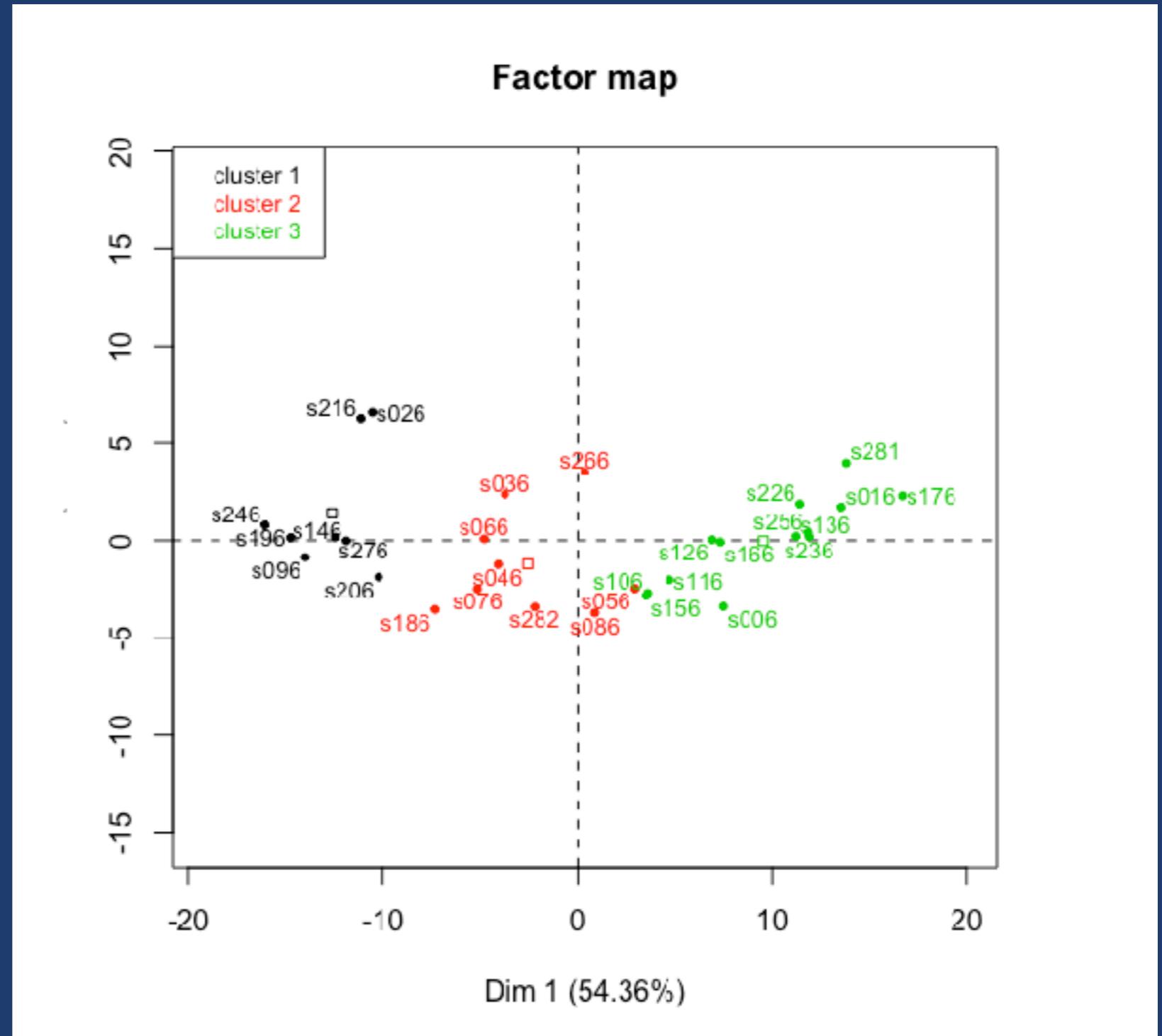
gr5 の刺激文の標準化なしの PCA (Factor Map)

- 3つのクラスターが形成されている
- 黒 acceptables
- 赤 questionables
- 緑 unacceptables



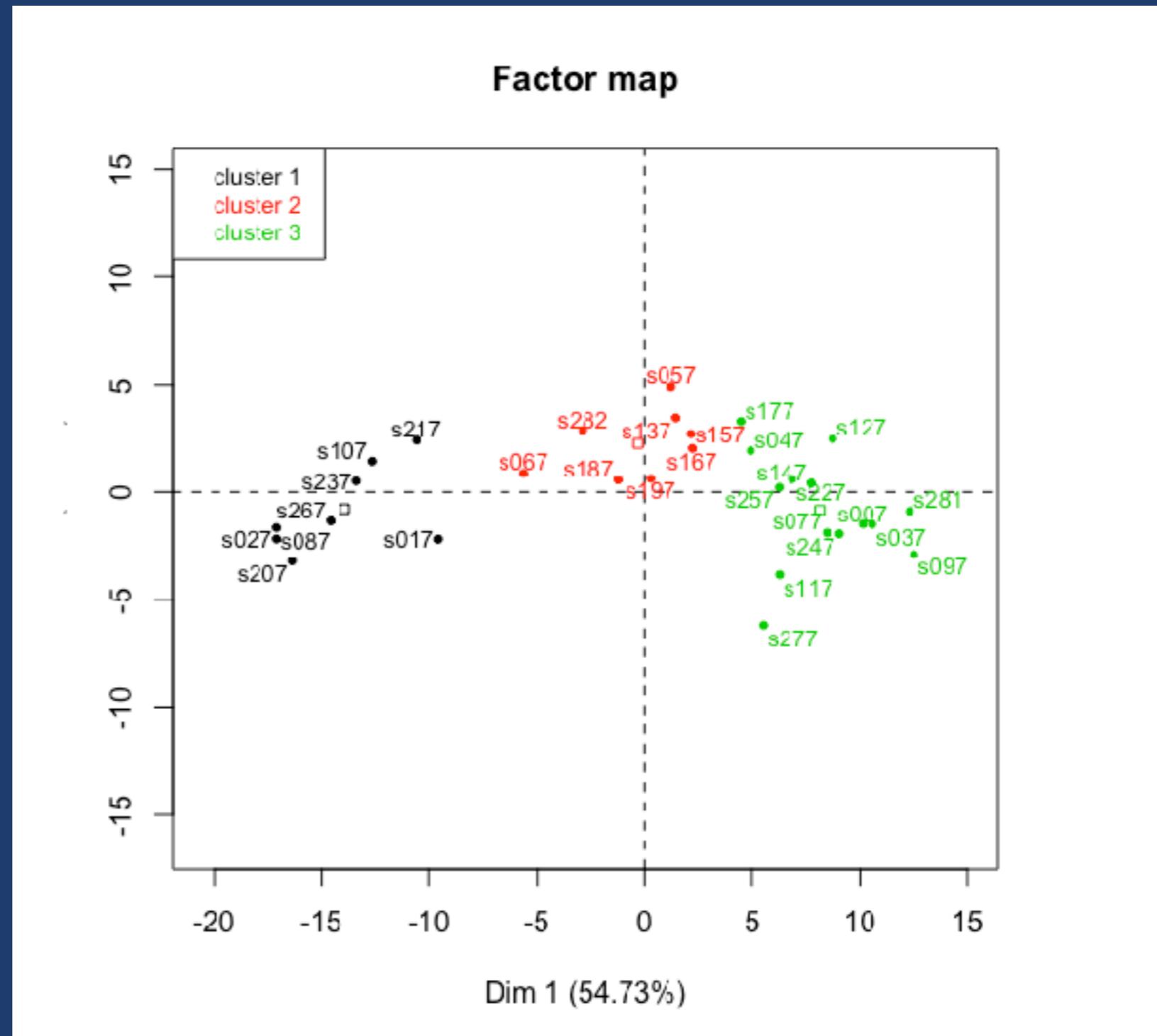
gr6 の刺激文の標準化なしの PCA (Factor Map)

- 3つのクラスターが形成されている
 - 黒 acceptables
 - 赤 questionables
 - 緑 unacceptables



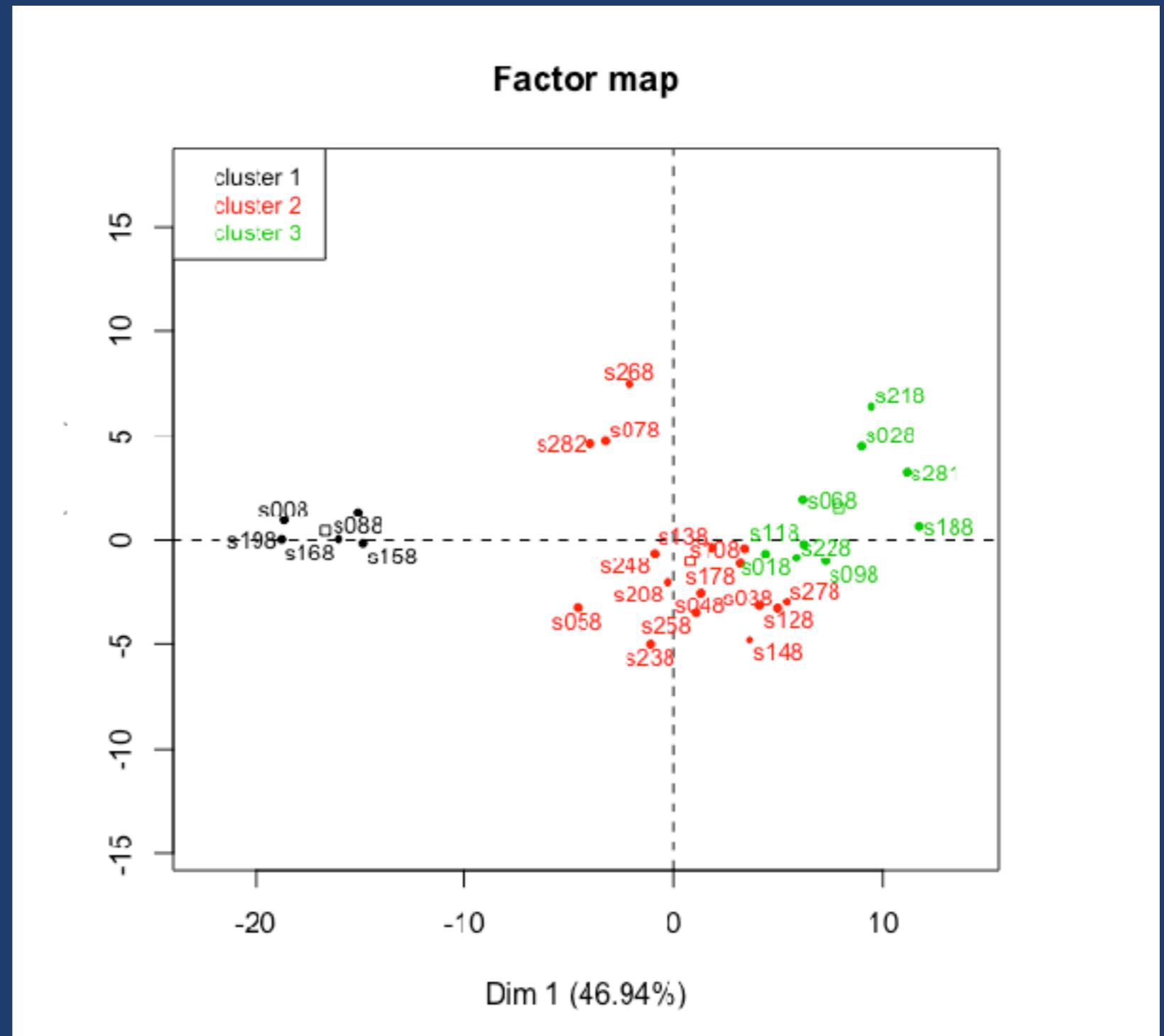
gr7 の刺激文の標準化なしの PCA (Factor Map)

- 3つのクラスターが形成されている
- 黒 acceptables
- 赤 questionables
- 緑 unacceptables



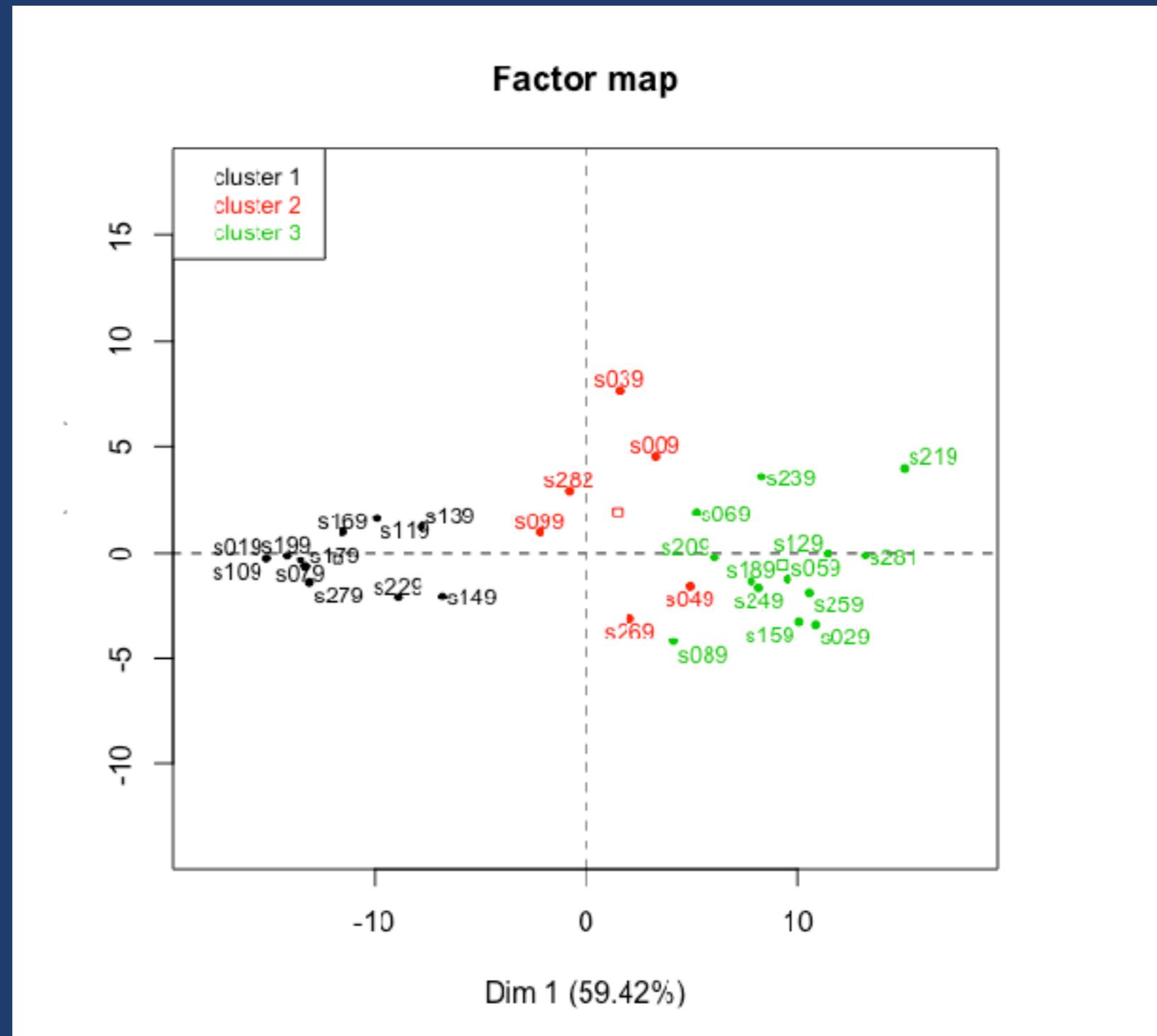
gr8 の刺激文の標準化なしの PCA (Factor Map)

- 3つのクラスターが形成されている
- 黒 acceptables
- 赤 questionables
- 緑 unacceptables



gr9 の刺激文の標準化なしの PCA (Factor Map)

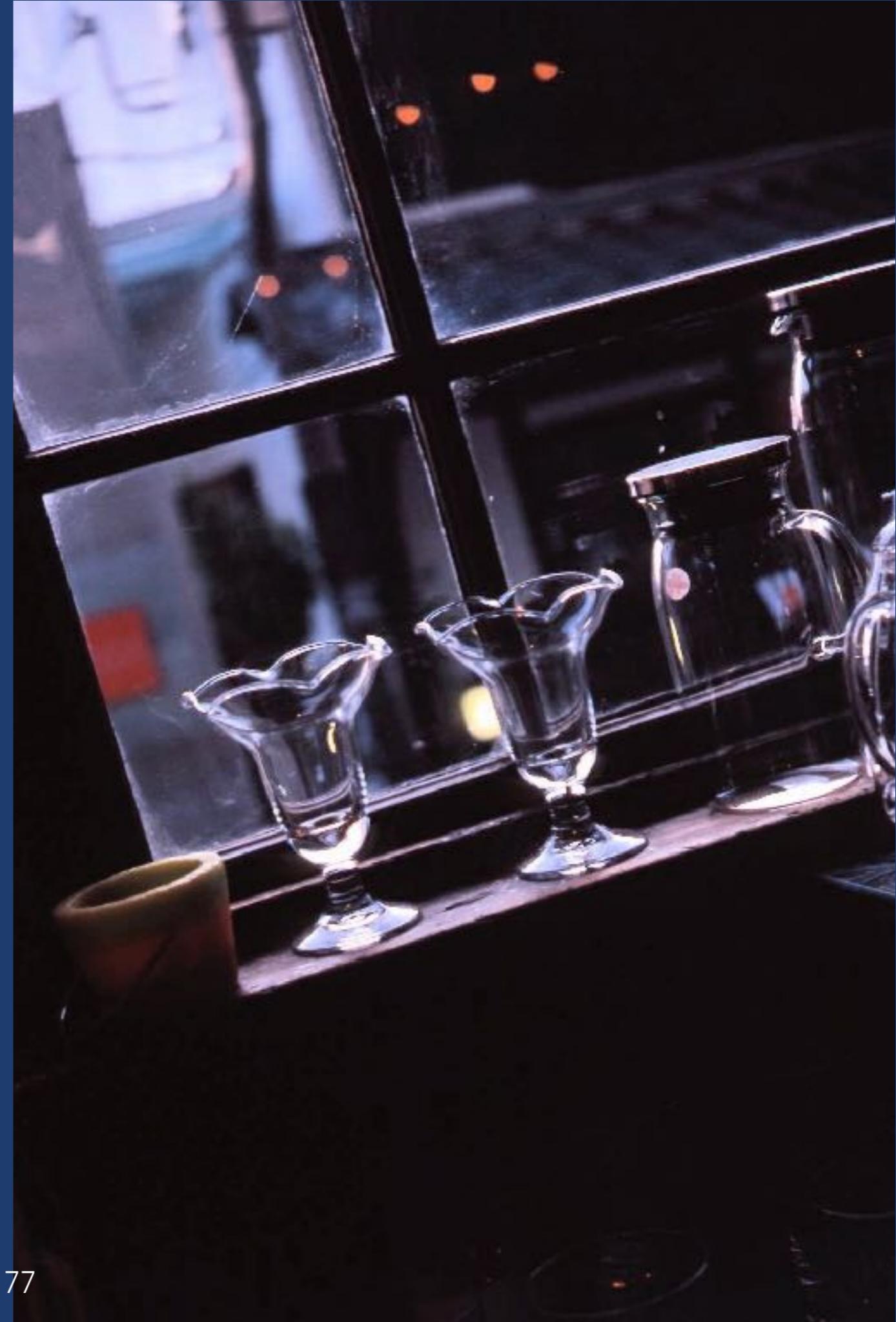
- 3つのクラスターが形成されている
- 黒 acceptables
- 赤 questionables
- 緑 unacceptables



概要

- 標準化ありの場合に競べて
 - cluster 1 (黒) と cluster 3 (緑) を隔てる凹みが生じない
 - 実際, Dim 1 の説明力が下がっている
 - そのため, 3D表現の時に表現力不足に繋がると予想されるが, それを確かめるのは難しい
 - normalization なしだと gr0, ..., gr9 を統合できない

PCA3Dで容認度
評定クラスター
を可視化



次元の問題

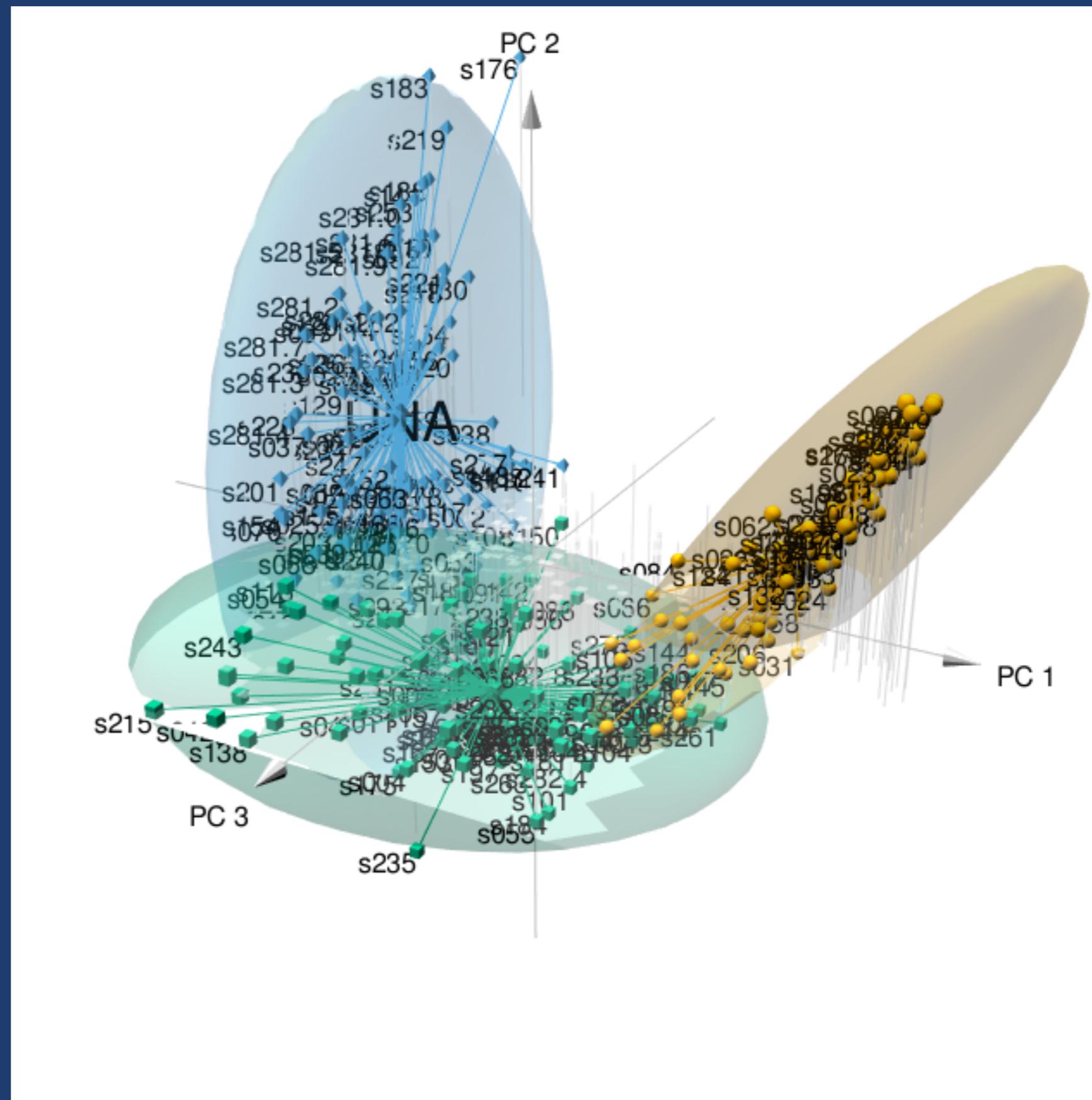
- 事実
 - $\text{SELF}(A)$ の結果は、本来3次元の構造を 2次元（平面上）に射影した結果
 - $\text{SELF}(A)$ を $r = 3$ でモデル化している
- 疑問
 - 3次元構造はどうなっているのか？

対策

- PCA3D を使って可視化
 - R の rgl パッケージ (Adler, Murdoch et al. 2018)
 - R の pca3d パッケージ (Weiner 2017)

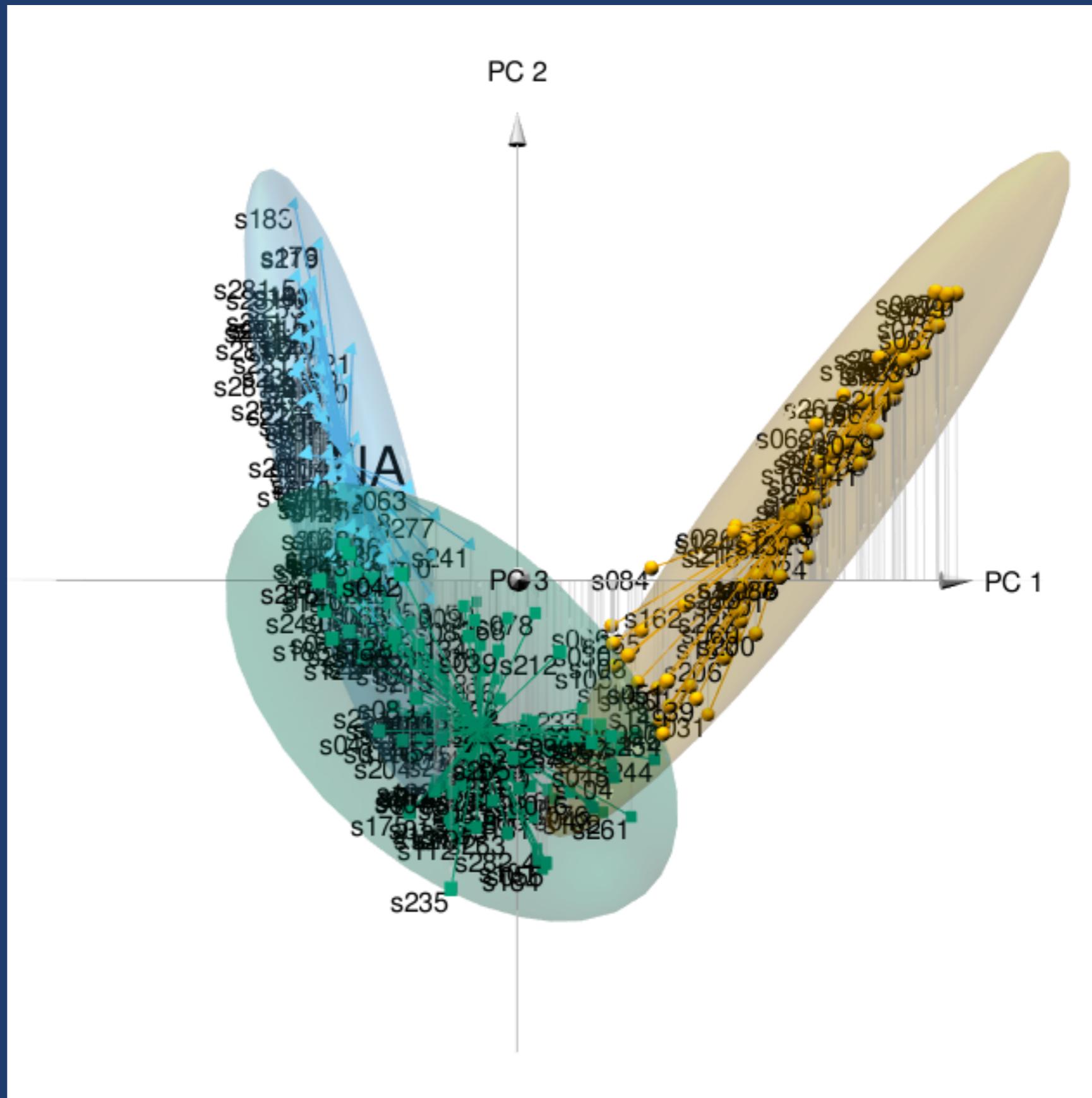
“pca3d” を使つ て A6 を可視化

- A/X/UNA の3値分類
 - p_{01} が最大なら A
(オレンジの領域);
 - p_{34} が最大なら
UNA (空色の領域);
 - その他なら X (薄緑の
領域)



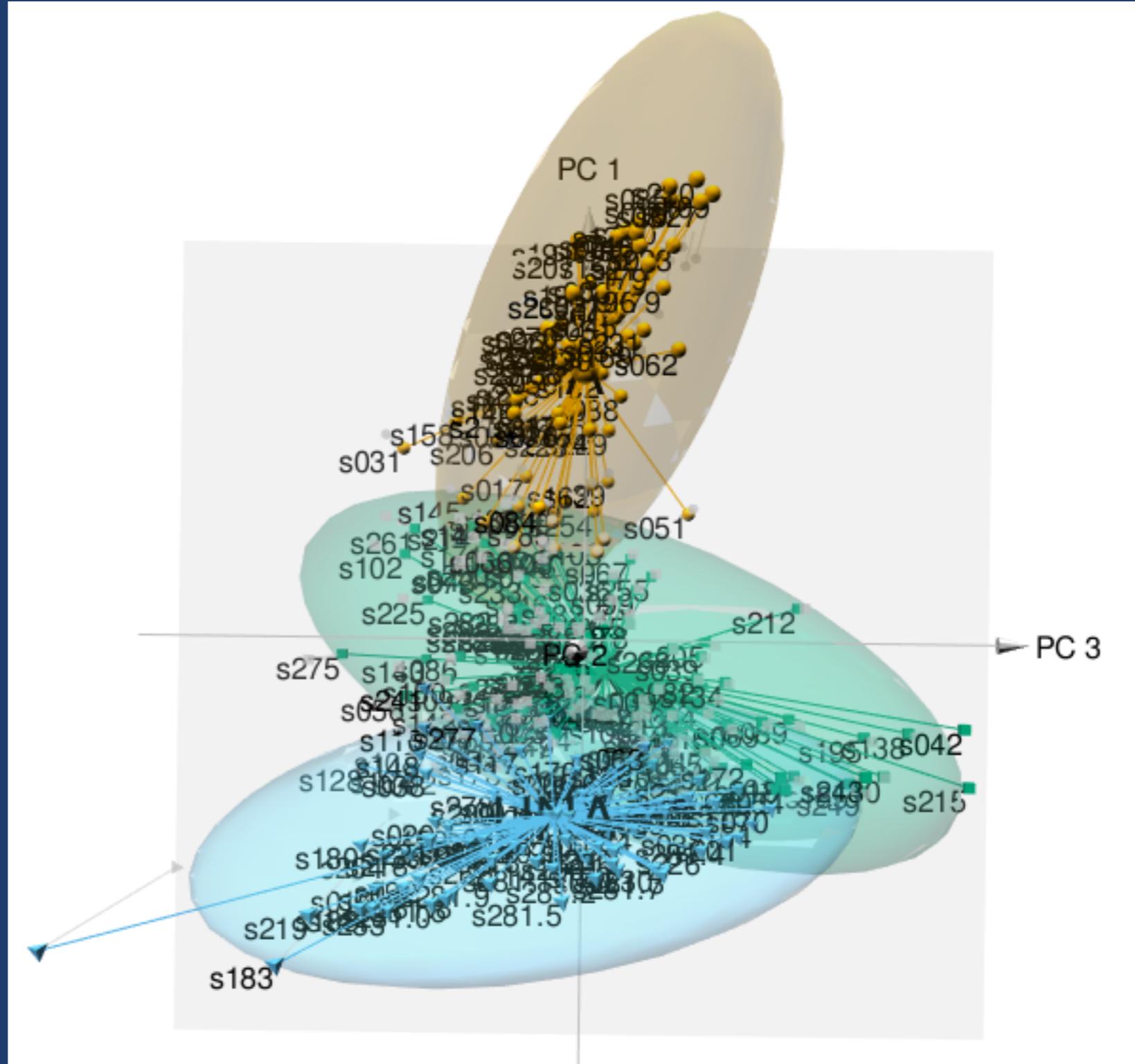
“pca3d” を使つ て A6 を可視化

- A/X/UNA の3値分類
 - PC3が見えない = 2Dで Dim 1 = PC1 と Dim 2 = PC2 が見えていたのと同じ
 - 空でない領域が狭い
=> Dim 2, Dim3 は相關している



“pca3d” を使つ て A6 を可視化

- A/X/UNA の3値分類
 - PC2が見えない角度
 - PC1 (= Dim 1) と
PC3 (= Dim 3) の関
係がわかる
 - 空でない領域が少し
広い=> Dim 1, Dim3
は少し相関している



“pca3d” を使って A6 を可視化

- A/X/U NA の3値分類
- PC1 が見えない角度
- PC2 (= Dim 2) と PC3 (= Dim 3) の関係がわかる
- 空でない領域が広い => Dim 2, Dim3 は独立している

