

都道府県別生命表掲載データからの作図法 (How to make charts by prefecture data)

最終更新 : Dec. 24, 2022 (Sat)

このページは、[鵜記](#)という日常メモに書いた内容から、都道府県別生命表に掲載されている数値データをわかりやすい図にする方法を抜粋してまとめたものである（2022年12月に大幅更新）。目標は、都道府県別平均寿命の推移を示す折れ線グラフと、死因別損失余命の都道府県別プロファイルを示すレーダーチャートを、男女別に作成することである。

用意するソフトウェア

1. **Excel形式のデータを読んで編集できるソフト** : Microsoft Excelの他、フリーソフトのLibreOfficeでも良い。ただし、このページからRで読めるように加工済みのタブ区切りテキストデータをダウンロードする場合は不要。
2. **R** : 詳細は[こちら](#)を参照。
3. **R用のfmsbライブラリ** : 詳細は[こちら](#)を参照。Rへのインストールは、Rを起動してから、`install.packages("fmsb")`でOK。ただし、長いラベルを折り返してつけるために、バージョン0.3.8以上のものでなくてはならない。

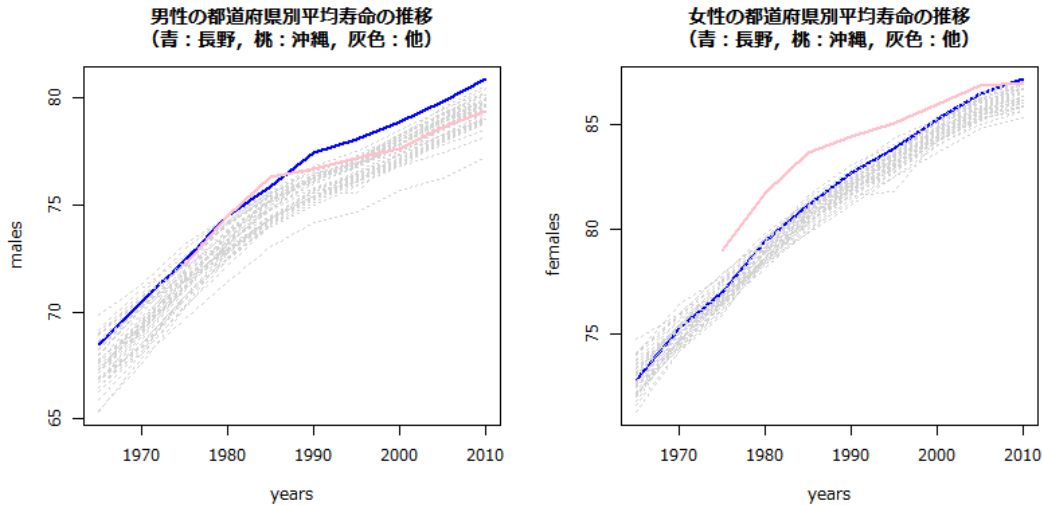
データの入手先

1. 厚生労働省のサイトで2013年2月28日に公開された、[平成22年都道府県別生命表の概況](#) (<http://www.mhlw.go.jp/toukei/saikin/hw/life/tdfk10/index.html>)の「図表データのダウンロード」からExcelファイル (<http://www.mhlw.go.jp/toukei/saikin/hw/life/tdfk10/dl/zuhyou.xls>)をダウンロードする（平成27年＝2015年データは[ここ](#)から、令和2年＝2020年データは[ここ](#)から入手できる）。
2. Microsoft ExcelかLibreOfficeのCalcでファイルを読み込み、都道府県別平均寿命の年次変化データが入っている「統計表 表2-1」と「統計表 表2-2」を新しいシートに横に並ぶようにコピー＆ペーストし、余計な行や列を削除し、最上行にアルファベットの変数名を付ける。Rで読み込む時には、全体をコピーした状態で、`read.delim("clipboard")`でも良いのだが、これをタブ区切りテキスト形式ファイルとして保存したものが[pref-e0-changes.txt](#)である。都道府県名の2バイトコードがCP932（いわゆるShift-JIS）なので、Windows以外のOSで読み込む時は適宜コード変換する方が良いかもしれない……と書いていたが、2020年までのデータを追加し、都道府県名をローマ字表記にした[Prefe0.txt](#)もダウンロードできるようにした。
3. 続いて、主な死因別の損失余命（この表では、「特定死因を除去した場合の平均寿命の延び」と書かれている）が都道府県別に入っている「参考2-1」、「参考2-1（続き）」、「参考2-2」、「参考2-2（続き）」の4枚のシートの内容を、新しいシートに横に並ぶようにコピー＆ペーストし、余計な行や列を削除し、最上行にアルファベットの変数名を付ける。これをタブ区切りテキスト形式ファイルとして保存したものが[pref-LLY-h22.txt](#)である。2バイトコードについての注意は同上……と書いていたが、都道府県名表記をローマ字にしたものを[PrefYLL2010.txt](#)としてダウンロード可能にした。また、令和2（2020）年版も[PrefYLL2020.txt](#)としてダウンロード可能にした。

折れ線グラフの作成

[pref-e0-changes.txt](#)をRの作業ディレクトリに置いて以下を実行すると、その下のグラフが画面に表示される（Windows環境で、直接pngファイルを出力させるコード[pngのサイズ指定や、フォントをメイリオに指定したもの]は、[change-e0.R](#)である。下に貼り付けたpngファイルはそうやって作ったもの）。

```
x <- read.delim("pref-e0-changes.txt")
males <- t(x[, 2:11])
colnames(males) <- x$PREF
females <- t(x[, 12:21])
colnames(females) <- x$PREF
COL <- ifelse(x$PREF=="長野", "blue", ifelse(x$PREF=="沖縄", "pink", "lightgrey"))
LWD <- ifelse(x$PREF=="長野", 2, ifelse(x$PREF=="沖縄", 2, 1))
LTY <- ifelse(x$PREF=="長野", 1, ifelse(x$PREF=="沖縄", 1, 3))
years <- 1965+0:9*5
layout(t(1:2))
matplot(years, males, type="l", col=COL, lwd=LWD, lty=LTY,
  main="男性の都道府県別平均寿命の推移¥n（青：長野、桃：沖縄、灰色：他）")
matplot(years, females, type="l", col=COL, lwd=LWD, lty=LTY,
  main="女性の都道府県別平均寿命の推移¥n（青：長野、桃：沖縄、灰色：他）")
```

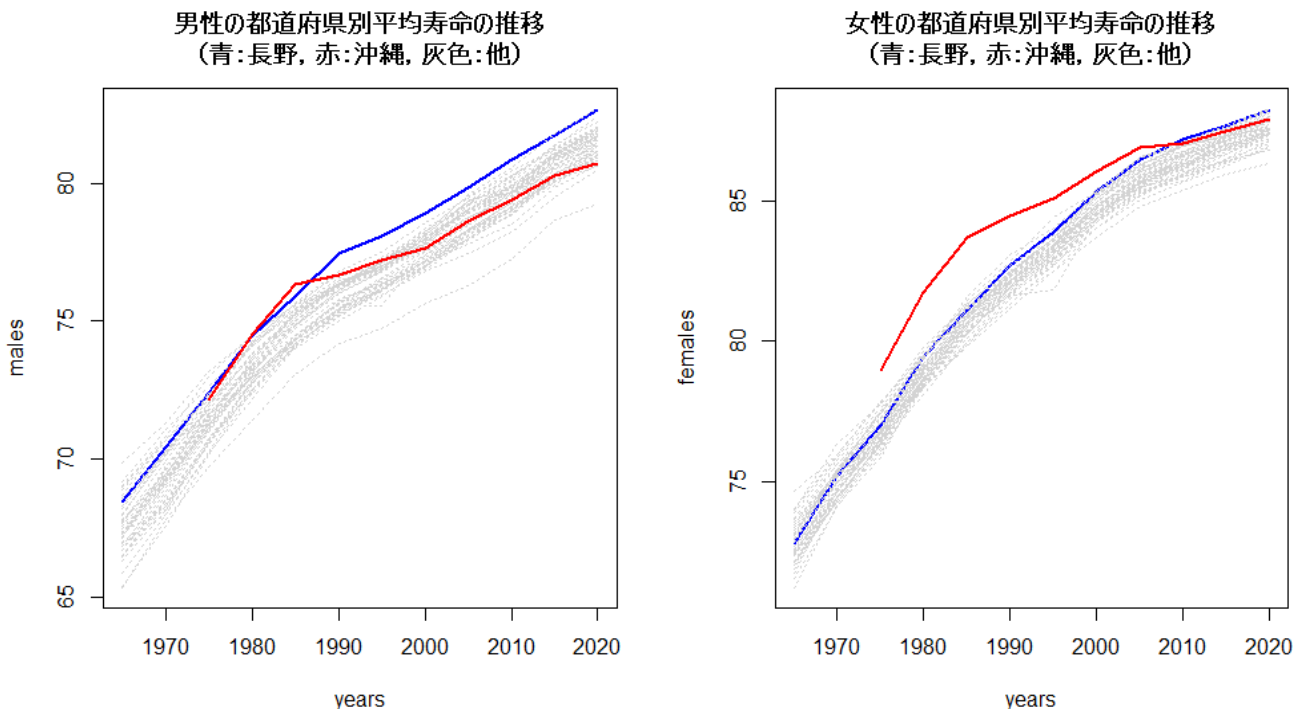


なお、ファイル読み込みの際に、`read.delim("pref-e0-changes.txt", fileEncoding="cp932")`のように、`fileEncoding`オプションを使えば、Windows以外のOSでもCP932のファイルを正しく読み込むことができる (Special thanks to: [裏RjpWikiさん](#)のご指摘)。

このグラフから読み取れることはそれほど多くないが、1985年までトップレベルだった沖縄男性の平均寿命が、1990年から急に伸びが鈍化したこと、長野県男性も1990年までの伸びに比べると1995年以降は伸びが鈍化していることがわかる。女性については、男性と違って、最近まで沖縄の平均寿命の高さは他都道府県とは段違いだったのに、2005年に追いつかれ、2005年から2010年には横這いになってしまったことが一目で分かる。数値だけ眺めるよりわかりやすいと思う。

ちなみに、これは折れ線グラフなので、縦軸がゼロから始まっていないことに注意されたい。2010年の男性の水準には、女性は1980年頃には既に到達していた。

2020年までのデータを使って描き直すコード ([pref-e0-changes-until2020.R](#)) の出力は下図。

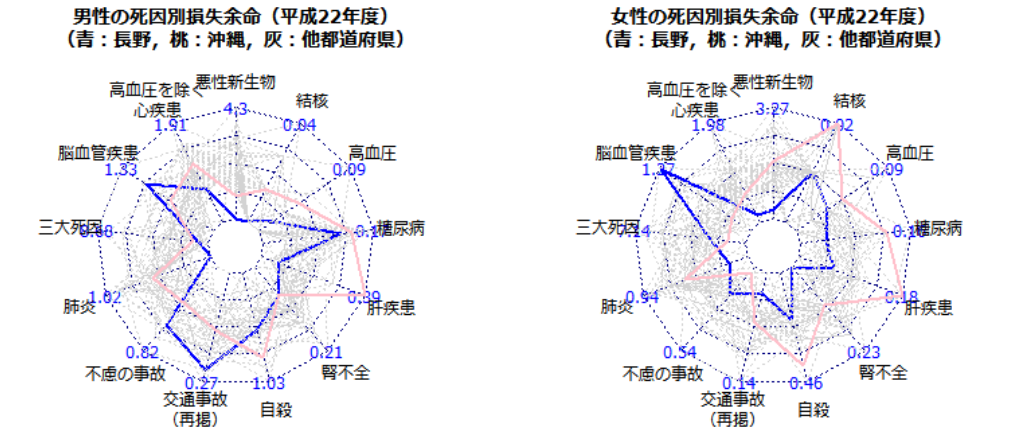


レーダーチャートの作成

[pref-LLY-h22.txt](#)をRの作業ディレクトリに置いて以下を実行すると、その下のグラフが画面に表示される (Windows環境で、直接pngファイルを出力させるコード[pngのサイズ指定や、フォントをメイリオに指定したもの]は、[LLY-profiles.R](#)である。下に貼り付けたpngファイルはそうやって作ったもの。最新のRを使う際は日本語文字コードに注意)。

```
x <- read.delim("./pref-LLY-h22.txt", fileEncoding="cp932")
COL <- ifelse(x$PREF=="長野", "blue", ifelse(x$PREF=="沖縄", "pink", "lightgrey"))
LWD <- ifelse(x$PREF=="長野", 2, ifelse(x$PREF=="沖縄", 2, 1))
LTY <- ifelse(x$PREF=="長野", 1, ifelse(x$PREF=="沖縄", 1, 3))
VX <- c("悪性新生物", "高血圧を除く心疾患", "脳血管疾患", "三大死因",
        "肺炎", "不慮の事故", "交通事故¥n(再掲)", "自殺", "腎不全", "肝疾患",
        "糖尿病", "高血圧", "結核")
males <- x[, 2:14]
females <- x[, 15:27]
```

```
require(fmsb)
layout(t(1:2))
radarchart(males, maxmin=FALSE, pcol=COL, axistype=2, pty=32, plty=LTY, plwd=LWD, vlabels=VX,
  title="男性の死因別損失余命（平成22年度）＼n（青：長野，桃：沖縄，灰：他都道府県）")
radarchart(females, maxmin=FALSE, pcol=COL, axistype=2, pty=32, plty=LTY, plwd=LWD, vlabels=VX,
  title="女性の死因別損失余命（平成22年度）＼n（青：長野，桃：沖縄，灰：他都道府県）")
```



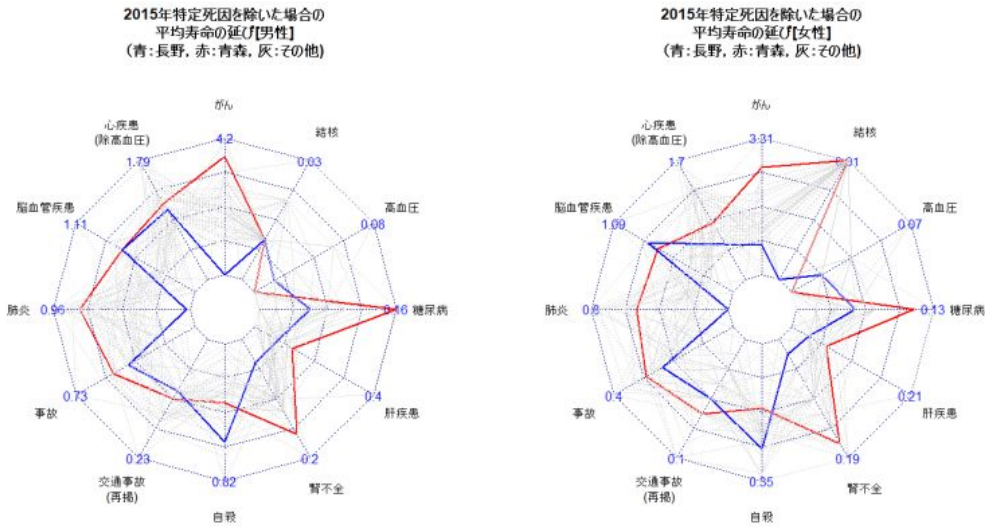
このグラフはいろいろなことを示唆してくれる。一見してわかることは、平均寿命が男女とも最長の長野県は、男女とも、がんと肺炎による死亡が少ないということだ。一方、脳血管疾患によって失われている余命は比較的大きい。これは、長野県の人は漬け物をよく食べるため、元々塩分摂取量が多く、そのために脳卒中が多かったのを、食生活改善推進員さんが歩き回って塩分摂取量を減らし、そのおかげで脳卒中が減ったと言われているのだが、それでもまだ塩分摂取が高いということかもしれない。ただし、くも膜下出血のリスク因子としては遺伝も大きいので、塩分摂取だけが問題とは言い切れないが。なお、長野県では、男性のみ交通事故によって失われている余命が大きい、これは子供の交通事故死だと思われる。細くて見通しが悪くて歩道が狭い道路が多いのに外遊びする子供は多いので、飛び出しによる交通事故が比較的多いのであろうことは想像に難くない。沖縄のプロファイルから目立つのは、肝疾患、糖尿病が高いことだ。たぶん飲酒が多いせいだろう。女性のみ結核による損失余命が大きかったが、これは流行があったのかもかもしれない。

たぶん、ここで書いた推測の裏付けを取るようなデータを探して分析すれば、地域関連研究だけれども、卒論が学会発表ネタくらいにはなるかもしれない。

2015年のデータで、長野と青森を強調して同様のグラフを描くコードは、[prefyll2015-nagano-aomori.R](#)としてダウンロードできる。PrefYLL2015はfmsbパッケージに入れたので、fmsbパッケージをインストールしてあれば、いちいちデータを読み込む必要はない。

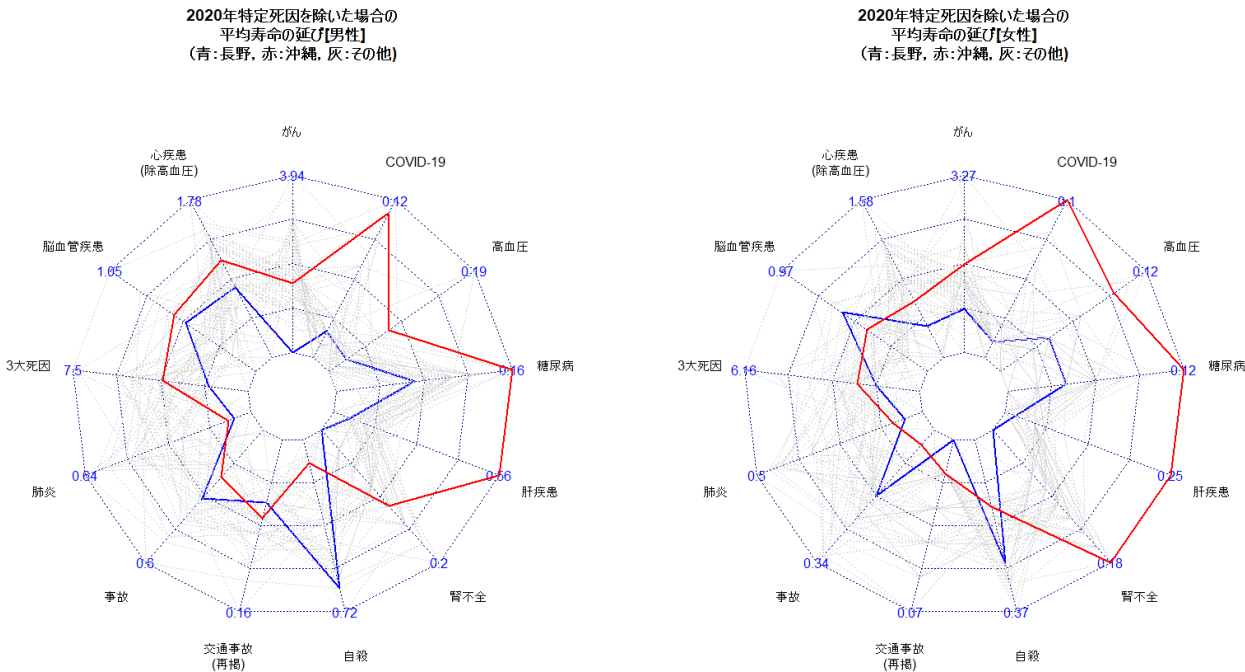
```
library(fmsb)
data(PrefYLL2015)
COL <- ifelse(PrefYLL2015$PNAME=="Nagano", "blue",
  ifelse(PrefYLL2015$PNAME=="Aomori", "red", "lightgrey"))
LTY <- ifelse(PrefYLL2015$PNAME=="Nagano", 1,
  ifelse(PrefYLL2015$PNAME=="Aomori", 1, 3))
LWD <- ifelse(PrefYLL2015$PNAME=="Nagano", 2,
  ifelse(PrefYLL2015$PNAME=="Aomori", 2, 1))
VLAB <- c("がん","心疾患＼n(除高血圧)","脳血管疾患",
  "肺炎","事故","交通事故＼n(再掲)","自殺","腎不全",
  "肝疾患","糖尿病","高血圧","結核")
males <- PrefYLL2015[, 3:14]
females <- PrefYLL2015[, 15:26]
layout(t(1:2))
radarchart(males, maxmin=FALSE, pcol=COL, axistype=2, pty=32,
  plty=LTY, plwd=LWD, vlabel=VLAB,
  title="2015年特定死因を除いた場合の＼n平均寿命の延び[男性]
  (青：長野，赤：青森，灰：その他)")
radarchart(females, maxmin=FALSE, pcol=COL, axistype=2, pty=32,
  plty=LTY, plwd=LWD, vlabel=VLAB,
  title="2015年特定死因を除いた場合の＼n平均寿命の延び[女性]
  (青：長野，赤：青森，灰：その他)")
```

実行すると下図が描かれる。



男女とも青森県では、がん、肺炎、腎不全、糖尿病による余命損失が大きいことがわかる。

2020年のデータもfmsbに入れる予定だが、サーバ上にアップロード済みなので、[このコード](#)を実行すれば下図が得られる。



Correspondence to: minato-nakazawa[atmark]umin.net.

[リンクと引用について](#)