

R と自己組織化マップを用いたパチンコ機種の分類

名列番号：029 千田 直樹

1. はじめに

R は、統計言語 S をオープンソースとして実装し直した統計解析ソフトであり、拡張パッケージをインストールして機能を自由に拡張できる。本研究課題では、自己組織化マップを用いてパチンコ機種の分類を行い、どのような結果が得られるかを検証する。

2. 研究に用いたデータ

2007 年に登場したパチンコ機種に関するデータを下記のサイトからダウンロードし、加工して用いた。

<http://www.ashh.net/total.html>

加工の結果、251 機種について、6 つの特徴次元（当たり確率、確率変動率、推定出球数、トータル確率（確変・時短での連チャンを含めて通常時何回転回せば 1 回当たるかの確率）、平均連チャン回数、等価ボーダー）から成るベクトルが得られた。

3. 自己組織化マップ

自己組織化マップとはコホネンが提案したアルゴリズムの一種で、以下の特徴がある。

- ・ どんな特徴ベクトルでも入力できる
- ・ 出力は入力と異なる次元でも構わない
- ・ 出力されたマスの周囲にはそれに類似したマスが集まる

本研究課題では、拡張パッケージ som を R にインストールし、som 関数を用いて分類を行った。som 関数のパラメータとして、以下の 3 つを指定した。

- ・ 縦方向のマップサイズ：8
- ・ 横方向のマップサイズ：8
- ・ マップの種類・・・hex(蜂の巣状のグラフ)

パチンコの機種データは、som 関数にかけの前に scale 関数

を用いて正規化を行った。

4. 結果

分類結果を図 1 に示す。自己組織化マップによる分類の結果、各マスに配置されたパチンコ機種の数は、n として示されている（n=0 の場合そのマスには機種の特徴（特徴次元）が該当するパチンコがないことを示している）。マスの中のグラフは、2 節で示した特徴を左から順に並べた時の各特徴の正規化された値を示している。例えば、右上のマ스에配置された 14 機種は確率変動率と平均連チャン回数の値が比較的高く、左下のマ스에配置された 9 機種は逆の傾向を示している。

図 1 から、中央付近には機種がほとんど配置されず、周辺部に集中して配置されていることがわかる。中央付近は 6 種類の特徴に関して平均的であり、ここに配置されるのは平凡で特徴の無い機種である。つまり、パチンコにおいてはそのような機種が非常に少ないことがわかる。一方、周辺部に配置された機種の傾向を見てみると、多くの場合に以下の 2 種類に大別されることがわかる。

- ・ よく当たるが稼げない機種：マップ右側
- ・ 当たりにくい稼げる機種：マップ左側

5. おわりに

R を使い SOM というグラフを作ってみて、このような結果が得られてよかった。統計学という講義もそれに類した授業もなかったので R を勉強でき R を通じて統計という分野を学ぶことができた。今後は R にさらにパッケージをインストールしてより複雑な統計を試みたい。

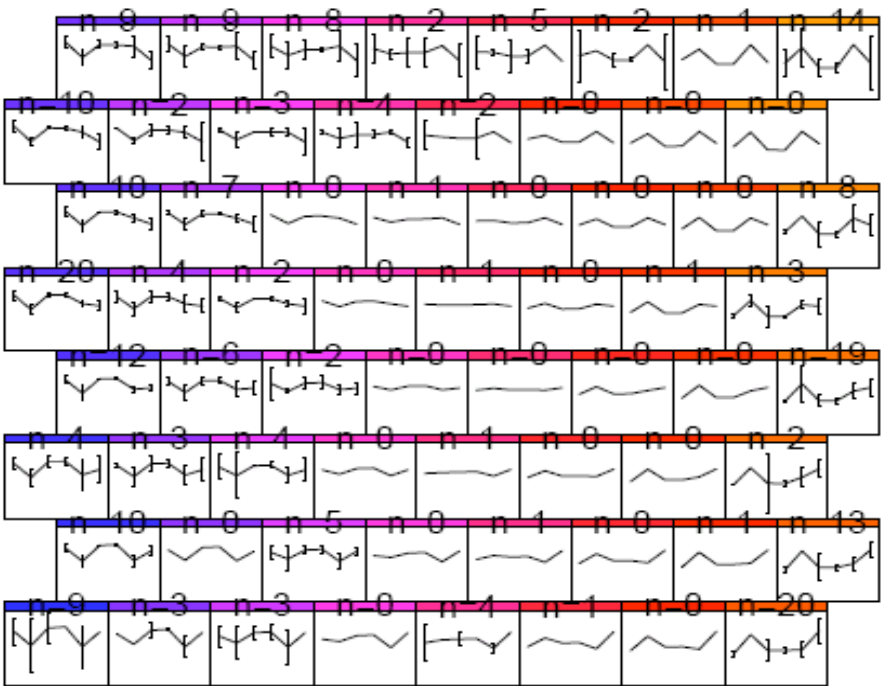


図 1．自己組織化マップによる分類結果