

インターネット上における評価データの解析

- 飲食店の評価データ解析事例と新たな図示法の提案 -

目白大学 心理学研究科
片所 強

概要

インターネット上には商品や飲食店に対して利用者が5段階評価をした数値データ、それらについての意見を自由に記述したテキストデータが存在している。こうした評価データは一商品、あるいは一店舗あたりで得られるデータ件数が少なく、ある種の統計量を算出することに実質的な意味を持たないことが多い。そこで本稿では、少数データをより分かりやすい形で図示する方法と、その際に用いる評価値の相対位置を表現する統計指標の計算方法を紹介する。また数値データとテキストデータそれぞれに基づいた評価者の分類をクラスター分析と主成分分析（もしくはコレスポンデンス分析）によって行った。その結果とテキストマイニングの結果とを併せて考察することで、評価者の特徴ごとにどのような店舗評価が行われているかを把握することが可能であるということを提示した。なお本稿で行われている解析は、実際にインターネット上より飲食店に対する評価データを収集して行われたものであり、同様の方法で他の様々な評価データを解析することが可能であることを示すものである。

目次

1	序論	2
2	方法	3
2.1	データの収集法	3
2.2	使用するデータセットと解析法	4
2.3	相対度の計算法	5
2.4	新しいグラフ化のテクニック	5
3	評価データの解析例	6
3.1	店舗間の比較	6
3.2	評価項目間の比較、および評価者間の比較	7
3.3	数値データによる評価者の分類	9
3.4	テキストデータによる評価者の分類	9
3.5	数値・テキストデータによる分類の違い	11
3.6	まとめ	12
4	おわりに	13

1 序論

今日インターネット上において商品やサービスに対しての評価データが存在しており、その量は膨大なものとなってきている。例えば Amazon.co.jp^{*1}では各種の商品に対して購入者が5段階の評価を行い、商品をレビュー (review) した結果が掲載されている。あるいは別の例として「ぐるなび」^{*2}という、飲食店を利用者が評価したデータを掲載している web サイトもある。その他にも「*** 口コミ」(** は任意のキーワードである) をキーワードとして Google などの検索エンジンを用いて検索してみると、指定したキーワードに関する評価データをいくつもの web サイトにおいて観覧することができる。統計学的な観点からみれば、どのような web サイトにおいても大概是5段階評価での数値データと自由記述 (レビュー) されたテキストデータが得られるということになる。

こうしたインターネット上に点在している商品やサービスの評価データは、少なからずとも利用者の行動に關与しているものと考えられる。例えば新たな電化製品を買う場合、その商品について「値段の割には多機能である」といったような評価がされていれば、おそらくはそれが購入を決定するためのきっかけとなるだろう。あるいは忘年会の会場を手配する時などにおいても、ある店舗についての利用者のレビューによる影響は大きいだろう。仮に「部屋の敷居がないのであまり騒げない」というレビューは忘年会の会場としては好ましくないと思われるので、利用するか否かを決定する際の参考となりうることは間違いないだろう。

このような評価は利用者側に限らず、商品やサービスを提供する側にとっても利点と欠点がある。もちろん、ある web サイト内において好意的な評価があったとすれば、それは購入数、もしくは来客数へ貢献されることだろう。しかし、web サイト上の評価に期待して購入や来店を決定した場合、それに合わないものを提供されたら再購入、再来店の可能性は低くなる。特に、事前により大きな価値を見出して (期待して) いる場合には、本来的には適切な程度のサービスにも不満を感じるかもしれないので、一概に高評価が提供する側にとって有益なものとはいえないと考えられる。

しかしいづれにしても、インターネット上の評価データが両者の立場それぞれにとっても重要な情報であることには変わりない。そこで本稿では利用者側の観点からして、いかに効率よく、かつ効果的にこうした評価データを商品の購入や飲食店へ足を運ぶための役に立たせることができるかを示すことにする。具体的には5段階評価された数値データを視覚的に表現するためのグラフィカルテクニックを紹介する。これには (a) 店舗間の比較、(b) 評価項目間の比較、(c) 評価者間の比較、といった3つの視点で数値データを視覚化して評価するという内容が含まれている。幸いにも S-PLUS に実装されている S 言語を利用することで、標準実装されているプロット機能を容易に拡張することができる。これは S-PLUS を使用する上での最も大きな利点の1つである。

一方で商品やサービスの利用者が自由記述によって個人の感想を掲載するような、いわゆる自由記述回答のデータ (テキストデータ) も重要な情報を有している。最近ではテキストマイニングという用語が社会科学における方々の分野で注目を浴びている。S-PLUS にはテキストマイニングを行うための機能 (関数) が実装されていないが、S と使い勝手が同様である R を併せて利用することでこれが可能になる。R には RMeCab パッケージ^{*3}が用意されており、一次的に R 上でテキストデータから文書行列 (term matrix) を作成し、それを S-PLUS 上で解析するといった応用ができるのも S-PLUS ならではの

*1 Amazon.co.jp の web サイト <http://www.amazon.co.jp/>

*2 「ぐるなび」の web サイト <http://www.gnavi.co.jp/>

*3 石田基広氏の web サイト (<http://rmecab.jp/wiki/index.php?RMeCab>) よりダウンロードすることができる。

利点であるといえよう。

以上のような解析は、特に本稿で取り上げた事例のように、解析の対象となるデータの件数が少ない場合に有効利用することができる。一般にインターネット上に存在する評価データの件数（一商品、あるいは一店舗あたりのレビュー件数）はそう多いものではない。したがって、分散分析や回帰分析といった線形モデルを当てはめようとするのは、推測統計学的な観点からしても無理があるといえよう。言い換えれば、単純に平均値や分散といった統計量を計算し、それについて検定を行うことに実質的な意味はなく、数値的な情報のみから何かを解釈しようとするのは困難であるといえる。そのような場合に数値データを上手く視覚的に表現することができれば、たとえ少数のデータ件数であっても十分な参考資料となりうる。この点がインターネット上の評価データを上手く活用するための重要な工夫である。

したがって本稿では、まずデータの収集法とそのデータを新たな図示法によってプロットする方法、そのために必要な統計指標の計算方法を示す。それから新たな図示法によるプロットの解釈（見方）を実例を通して説明していく。続いて評価者の分類とその特徴を把握するために、数値・テキストデータに対してクラスター分析や主成分分析（もしくはコレスポンデンス分析）を用いた解析例を紹介し、テキストマイニングによる分析結果とを併せて考察していくこととする。

2 方法

2.1 データの収集法

評価データの解析にあたって「ある飲食店に行きたいが、どの店がよいだろうか。」という利用者（状況）を想定してインターネット上の評価データを収集する。一般に最も手っ取り早い方法は Google で「イタリア料理 新宿 * * *」といったキーワード検索をすればよい。検索の際に * * * には、「口コミ」、「評価」、「評判」、「印象」などといったキーワードを選定すると効果的である。あるいは既に候補店舗があげられている場合、「イタリア料理」ではなく具体的な店舗名を用いればよいだろう。なお、場所を限定する必要がない場合、もしくは店舗名が限定的な地域にしか存在しないような場合には「新宿」というキーワードも不要であろう。

本稿ではエンターテイメント系飲食店（俗に「メイド喫茶」と呼ばれる飲食店）を対象とし、いくつかあるメイド喫茶の店舗より代表的な 4 つの店舗名をキーワードとして採用した。その店舗名は「ぴなふおあ」、「メイリッシュ」、「@ほお～むカフェ」、「めいどりーみん」である。これらのキーワードを用いたのは、店舗名が特徴的で他に類似する（検索エンジンが混同する）キーワードが含まれていないこと、店舗が存在する地域が東京の秋葉原に限定されていること、得られる評価データ量が多すぎないことから事例データとして解析するのに適切であると判断されたからである。

2010 年 7 月 19 日現在で参照できることが確認されている web サイト、検索に用いた具体的なキーワードの組み合わせ、および各サイトにおける店舗レビュー数（評価数）を表 1 にまとめている。これらのサイトはいずれも 5 段階評価と自由記述評価の 2 つが評価項目として設置されている。ただし、サイトによっては 5 段階評価における項目の種類が少し異なっている部分もある。例えば、「食べログ」というサイトでは総合評価の他に料理の味や店の雰囲気などといった項目も設けられており、一方で「Yahoo グルメ」では総合評価 1 つのみであったりする。

表1 検索ワードと参照した web サイト一覧

検索ワード	サイト名	レビュー数
びなふぉあ 評価	食べログ	5
	ivedoor グルメ	3
	Yahoo グルメ	2
メイリッシュ 評価	食べログ	5
	ivedoor グルメ	2
	Yahoo グルメ	1
@ほお ~むカフェ 評価	食べログ	5
	ivedoor グルメ	3
	Yahoo グルメ	6
めいどりーみん 評価	食べログ	3

2.2 使用するデータセットと解析法

2.2.1 数値データのグラフ化

まず前述した4つのメイド喫茶：びなふぉあ、メイリッシュ、@ほお ~む、めいどりーみんに対する5段階評価による評価データについて店舗間比較を行えるような図を作成する。これには1列目に評価値、2列目に店舗名がおかれたようなデータフレームが用いられる。データフレームは `data.frame()` によって作成される。評価値は各サイトにおける総合評価の値であり、店舗名は4つの水準をもつカテゴリカル型の変数ということになる。

次にある特定の店舗についての評価について、評価項目間の比較と評価者間の比較を行うための図を作成する。これには行方向(行要因)に評価者がおかれ、列方向(列要因)に評価項目がおかれたようなデータ行列が用いられる。データ行列は `matrix()` によって作成されるものである。ここで評価項目には総合、料理・味、サービス、雰囲気といった4項目が含まれており、本稿では特に「びなふぉあ」という店舗のデータセットについて解析を行う。

解析には S-PLUS 6.2J for Windows を用いる。なお使用するデータセットについて、店舗間比較を行うために用いるデータフレーム形式のものを表4、項目間比較と評価者間比較に用いるデータ行列形式のものを表5として本稿末尾に付録として載せてある。

2.2.2 テキストデータの解析

テキストデータの解析といっても、ここでは数値データの解析と比較するために行う。まずは項目間比較と評価者間比較で用いるデータセット(表5)に対して主成分分析とクラスター分析(ward法)を行い、評価者を分類することを試みる。それに続いてテキストデータ(R上で作成された文書行列)に基づいたコレスポネンス分析とクラスター分析(ward法)を実行する。なお、コレスポネンス分析は主成分分析と同等な解析法であり、主成分分析は連続型のデータに用いられるのに対して、コレスポネンス分析はカテゴリカル型のデータに対応したものである。S-PLUSでは `mass` パッケージに `corresp()` が用意されている(Rの場合は `MASS` パッケージに含まれる)。

解析には S-PLUS の他に R 2.11.1 とパッケージ `RMeCab_0.89` を用いる。

2.3 相対度の計算法

本稿で紹介する新たな図示法（プロット）を用いるために、ここで相対度（Relative Value, RV）という指標を定義する。これはある評価値を基準（中心）として、他の評価値がどの程度のバラつきを示しているかを表すものである。換言すればある評価値が相対的にどれほど評価値から逸脱しているのかを表すものである。

相対度の計算

n 個の評価値が格納されているデータベクトル $X = (x_1, x_2, x_3, \dots, x_n)$ があるとする。ここで基準となる評価値 x_i について相対度 RV を計算するための公式は以下のようになる。

$i = 1$ のとき、

$$RV = \frac{\sum_{i=2}^n (x_i - x_1)^2}{n} \quad (1)$$

$i = k (1 < k < n)$ のとき、

$$RV = \frac{\sum_{i=1}^{k-1} (x_i - x_k)^2 + \sum_{i=k+1}^n (x_i - x_k)^2}{n} \quad (2)$$

$i = n$ のとき、

$$RV = \frac{\sum_{i=1}^{n-1} (x_i - x_n)^2}{n} \quad (3)$$

式（1）（2）（3）^{*4}はそれぞれ分散を求める公式に似ており、分散の公式では各観測値と平均値の差の自乗和（偏差平方和）を計算するが、RV では平均値の代わりに基準となる評価値が用いられることになる。これは分散が平均値を中心としたバラつきの程度を表す指標であるのに対して、基準となる評価値を中心としたバラつきの程度を表す指標であると解釈できる（この指標の解釈についての詳細は実際の解析例を通して示す）。なお、これを S-PLUS を用いて計算するための関数を $RV()$ と $RV2()$ として、付録にソースコードを載せてある。

2.4 新しいグラフ化のテクニック

評価データの解析例を示す前に、本稿で新たに提案するグラフ化のテクニックについて説明しておく。基本的には従来より用いられている散布図と同じ要領で描かれることになる。まず横軸に評価値、縦軸に先ほど紹介した相対度をとって散布図を出力する。しかし 5 段階評価のような数パターンの回答に制限されているようなデータの場合、散布図として出力してもデータポイントが重なってしまい、全体の特徴を把握するのは難しいことが多い。そこである評価値の回答数（度数、頻度）に比例して、描かれるデータポイントの大きさを調節する工夫が効果的であると考えられる（図 1）。

^{*4} これらの式はいずれも実質的には同じもので、 $i = 1$ から k までの任意の評価値をベクトルの先頭（第 1 要素）に置くと考えれば、全ては式（1）によって計算できる。付録にある $RV()$ は式（1）を定義したもので、 $RV2()$ は単にベクトルの順序を入れ替えて $RV()$ を用いて計算するためのものである。

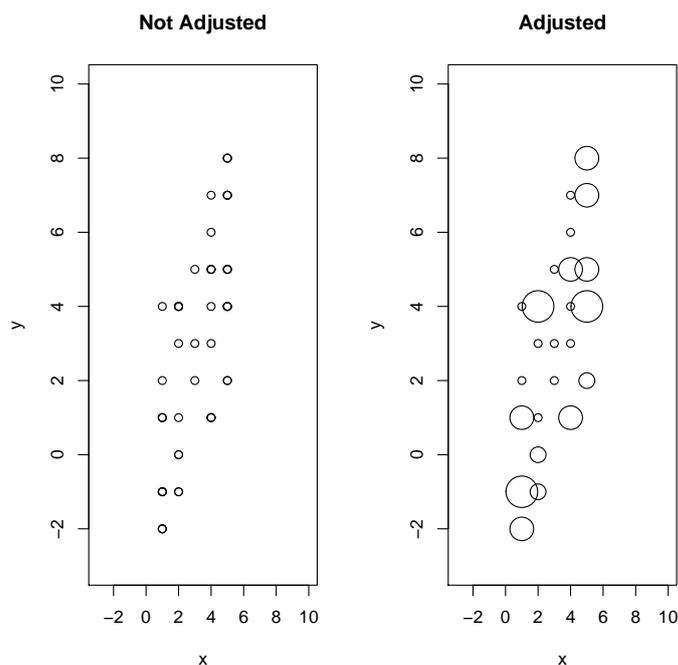


図1 度数に対してデータポイントの大きさを調節するプロット

そもそも図1の右図(Adjusted)のようなプロットはバブルプロット(Bubble Plot)もしくはバブルチャート(Bubble Chart)として知られているもので、3次元データを2次元平面上に表現するグラフィカルテクニックの1つである。ところが、このような方法はサンプルサイズが大きくなるとデータポイントが重なって、かえって見にくいものとなってしまいう欠点もある。また、データポイントの大きさを度数に比例して描くような場合、その値が大きくなるほどデータポイントも大きくなるので、平方根をとるなどの工夫も必要になってくる。

しかし本稿で取り上げたような、インターネット上における評価データのように、さほど大規模ではないデータに対しては非常に有効な方法であるといえる。グラフ化の発想の原点はバブルプロットにあるが、さらにいくつかの工夫を加えることで、評価データを考察するために効果的な図として用いることができる。それを次章より実例をふまえて示すこととする。

3 評価データの解析例

3.1 店舗間の比較

まず表4に基づいて、4つの店舗それぞれがどのような評価を受けているかを比較するためのプロットを図2に示す。横軸に評価値(evaluation)、縦軸に式(1)によって計算された相対度(RV)をとって散布図をプロットしてある。この際、評価値の度数に比例してデータポイントが大きくなるようにしており、各ポイント同士を直線で結ぶかたち(折れ線グラフ)にしてある。各色で描かれている十字は、x軸について評価値の平均値、y軸についてRVの平均値が直交する座標に打たれたものである。

図を解釈するとき、最初に各色の十字ポイントを参照するとよい。横軸は評価値であるから、単純に右側に位置する店舗ほど平均的に高い評価を受けていることが分かる。この場合だと@ほお~むカフェ(athome)とぴなふぉあ(pina)の評価が高いという結果になっている。一方で縦軸はRVであるが、

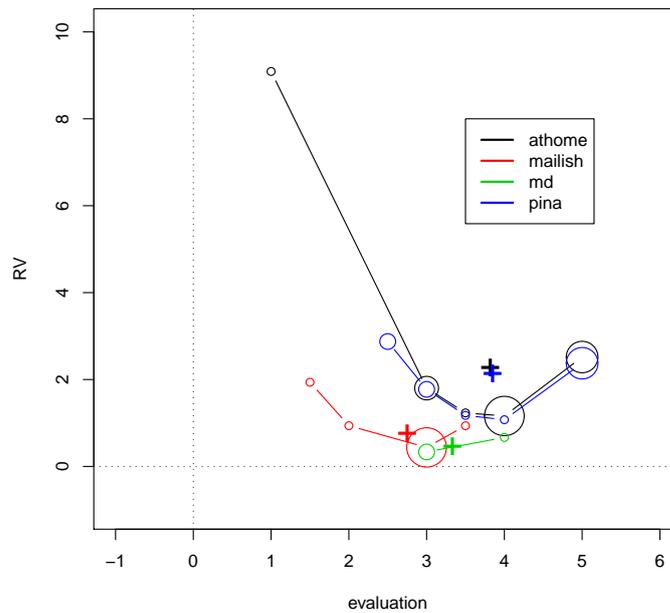


図2 店舗間の比較をするためのプロット

RV は簡単にいえば、高い値を有しているほど中心的な評価値（例えば「3」である）から離れた評価値（例えば「1」や「5」である）が得られている程度が大きいと解釈することができる。換言すると得られている評価値のうち、両極端な評価値が多く存在しているほど、この値は高くなる。

例えば、メイリッシュ (mailish) を見てみると、図からも「1.5」、「2」、「3」、「3.5」という評価が得られていることが確認できる。そして円の大きさは「3」が最も大きく、他の円は小さいものとなっている。これは得られた評価値のうち、中心的な値である「3」に評価（回答）が集まっているためである。これに対してぴなふぉあ (pina) について、こちらは中心的な評価値である「3.5」や「4」に比べて、両端の評価「1」や「5」に多くの評価が集まっているので RV 値が高くなっているのである。評価値の平均値が高いことから、ぴなふぉあは高い評価値に偏った評価、つまり「総合的には良い店」とされていることが分かる。

一方、RV 値が高い別の場合として@ほお～むカフェ (athome) が良い例となる。これは円の面積が「3」、「4」、「5」共に大きくなっているが、RV の平均値は高くなっている。特に中心的な「4」が最も大きな面積を有しているにもかかわらず RV が高い理由は、「1」という極端な評価が得られているためである。ある意味ではずれ値ともとれるが、口コミ評価などでは、こうした他の評価者から逸脱した評価に、他の評価者からは得られない重要な情報が潜在していることが多いので無視できない。これについては、後ほどテキストデータを用いた解析と併せて考察してみることにする。

3.2 評価項目間の比較、および評価者間の比較

先ほどは店舗間、すなわち「どの店舗が良いか？」という観点からみたものであったが、今度はある特定の店舗に限定して4つの評価項目（総合、料理・味、サービス、雰囲気）と、その4つの項目について各評価者がどのような評価をしているかという評価特徴をみていくことにする。ここではぴなふぉあを対象としてみていくことにする。

まず図 3 は評価項目（表 5 の列要因）についてプロットしたものである。凡例のラベルは総合：comprehensive、料理・味：food、サービス：service、雰囲気：atmosphere となっている。

十字のデータポイントに着目すると、料理・味（food）以外についてはどれも「3」から「4」の範囲に位置しており、全体としては高い評価を受けていることが確認できる。その中でも雰囲気（atmosphere）は高い評価を受けており、RV も小さいことから平均値から離れている回答が少なめであることが分かり、円の面積をみても「4」と「5」が相対的に大きくなっていることから「良い雰囲気」である店舗だといえる。総合（comprehensive）とサービス（service）は RV が雰囲気よりも少し高く、平均値から離れた回答が多いことが円の面積からも確認できる。より良い評価である「5」もあるが、対してより低い評価である「2.5」もあるので平均値に惑わされないように注意する必要がある。料理・味については RV が高い値となっているが、これは「0」という評価の影響を受けているためであり、これは欠損値を「0」で補完したためであり、実際にはもっと小さな値となるだろう。それを考慮すると「3」付近に回答が集中しているが、平均値で見れば全体的には好印象を受けている店舗であることが推察される。

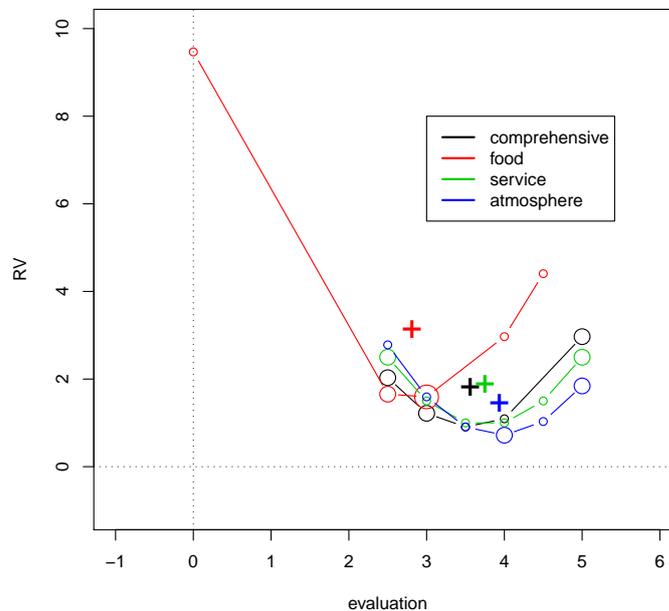


図 3 びなふおあの評価項目間の比較

続いては評価者間の比較を行う。図 4 は評価者（表 5 の行要因）についてプロットされたものである。この図においてとりわけ目立つのは評価者 1 であるが、先に述べたようにこの評価者の「0」という値は欠損値を補完したものである。したがって、この値を除けば黒丸は $x = 0$ 軸上に乗るかたちに（被験者 2 のように）なる。

どの評価者もほぼ一点にデータポイントが集中しており、RV も総じて低いことから、どの評価項目についても同様な評価を下していることが分かる。このように RV はある特定の評価者がどれほどバラツいた（あるいは一点に偏った）回答をしているかを確認する場合に役立ち、評価値との散布図として図示することでより直感的に解釈しやすいものとなりうる。

またこの図は、例えば平均値（十字のデータポイント）に着目すると、評価者 5 と評価者 7 は他の評価者と比べて高い評価を行っている。そのうえ RV と円の面積から「5」に偏った回答をしていることが確認できる。これに対して評価者 2 や評価者 3 は逆に「2」と「3」の辺りに集中している。こういった

ことから即座に低評価者と高評価者、あるいは平均的な評価者を見つけることができ、他の自由記述回答を見比べるときの良い判断材料となりうるだろう。

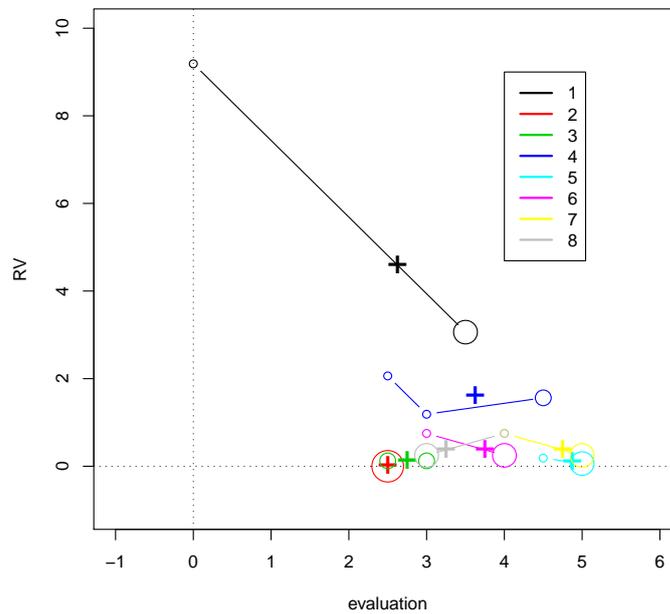


図 4 びなふおあの評価者間の比較

3.3 数値データによる評価者の分類

ここまで店舗間の比較（図 2）、ある特定の店舗（びなふおあ）についての評価項目間の比較（図 3）、そして評価者間の評価の特徴（図 4）をみてきた。これ以降は 5 段階評価の数値データと併せて得られている、テキストデータも加えてより詳しく評価データを解析していく。これに先立って、まずは数値データ（表 5）に基づいて主成分分析、およびクラスター分析（ward 法）によって評価者の分類を行う。その結果が図 5 に示されている。

図 5 の左図はクラスター分析の結果であるが、デンドログラムから 4 つのクラスター（グループ）にわけられることが見て取れる。また右図は主成分分析の第 1 主成分と第 2 主成分をプロットしたものであるが、これも 5, 7; 8, 6, 4; 3, 2; 1 という 4 つのそれぞれ類似した特徴を持った回答者が集合していることが確認できる。ここで図 4 とを合わせて見ると、大まかに分類ラベルをつけてしまえば 5, 7 は高評価群、8, 6, 4 は中評価群、3, 2 は低評価群、そして 1 はその他としてまとめられるだろう。

3.4 テキストデータによる評価者の分類

さて数値データに基づいた分類の一方で、テキストデータに対して文書行列を作成し、それをクラスター分析とコレスポンス分析によって解析した結果が図 6 である。左図のデンドログラムによれば 3 つのクラスターに分類されるが、コレスポンス分析の評価者についてプロットした結果とは異なるようである。前者は 1; 7, 2, 4, 5; 3, 6, 8 という 3 分類に対して、後者は 3, 5, 7; 1, 4; 6, 8; 2 といった括りであると読み取られる。

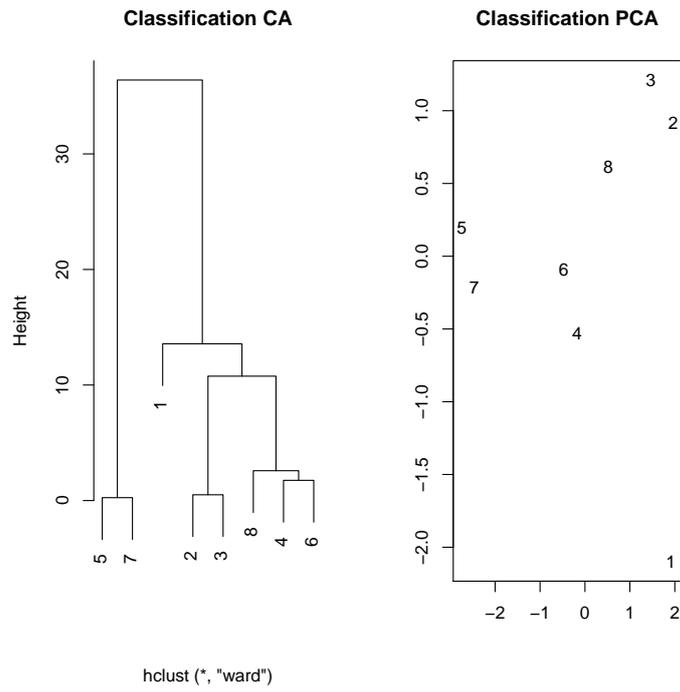


図 5 数値データに基づく分類

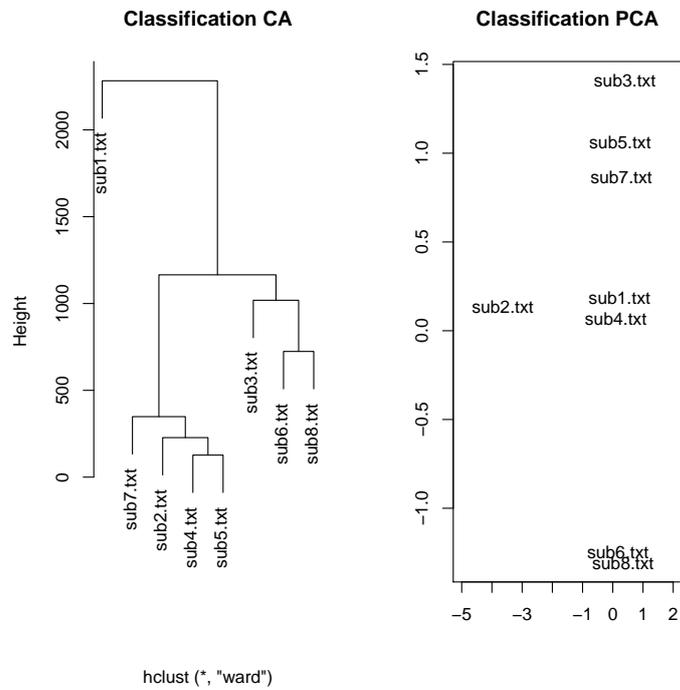


図 6 テキストデータに基づく分類

3.5 数値・テキストデータによる分類の違い

まとめると数値データに対してクラスター分析と主成分分析を行った場合、両者は同様の結果として4つの分類がなされるようであった。それに対してテキストデータの場合は、クラスター分析とコレスポネンス分析とで数値データほど共通した分類がなされていないようであった。また、数値データによる分類とテキストデータによる分類の両者を見比べても、はっきりとした共通分類が見られなかった。

こうした結果より、インターネット上における評価データにおいて、数値(5段階評価による)データで表される評価とテキスト(自由記述による)データで表現される評価とでは、両者はかなり相違した評価が下されている可能性があるかと推察される。例えば、数値による評価では高い評価を下していても、自由記述による評価ではあまり高い評価をしていないかもしれないということである。これは評価者本人が意識していなくても、自由記述による評価にはネガティブな評価(形容詞)を書いているということの現われかもしれない。

3.5.1 自由記述の内容と数値評価をふまえた分析

そこで今度は数値・テキストデータに基づいた分類結果を参考にし、いくつかの自由記述回答の内容について見ていくことにする。ここでは1つの例として、特に評価者5と評価者7について見てみる。数値データによる分類(図5)によれば、彼らは高評価群に含まれる。テキストデータによる分類(図6)でも、コレスポネンス分析の結果で2人は近い位置に布置されている。つまり、数値評価でも自由記述による評価でも、この2人に限っていえばかなり類似した特徴を有しているものと察することができる。

評価者5の自由記述の中から形容詞を抽出^{*5}すると、「安い」と「良い」という単語がある。これだけでは何が「安い」のか、もしくは「良い」のか分からないので共起語分析^{*6}を行う。まず「安い」という単語をノード(node, 中心語)とし、スパン(span, 単語の範囲)を3として実行したところ「思ったより安い」という文章が想起され、同様に「良い」の場合は「サービスが良い」という評価をしていることが分かった。

他方で評価者7についても同様の分析を行ったところ、まとめると「メイドさんがオムライスにケチャップで描いてくれる絵がかわいい」、「思っていたより安い」、「ファミレスよりは少し高い」、「店内の雰囲気は明るい」、「観光客っぽいお客さんがたくさんいる」という表現があげられた。以上の結果を表2にまとめてあるが、これを概観すると、どうやら最初は値段が高いという印象をもっているが、実際には思っていたほど高くないと感じているようであることが分かる。本稿の冒頭でも述べたが、期待していたよりも大きな価値(高いと思っていたのに安かった)が得られたことで高評価へと繋がったものであるといえよう。

さて、ここで改めて図4と図5を見てみると、数値評価的には評価者5と評価者7と対極な位置にあるのは評価者2と評価者3である。数値データに対するクラスター分析と主成分分析の結果の節で、彼らは「低評価群」とラベル付けできると述べた。ところがテキストデータによる分類図(図6)を見ると、デンドログラムでは全く異なる枝の終端に位置しているし、コレスポネンス分析の図においても両者はかなり離れた位置に布置されている。そこで両者の自由記述内容がどのような点で異なるのかを、先ほどの同じ方法(共起語分析)で確かめる。

*5 RMeCab パッケージに含まれる RMeCabFreq() を用いた。

*6 RMeCab パッケージに含まれる collocate() を用いた。

表 2 形容詞についての共起語分析（高評価者の意見）

ノードとされた形容詞	共起語分析より想起される表現
安い（評価者 5）	思ったより安い
良い	サービスが良い
かわいい（評価者 7）	メイドさんがオムライスに描いてくれる絵がかわいい
安い	思っていたより安い
高い	ファミレスよりは少し高い
明るい	店内の雰囲気は明るい
っばい	観光客っばいお客さんがたくさんいる

評価者 2 の自由記述に出現する（抽出された）形容詞は「かわゆい」と「すっごい」であった。この 2 つの語をノードとし、それぞれ適宜スパンの値を変更して分析したところ「描いてくれる絵がすっごいかわゆい」という意見が想起された。したがって、数値による評価の割には否定的な形容詞が存在しておらず、むしろ肯定的な表現がなされていることが明らかになったといえる。

これとは逆に評価者 3 は、例えば「悪い」、「古い」、「無い」といった否定的な形容詞が記述に含まれていた。実際に前後の文脈を読み取ってみると「素人っばい子が良いのか悪いのか分からない」、「壁が黄ばんで古さを感じる」、「喫煙席との仕切りが無い」という意見を述べている。なお、共起語分析によってノードとなる単語に共起する単語が出力されるが、それらから特定の意見（記述）が推察できない場合、テキストファイルからこうした形容詞を検索してみれば、前後の文脈が読み取ることができ効率的である。この評価者 3 についてはその方法を採用した。

以上、評価者 2 と評価者 3 の意見についてまとめたものを表 3 に示す。

表 3 形容詞についての共起語分析（低評価者の意見）

ノードとされた形容詞	共起語分析より想起される表現
かわゆい + すっごい（評価者 2）	描いてくれる絵がすっごいかわゆい
悪い（評価者 3）	素人っばい子が良いのか悪いのか分からない
古い	壁が黄ばんで古さを感じる
無い	喫煙席との仕切りが無い
古い + 無い	喫煙席との仕切りが無いため、壁が黄ばんで古さを感じる

3.6 まとめ

3.6.1 どの店舗が良いか

店舗間の比較について図 2 より、全体的に良い評価である店舗は「@ほお～むカフェ」と「びなふぉあ」であった。前者は 1 人だけ逸脱した評価を下している評価者が居るので RV が高かったが、それ以外は「4」から「5」の評価数が多いという特徴があった。対して後者は平均評価値は高いが、中心評価値よりも両端の評価値に回答度数が偏っているという特徴があった。これらとは別にメイリッシュは「3」に回答が集まり、RV 値も低いことから定評を受けているという特徴であった。それから「めいどりーみん」はそもそもの回答数が少ないので比較対象としては少し不十分であった。

3.6.2 その店舗の評価はどうか

本稿では特に「ぴなふぁあ」を選定して評価データを解析した。まず4つの評価項目：総合、料理・味、サービス、雰囲気はどれも平均的に高い評価を受けていた（図3）。また各評価者の評価特徴を見ると、ある評価値一点に集中して回答している（例えば「3」だけといったように）傾向が見て取れた（図4）。これに続いて、さらに評価者の特徴を分類するために数値データとテキストデータに対してクラスター分析と主成分分析、テキストデータにはコレスポネンス分析を行った（図5および図6）。

特に高評価群と低評価群との自由記述の内容を分析したところ、高評価者の記述内には否定的な形容詞が含まれておらず、抽出された肯定的な形容詞から想起される意見も好評なものであった（表2）。それに対して低評価者の記述には、否定的な意見がなく肯定的な意見である者（評価者2）と否定的な意見が顕著に目立った者（評価者3）とが確認された。

数値評価上の分類では、両者は同一グループに属するが、テキストデータに基づいた分類ではほぼ両極的な位置に布置されていた。このことより、数値・テキストデータそれぞれによる分類精度はかなり原データの評価を反映した結果として得られることが示されたといえる。また、数値上では低評価者とされる場合であっても、自由記述では比較的に高評価をしていたり、数値上での評価と同じく低評価を下すこともあることが示された。これはどちらか一方のデータを吟味するだけでは、十分な情報が得られないという事実を表したものであるといえる。

4 おわりに

本稿ではインターネット上にある評価データ：数値データとテキストデータの解析例を示した。特に数値による評価データは5段階評価（5件法）によって回答される場合が多い。こうした数値データは店舗ごとに見れば少なければ2~3件程度、多くても10件に満たないほどである。そして、そうした少数の件数において平均値や分散といった統計量を計算したところで実質的な意味を持たないことが多い。そこで少数しかないデータであっても、図示する方法を工夫することでより高度な解析（解釈）が可能であることを紹介した。また、多くの口コミサイトなどには5段階評価の項目とは別に、自由記述回答の項目が設置されており、こうしたテキストデータも重要な情報である。したがって、数値データとテキストデータを相互的に考察することで、より有益な情報が得られるといえる。

具体的に本件で示した解析の流れをまとめてみると、まず数値データを新しい方法で図示し（図2、3、4）、店舗間比較、評価項目間比較、評価者間比較を行った。つぎに数値データ（表4、5）を用いてクラスター分析と主成分分析を行い評価者の分類を試みた。またこれに加えてテキストデータについても、R上で文書行列を作成し、それをS-PLUS上でクラスター分析とコレスポネンス分析を行った。そして数値・テキストデータに基づいた分類を参考にして自由記述回答の内容について吟味した。

利用者側の立場からすれば、日常的な生活場面において、これほど手間のかかることをする必要を感じることは少ないかもしれない。しかし、今日ほどインターネット上に様々なサービスコンテンツが提供されている中で、こうした利用者側から提供されたデータを統計学的に考察することは提供する側として重要な意味を持つことだろう。例えば、自分の経営するいくつかの店舗の評価を比較することで、低評価の店舗について接客態度を見直すきっかけとなるかもしれない。あるいは、自分の会社で提供している製品と他者が提供している製品との評価を比較することは、特定の商品を宣伝する必要性を知らされることになるかもしれない。いずれにせよ本稿で紹介したグラフは、付録のソースコードを貼り付けるだけで簡単に作図できるので多くの分野で活用されることを期待するところである。

また、今回は店舗間の比較から、特定の店舗とその評価者の特徴についての解析例は1つの店舗(びなふぉあ)のみであったが、同様な作業で他店舗の特徴を把握することができる。もちろん飲食店に限らず、書籍や電化製品でも良いし、動画配信サービスにおける動画の評価であったりなど様々な評価データに対して応用できる。本稿の序論でも述べたように、現在はインターネット上に今回取り上げたような評価データが大量に埋もれているので、これらを活用しない手はないだろう。

今回の解析の所々でRを利用しているが、数理システムからText Mining Studioというテキストマイニングツールが製品として提供されているので、本来ならばこちらを積極活用すべきかもしれない。しかしS-PLUSユーザーとすれば、RはSとほぼ同様の作業(コマンド)で解析を行えるので、S-PLUSユーザーであっても容易にテキストマイニングを行える可能性を示すためにあえて紹介することとした。もちろん、より高度な解析となればGUIによる操作が可能であるText Mining Studioのような製品が威力を発揮するだろう。ただし、あまり大規模でないテキスト分析であればS-PLUSの補助としてRを使うことも効果的であるといえる。いずれにしてもS-PLUSはRと併せて使いこなせる環境であるという利点を改めて強調しておきたい。

現在は多変量解析といった複雑で高度な解析を容易に実行できる環境が整っているため、基本的な統計量を吟味したり、図を作成してデータの特徴を把握するといったことが軽視されがちな場面も少なくない。あるいは、あまりにも作業的に統計的仮説検定が行われるため(有能な製品によって実行できてしまうため)実質的な、固有科学としての価値判断を忘れがちであるかのような風潮もうかがえる。今回、web上の評価データを単純でも少しの工夫を加えた図で表現する方法を提案したのは、こうした現状において見過ごしがちな点を再考するためのきっかけの1つとなれば幸いであると考えたからである。データ解析は決して作業的なものではなく、常に分析者の創意工夫が要されるべきものであることを加えて主張して結びとする。

参考文献

- [1] Brian S. Everitt. *An R and S-PLUS Companion to Multivariate Analysis*. Springer-Verlag, 2005. (石田基広 [訳]. RとS-PLUSによる多変量解析. シュプリンガー・ジャパン, 2007.)
- [2] 早川清, 山崎竜, 木全直弘, 清水銀嶺, 佐藤楓. *メイド喫茶で会いましょう*. アールズ出版, 2008.
- [3] 石田基広. *Rによるテキストマイニング入門*. 森北出版株式会社, 2008.
- [4] 金明哲. *テキストデータの統計科学入門*. 岩波書店, 2009.
- [5] Uwe Ligges. *Programmieren mit R*. Springer-Verlag, 2004. (石田基広 [訳]. Rの基礎とプログラミング技法. シュプリンガー・ジャパン, 2006.)
- [6] 間瀬茂. *Rプログラミングマニュアル*. 数理工学社, 2007.

付録

```
# 相対逸脱度値を計算する
# 評価データのベクトルを指定すれば OK
#   ただし第 1 要素に基準となる評価データを置くこと
# 使用例：
# x <- c(4, 1, 1, 2, 5) # 評価データ
# y <- RV(x)           # 相対逸脱度値
RV <- function(x){
  n <- length(x)
  res <- sum((x[2:n] - x[1])^2) / n
  res
}
```

```
# 相対逸脱度値を計算する 使用の際には RV() が別途必要
# 評価データのベクトルを指定すれば OK
# 使用例：
# x <- c(4, 1, 1, 2, 5) # 評価データ
# y <- RV2(x)           # 相対逸脱度値
RV2 <- function(dat){
  n <- length(dat) # ベクトルの長さを取得
  res <- numeric(n) # 大きさ n の変数 res を用意
  for(i in 1:n){
    if(i == 1){ # 1 回目の繰り返しならば,
      num1 <- dat[i] # 1 番目の要素を num1 に代入
      num2 <- dat[(i+1):n] # 2 から n までの要素を num2 に代入する
      res[i] <- RV(c(num1, num2))
    }else if(i == n){ # n 回目の繰り返しならば,
      num1 <- dat[i] # n 番目の要素を num1 に代入
      num2 <- dat[1:(i-1)] # 1 から n-1 番目までの要素を num2 に代入
      res[i] <- RV(c(num1, num2))
    }else{ # それ以外: 2 から n-1 回目までの繰り返しならば,
      num1 <- dat[i] # i(1 > i > n) 番目の要素を代入
      num2 <- c(dat[1:(i-1)], dat[(i+1):n])
      res[i] <- RV(c(num1, num2))
    }
  }
}
```

```

}
return(res)
}

```

```

# arw.plot.df() と arw.plot.mat() を利用するために必要な関数
arw.plot2 <- function(x, y, arw=FALSE,
                      renew=FALSE, pcol, ylimit=FALSE, xlimit=FALSE){
  xy.tbl <- table(x, y) # テーブル集計
  nx <- nrow(xy.tbl)    # 行数の取得
  ny <- ncol(xy.tbl)    # 列数の取得
  vx <- as.numeric(rownames(xy.tbl)) # 行ラベルを numeric に変換
  vy <- as.numeric(colnames(xy.tbl)) # 列ラベルを numeric に変換
  # 空のプロットエリアを作成する
  # renew == TRUE なら plot() で空のプロットエリアを作成しない
  if(renew == FALSE){
    plot(
      c(-1, ifelse(xlimit=="TRUE", as.numeric(readline("xlim : ")), max(x)+1)),
      c(-1, ifelse(ylimit=="TRUE", as.numeric(readline("ylim : ")), max(y)+1)),
      xlab="evaluation", ylab="RV", type="n")
    abline(h=0, v=0, lty=3)
  }
  # 散布図：データポイントの大きさを調節しながら描く
  # xy.tbl における各セルの度数が cex の値として指定される
  for (i in 1:nx) {
    for (j in 1:ny) {
      if (xy.tbl[i, j] > 0) points(vx[i], vy[j], cex=xy.tbl[i, j], col=pcol)
    }
  }
  a <- cbind(x, y)
  a <- a[order(a[,1]), ]
  points(a[,1], a[,2], col=pcol, type="c")
  xm <- mean(x) # 引数 x に指定されたデータの平均値
  ym <- mean(y) # 引数 y に指定されたデータの平均値
  points(xm, ym, pch="+", cex=2, col=pcol) # 座標 (xm,ym) に点を打つ
}

```

```

# データフレームに対応バージョン (arw.plot2() と RV2() が別途必要)
# データフレームの第 1 列には評価値, 第 2 列には店舗名 (カテゴリカル型) をおくこと
# 引数には DF:データフレーム, adj:軸の調整をする場合は TRUE とする
arw.plot.df <- function(DF, adj=FALSE, legend.adj=FALSE){
  nc <- ncol(DF)
  if(nc > 2) stop(message="適切なデータフレームではありません。")
  lab <- levels(DF[, 2]) # 第 2 列の水準を調べる
  lab.n <- length(lab) # 水準の数
  for(i in 1:lab.n){
    x <- DF[DF[,2] == lab[i], ] # 第 i 水準に該当するデータを抽出
    x <- x[,1] # 評価値のみを x に代入しなおす
    y <- RV2(x) # 評価値に対する RV を計算し y に代入
    if(i == 1){
      # adj=TRUE ならば軸の設定をコンソール画面上で行う
      arw.plot2(x, y, pcol=i, xlimit=adj, ylimit=adj)
    }else{
      # 重ねて他の水準のデータポイントを描いていく
      arw.plot2(x, y, pcol=i, renew=TRUE)
    }
  }
}
# 凡例
legend(
  ifelse(legend.adj=="TRUE", as.numeric(readline("x : ")), -1),
  ifelse(legend.adj=="TRUE", as.numeric(readline("y : ")), 1),
  c(lab), col=1:lab.n, lwd=2)
}

```

```

# 行列に対応バージョン (arw.plot2() と RV2() が別途必要)
# 引数 dat:データ行列, rc:1 ならば行, 2 ならば列についてプロットされる
# adj:作図領域の横軸と縦軸の調整を自分でするかどうかの指定
arw.plot.mat <- function(dat, rc, adj=FALSE, legend.adj=FALSE){
  nr <- nrow(dat) # 行数を取得
  nc <- ncol(dat) # 列数を取得
  # rc=1 とされたら行についてプロットする
  if(rc == 1){
    for(i in 1:nr){

```

```

x <- c(dat[i, ]) # 第 i 行のデータベクトル
y <- RV2(x)      # それに対する RV 値
if(i == 1){
  arw.plot2(x, y, pcol=i, xlimit=adj, ylimit=adj)
}else{
  arw.plot2(x, y, pcol=i, renew=TRUE)
}
}
}
# 凡例の描写を行う
# 行ラベルを取得し, NULL ならば数値がラベルに使われる
row.name <- rownames(dat)
if(is.null(row.name) == TRUE) row.name <- as.character(1:nr)
legend(
  ifelse(legend.adj=="TRUE", as.numeric(readline("x : ")), -1),
  ifelse(legend.adj=="TRUE", as.numeric(readline("y : ")), 1),
  row.name, col=1:nr, lwd=2)
}else{ # rc=1 以外なら列についてプロットする
for(i in 1:nc){
  x <- c(dat[, i]) # 第 i 列のデータベクトル
  y <- RV2(x)      # それに対する RV 値
  if(i == 1){
    arw.plot2(x, y, pcol=i, xlimit=adj, ylimit=adj)
  }else{
    arw.plot2(x, y, pcol=i, renew=TRUE)
  }
}
}
# 凡例の描写
# 列ラベルを取得し, NULL ならば数値がラベルに使われる
col.name <- colnames(dat)
if(is.null(col.name) == TRUE) col.name <- as.character(1:nc)
legend(
  ifelse(legend.adj=="TRUE", as.numeric(readline("x : ")), -1),
  ifelse(legend.adj=="TRUE", as.numeric(readline("y : ")), 1),
  col.name, col=1:nc, lwd=2)
}
}
}

```

表 4 店舗間比較用のデータフレーム

評価値	店名
3.5	pina
2.5	pina
2.5	pina
3	pina
5	pina
4	pina
5	pina
3	pina
5	pina
5	pina
3	mailish
3.5	mailish
3	mailish
2	mailish
1.5	mailish
3	mailish
3	mailish
3	mailish
4	athome
3	athome
4	athome
3	athome
3.5	athome
4	athome
5	athome
3	athome
4	athome
5	athome
1	athome
5	athome
5	athome
4	athome
3	md
3	md
4	md

表 5 項目間および評価者間比較用のデータ行列

	総合	料理・味	サービス	雰囲気
マロ猫	3.5	0	3.5	3.5
えみみ	2.5	2.5	2.5	2.5
みゅーぱぱ	2.5	3	2.5	3
aice	3	2.5	4.5	4.5
rikimaru 軍曹	5	4.5	5	5
すぼだま	4	3	4	4
popopopo	5	4	5	5
グッドスピード	3	3	3	4