

# 悪性新生物の地域性に関する統計的分析

2010SE279 吉田里穂

指導教員：松田眞一

## 1 はじめに

1950年から悪性新生物、いわゆるがんが日本の死亡要因上位5位以内に登場し始め、1981年から現在まで、死亡要因第1位である。現在、がん＝不治の病ではないが、日本人の多くはがんで亡くなっている。本研究では、生活環境の相違が悪性新生物の罹患や発生部位にどのような影響を及ぼすかを明らかにし、その地域性を見ていく。

## 2 使用データについて

web[1]にある、悪性新生物部位別75歳未満年齢調整人口死亡率人口10万人対(2011年)から「全部位、食道、胃、肝臓、胆のう、膵、肺、膀胱、悪性リンパ腫、白血病、大腸」の男女合計がん死亡率データを入手し、これを目的変数とする。また、説明変数はがんの発生に関係がありそうな変数(酒消費量や行動率など)を44個用意した。全て、量的変数である。使用した都道府県別データはweb[1][2][3][4][5][6][7]から入手した。

## 3 解析方法について

本研究では、分析手法を3つ用いた。生活環境と悪性新生物の相関関係を調べるため、重回帰分析、正準相関分析を行い、地域性を見るため、クラスター分析を行った。クラスター分析は、鎖効果が見られなかったワード法を用いた。重回帰分析では修正決定係数を0.7以上、正準相関分析では0.8以上を目安とし、第5正準変量まで扱った。(大野[9]、金[10]参照)

## 4 重回帰分析

AIC(赤池情報量基準)値を参考に、減少法・増減法・増加法・増加法を組合せて変数選択を行った。また、VIF(分散拡大要因)を用いて、選ばれた変数すべてのVIF値<10となるまで変数選択を繰り返し、満たされたとき多重共線性が回避されたと思なした。一部の分析結果のみを示す。

### 全部位(修正決定係数:0.736)

男女喫煙率、しょう油がプラスに、みそがマイナスに強く反応した。がんの発生要因は、たばこ塩分の影響が大きいことが確認できた。みそは多くの塩分を含むが、渡邊[8]より、新鮮なみそに限って、みその塩分濃度にかかわらず、がん抑制効果があるため、マイナスとなった。みそ消費量やみその蔵元が多い長野県は、がん死亡率が低い。

### 肝臓(修正決定係数:0.8895)

沖縄県はがん死亡率は低い、一部部位の死亡率が高い点が他県と異なるため、外れ値とした。みそがマイナス、酒や塩分に関する変数の多くがプラスとなった。ここでも、みその抑制効果が確認できた。また、酒のつまみに味の濃いものが多いためか、調味料が多く反応した。酒に繋がる要因はプラス、他はマイナスという構造が見られる。

### 白血病(修正決定係数:0.8301)

単式焼酎(本格焼酎や泡盛)のみプラスに強く反応した。ゆえに、白血病は食や生活環境の影響は受けにくい、アルコール度数が非常に高い酒の影響は受けると分かった。

## 5 正準相関分析

重回帰分析と異なり、有効な変数選択法が存在しない。ゆえに、VIFを用いてVIF値が大きい変数から削り、全ての変数のVIF値<4となった組合せを採用した。また、このとき、多重共線性が回避されたものと見なした。

### 第1正準変数:「白血病に関する軸」

(第1正準相関係数:0.9653)

表1 第1正準変数の結果(一部削除)

第1群		第2群	
単式焼酎	0.537	食道	0.016
ビール	-0.181	胃	-0.297
ブランデー	-0.255	肝臓	-0.384
喫煙率男	-0.316	胆のう	-0.196
旅行	-0.197	膵	0.001
教養娯楽費	0.285	気管と肺	-0.133
食塩	-0.118	膀胱	0.169
しょう油	-0.252	悪性リンパ腫	0.238
みそ	0.745	白血病	0.746
酢	-0.266	大腸	-0.233
マヨネーズ	-0.224		

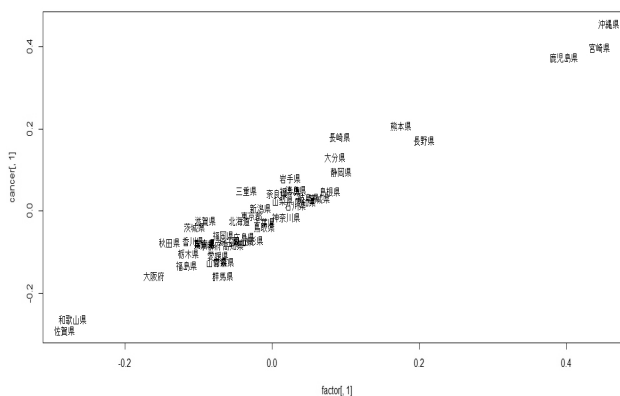


図1 第1正準変数の散布図:都道府県の値

表1より、第1群は単式焼酎とみそ、第2群は白血病がプラスに強く反応した。また、絶対値が他と比べ大きく、3変数とその他という構図が見える。図1を見ると、沖縄、宮崎、鹿児島が右上で孤立し、九州以南に地域性が見られる。この3県は、単式焼酎消費量トップ3、白血病死亡率ワースト3位である。これは重回帰分析の結果と一致する。一方、みそは白血病以外のがんには強い抑制効果を示すため、相対的に白血病と同じ向きになったと考えられる。

## 第2 正準変数：「肝臓がんに関する軸」

(第2 正準相関係数：0.9206)

第2 群は肝臓，第1 群は単式焼酎，ブランデー，喫煙率男，みそが強く反応した。みそは正，他の4 つは負となった。重回帰分析より，みそは肝臓がんを抑制し，酒とたばこは発生要因となることが分かっている。また，正準変数の散布図では，酒の消費量が多い九州や東北，肝臓がんの死亡率が高い県が図の左側に集中していた。このことから，肝臓とアルコールの相関が確認できた。

## 第3 正準変数：「血液がんに関する軸」

(第3 正準相関係数：0.8928)

第2 群は白血病が正，悪性リンパ腫が負に強く反応した。また，絶対値が0.7 前後と大きいため，これらを分ける軸と予想した。さて，第1 群はみそが負に強く反応した。重回帰分析から，悪性リンパ腫の方がみそに反応し，そして悪影響を受けることが分かっている。また，白血病は単式焼酎だけ強く反応した。これは，正準相関分析結果と一致する。ゆえに，血液がんに関する軸であると結論付けた。

## 第4 正準変数：「消化器官と排泄器官に関する軸」

(第4 正準相関係数：0.8488)

大腸や悪性リンパ腫が正，胃が負に強く反応した。絶対値は小さいが，排泄器官が正，消化器官が負という形になった。また，第1 群は喫煙率が正，しょう油が負となった。散布図を見ると，大腸がんや肺がんが多い地域が正に位置しているため，消化器官と排泄器官に関する軸とした。

## 第5 正準変数：「油分に関する軸」

(第5 正準相関係数：0.8192)

第2 群は大腸，胆のう，胃が負，それ以外が正となり，消化系で溜める器官が否かで違いが出た。また，第1 群は，酒が正，油が負となった。油は消化吸収に時間がかかり，消化されないこともある。ゆえに，食物が長時間とどまる器官は影響を受けやすいと考え，油分に関する軸とした。

## 6 クラスタ分析

### 第1 群：「癌の死亡率が高い群」

(北海道，青森，秋田，大阪，和歌山，鳥取，佐賀)

がん死亡率上位7 位が集まった。部位別死亡率も高い。また，塩分摂取量が多い地域で，塩またはしょう油の購入量が他県と比べて多い。伝統的に塩分が多い保存食を有する地域や，海に面した地域が多く，日常的に塩分が多い食事をしている地域だと考えられる。

### 第2 群：「平均的な群」

(宮城，山形，福島，茨城，栃木，埼玉，東京，神奈川，富山，愛知，京都，兵庫)

人口が100 万人超の大都市がある都道府県やその周りの県が集まった。大都市圏は様々な地域から人が集まりやすいため，平均的な値となったと考えられる。特に，関東地域はその傾向が強く見られる。

### 第3 群：「男性の全部位の死亡率が高い群」

(岩手，山口，高知，福岡，長崎)

第1 群の次に死亡率が高い。しかし，第1 群は男女と

もに死亡率が高いが，第3 群は男性死亡率が高い点で異なる。例えば，大腸がんは女性に多いがんであるが，第3 群は男性に多い傾向にある。

### 第4 群：「がん死亡率は低いが，消化腺は弱い群」

(福井，山梨，長野，滋賀，岡山，香川，熊本，大分)

他の群に比べ，がん死亡率が低い。しかし，消化腺（肝臓と膵臓）に着目すると，他の群と差はなく，むしろ（算術）平均よりも高い。中部地方など酒に弱い地域が集まっているためだと考えられる。熊本，大分は酒に強いが，強い酒の飲酒量も多い。肝臓の負担は大きい，分解能力が高いため，体全体への影響は小さいのだと推測される。

### 第5 群：「平均よりがん死亡率が低い群」

(群馬，千葉，新潟，石川，岐阜，静岡，三重，奈良，島根，広島，徳島，愛媛，宮崎，鹿児島，沖縄)

がん死亡率で（算術）平均，中央値を共に下回る。しかし，部位によって，死亡率がワースト3 位に入る県が集まっている点に特徴がある。

## 7 まとめ

一般的に言われているように，酒，たばこ，塩はがんの発生要因になっていることが確認できた。特に，アルコール度数が高い酒はがんへの影響が大きい傾向にあった。また，臓器がんに対し，みその抑制効果が確認された。しかし，血液がんには効果確認できず，中でも白血病は生活環境の影響を受けにくいという点で，他のがんとは違う性格をもつようである。

## 8 おわりに

がんの罹患には，遺伝子要因が大きいと思っていたが，食生活，特に飲酒の影響が大きいということが分かった。つまり，後天的な要因の方が影響を与えるということが確認できた。しかし，最近ではウイルスとがんの因果関係が示されており，ウイルスの感染率等のデータがあれば，面白い結果が得られたらと思う。

## 参考文献

- [1] がん情報サービス <http://ganjoho.jp>
- [2] 国税庁：平成23年，都道府県別の酒類販売数量
- [3] 総務省統計局：「日本の統計2013」，第2章人口・世帯
- [4] 総務省：「家計調査」，2011年
- [5] e-Stat 政府統計の総合窓口：「社会生活統計指標—都道府県の指標—2013」
- [6] 一般財団法人 自動車検査登録情報協会
- [7] 『日本経済新聞』2010年7月2日電子版「“酒豪”どこに多い？「全国酒豪マップの謎」」<http://www.nikkei.com/article/DGXBZO10187530R00C10A7000000/>
- [8] 渡邊敦光：『お味噌の効能』。日本醸造協会誌，Vol.105，no.11，pp714-723（20101115）
- [9] 大野高裕：『多変量解析入門：自由自在に使いこなすコツ』。同友館，1998
- [10] 金明哲：『Rによるデータサイエンス2：多次元データ解析法』。共立出版，2009