

# サザエさんのじゃんけん

## R言語によるデータ分析入門

やじゅ@静岡Developers勉強会

# はじめに

サザエさんのじゃんけんは、1991年10月20日の放送から始まりました。初回の手は「グー」でした。

今現在でも続いており、2013年12月の段階で1150回以上のじゃんけんのデータになっております。

ちなみに、実写版のサザエさんは毎回「チョキ」です。

過去の手一覧や次の手の予測しているサイトを参考に  
今回R言語を使ってデータを分析してみます。

[サザエさんのジャンケン学](#)

[サザエさんじゃんけん研究所](#)

# 自己紹介／資料置場

▶ やじゅ@静岡・・・漢字名は「八寿」

平口 八寿人(<https://www.facebook.com/yasuhito.hiraguchi>)

アラフォーエンジニア、元M S M V P (Visual Basic)

静岡県島田市のS L (大井川鉄道)が走っている所に在住

Twitter : [yajū](#)

<http://blogs.wankuma.com/yaju/>

<http://yajū3d.hatenablog.jp/> (数学と物理を基礎からやり直す)

▶ 資料場所

[https://github.com/yaju/Sazae\\_R](https://github.com/yaju/Sazae_R)

# R言語について

## ▶ R言語って何？

誰でも無料で、自由に利用できるデータ分析のための言語です。  
どんなOS（Windows, Macintosh, Linux）でも動作します。

利用可能な分析手法やツールを、世界中のユーザーが開発し公開しているため、  
極めて豊富です。

ダウンロード（Windows版）はこちらから  
<http://cran.md.tsukuba.ac.jp/bin/windows/base/>

## ▶ R Studioって何？

統計ソフトR用の統合開発環境(IDE) です。  
Rをもっと 便利に使うためのフリーソフト

## ▶ R Studioのインストール

ダウンロードはこちらから <http://www.rstudio.com/ide/>  
使い方は、G o o g l e 先生に尋ねてください。

# アジェンダ

- ▶じゃんけんデータの生成
- ▶じゃんけんデータのグラフ化
- ▶じゃんけんの癖
- ▶人間乱数の分析
- ▶次の手の予測と勝敗結果
- ▶まとめ

## ■ じゃんけんデータの生成

- ▶ Webサイトからサザエさんの過去の手の一覧を取得する

RCurl(HTTP リクエストインターフェース )を使うことで、Webサイトから随時データを取得します。

### サザエさんの過去の手の一覧

第1153回	13.12.01	チョキ (観月ありさ)
第1152回	13.11.24	グー
第1151回	13.11.17	パー
第1150回	13.11.10	チョキ
第1149回	13.11.03	チョキ
第1148回	13.10.27	パー
第1147回	13.10.20	グー
第1146回	13.10.13	グー
第1145回	13.10.06	チョキ
第1144回	13.09.29	パー
第1143回	13.09.22	パー
第1142回	13.09.15	チョキ
第1141回	13.09.08	グー
第1140回	13.09.01	チョキ
第1139回	13.08.25	パー
第1138回	13.08.18	休み

1. 文字コードはShift-JIS
2. テーブルタグは未使用
3. 第7回 91.11.31と  
ありえない日付がある

# Webスクレイピング ソースリスト

```
library(RCurl)

txtfile <- iconv(getURL("http://www.asahi-net.or.jp/~tk7m-
ari/sazae_ichiran.html", .encoding="Shift-JIS"), "Shift-JIS","UTF-8")

txtvec <- strsplit(txtfile,'\\n')[[1]]

flg <- FALSE

i <- 1

while(i > 0) {
  tag <- substr(txtvec[i],1,1)
  if(tag == '<' || tag == "){ i <- i + 1; next }
  str <- strsplit(txtvec[i],'\\t')[[1]]
  tmp <- regexpr('\\\\d+',str[1])
  seq <- as.integer(substr(str[1],tmp,tmp+attr(tmp,'match.length')-1))
```

```
  if(seq == 7){ str[2] <- '91.12.1'}
  dt <- as.Date(str[2],format='%y.%m.%d')
  tmp2 <- regexpr('<| (' ,paste(str[3],'<'))
  str2 <- substr(str[3],0,tmp2-1)
  idx <- 9
  kind <- '-'
  if(str2 == 'グー') { kind <- 'G'; idx=1 }
  if(str2 == 'チヨキ'){ kind <- 'C'; idx=2 }
  if(str2 == 'パー'){ kind <- 'P'; ; idx=3 }
  if(flg == FALSE){
    tbl <- data.frame(seq,dt,kind,idx,yy,mm,dd,)
    flg <- TRUE
  }
  else{
    tbl <- rbind(xy,data.frame(seq,dt,kind,idx))
  }
  #print(seq)
  i <- i + 1
  if(seq == 1) break
}
```

# じゃんけんデータの説明

- ▶ Webスクレイピングにより生成されたデータフレーム

	seq	dt	kind	idx
1	1153	2013-12-01	C	2
2	1152	2013-11-24	G	1
3	1151	2013-11-17	P	3
4	1150	2013-11-10	C	2
5	1149	2013-11-03	C	2
6	1148	2013-10-27	P	3
7	1147	2013-10-20	G	1
8	1146	2013-10-13	G	1
9	1145	2013-10-06	C	2
10	1144	2013-09-29	P	3

列名

seq 回数

dt 日付

kind 種類

グー:G、チョキ:C、パー:P、休み:-

Idx 表示順

グー:1、チョキ:2、パー:3、休み:9



# ■ じゃんけんデータのグラフ化

## ▶ データ分析のグラフ化

サザエさんのジャンケン学のサイトでは、年毎、回の1の位毎、月毎及び週毎のサザエさんの過去の手の分布を分析しています。

[http://www.asahi-net.or.jp/~tk7m-ari/sazae\\_analysis\\_distribution.html](http://www.asahi-net.or.jp/~tk7m-ari/sazae_analysis_distribution.html)

## ▶ データフレームをSQLで操る「sqldf」の活用

RでSQLで行うようなデータ加工・集計を行うには、基本的にsubsetやorder、merge、aggregateといった関数を利用します。

SQLとRの関係については、下記サイトがわかりやすいです。

[http://d.hatena.ne.jp/a\\_bicky/20110529/1306667230](http://d.hatena.ne.jp/a_bicky/20110529/1306667230)

普段データのデータ加工や集計作業にSQLを使用している側にとっては、SQLでクエリを扱える方がとっつきやすい。

「sqldf」パッケージを用いれば、直接SQLを利用してデータを扱うことが可能になります。

# じゃんけんの種類別総数グラフ化の準備①

## ▶ じゃんけんの種類別総数SQL

```
sumtbl <- sqldf("SELECT kind,COUNT(kind) cnt,idx FROM tbl GROUP BY kind ORDER BY idx")
```

	kind	cnt	idx
1	G	359	1
2	C	387	2
3	P	370	3
4	-	38	9

8

## ▶ 棒グラフ化が面倒

R標準の棒グラフ「barplot」では、横軸(X)と縦軸(Y)のデータ列の指定が出来ないため、下記のように「barplot」に則ったデータに加工する必要があって面倒である。

	G	C	P	-
1	359	387	370	38

## じゃんけんの種類別総数グラフ化の準備②

### ▶ ggplot2による棒グラフ作成

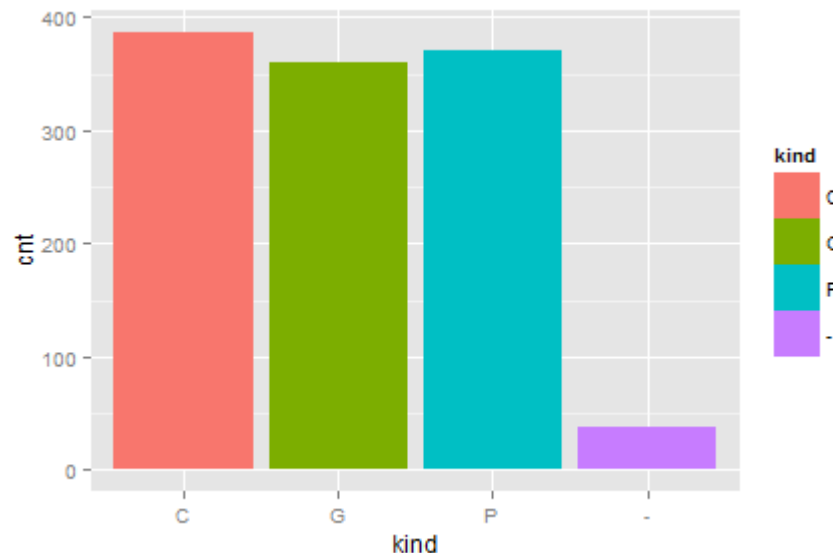
「ggplot2」では、横軸(X)と縦軸(Y)のデータ列の指定が出来るので、データフレームのまま扱うことが出来る。

```
library(ggplot2)
```

```
p <- ggplot(sumtbl, aes(x=kind, y=cnt))
```

```
p + geom_bar(stat="identity", aes(fill=kind))
```

	kind	cnt	idx
1	G	359	1
2	C	387	2
3	P	370	3
4	-	38	9



でも、標準指定では横軸がC,G,P,-と順番が勝手に並び替えられてしまっているので、G,C,P, -の順番にしたい。また、横軸(kind)と縦軸(cnt)と色名(kind)を日本語にしてみたい

## じゃんけんの種類別総数グラフ化の準備③

- ▶ reorder関数を使って横軸(X)の順序を指定する

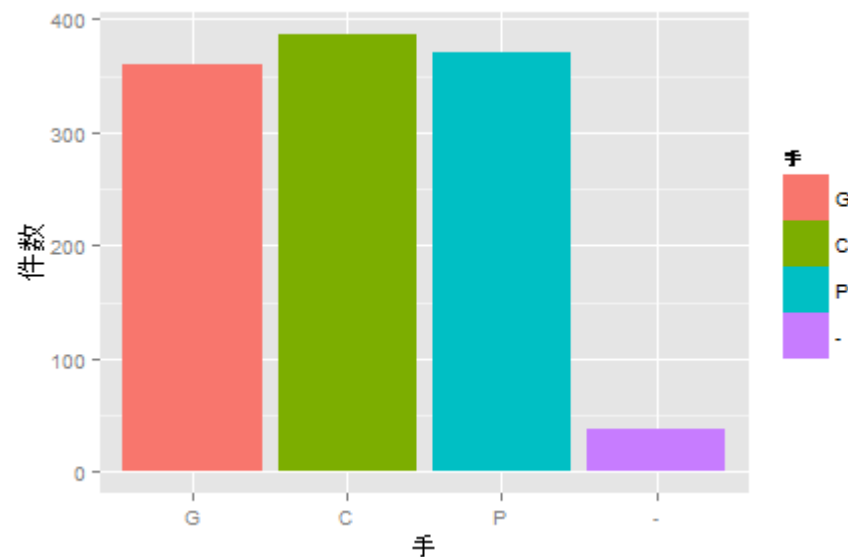
データフレームに横軸(X)の順序を規定する変数を含めておき、aes コンポーネントで reorder 関数を使うだけで好きな順序に変更することができます。

また、xlab,ylab,labs(fill)に横軸(X)と縦軸(Y)と色名に日本語を指定する。

```
p <- ggplot(sumtbl, aes(x=reorder(kind, idx), y=cnt))
```

```
p + geom_bar(stat="identity", aes(fill=reorder(kind, idx))) + xlab("手") + ylab("件数") + labs(fill = "手")
```

	kind	cnt	idx
1	G	359	1
2	C	387	2
3	P	370	3
4	-	38	9

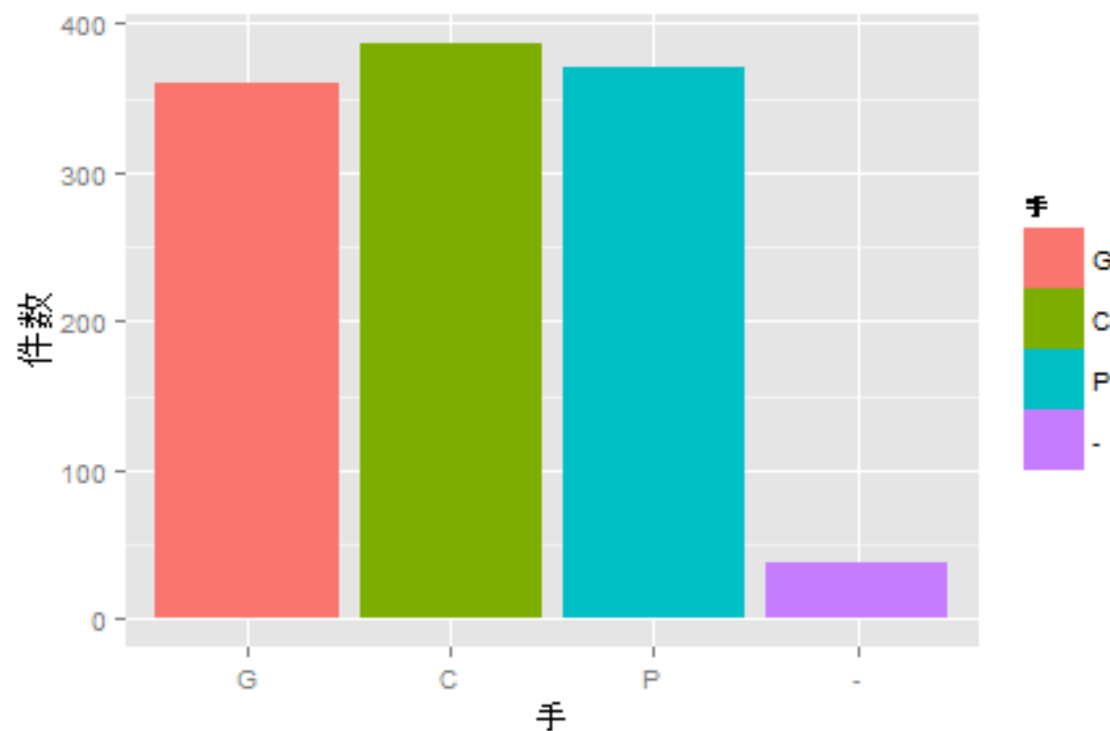


# じゃんけんの種類別総数による分析

## ▶ 分析結果

約20年のじゃんけんの手の総数の割にはあまり片寄りがなく均衡した結果となっている。  
毎年、各手の総数がバランスよく振り分けられていると思われる。

	kind	cnt	idx
1	G	359	1
2	C	387	2
3	P	370	3
4	-	38	9



# じゃんけんの年別種別総数グラフ化の準備①

- ▶ 10年前(2003年以降)のデータの年別種別総数のSQL

#日付を年のみに変換する関数を生成

```
year = function(data){ data$dt <- format(data$dt, "%Y"); data }
```

#関数を使って日付を年のみにしたデータフレームを生成

```
ytbl <- sqldf("SELECT * FROM tbl", method = function(x)  
  data.frame(year(tbl)))
```

# 10年前(2004年以降)から休みを除いた年別種別総数のデータフレームを生成

```
yeartbl <- sqldf("SELECT kind,dt,COUNT(kind) cnt,idx FROM yeartbl  
WHERE kind<>'-' AND dt>=2004  
GROUP BY kind,dt  
ORDER BY dt,kind")
```

# じゃんけんの年別種別総数グラフ化の準備②

- ▶ 10年前(2004年以降)から休みを除いた年別種別総数のデータフレーム

	kind	dt	cnt	idx
1	C	2004	16	2
2	G	2004	18	1
3	P	2004	15	3
4	C	2005	18	2
5	G	2005	14	1
6	P	2005	17	3
7	C	2006	18	2
8	G	2006	15	1
9	P	2006	17	3
10	C	2007	18	2
11	G	2007	19	1
12	P	2007	14	3
13	C	2008	21	2
14	G	2008	14	1
15	P	2008	16	3

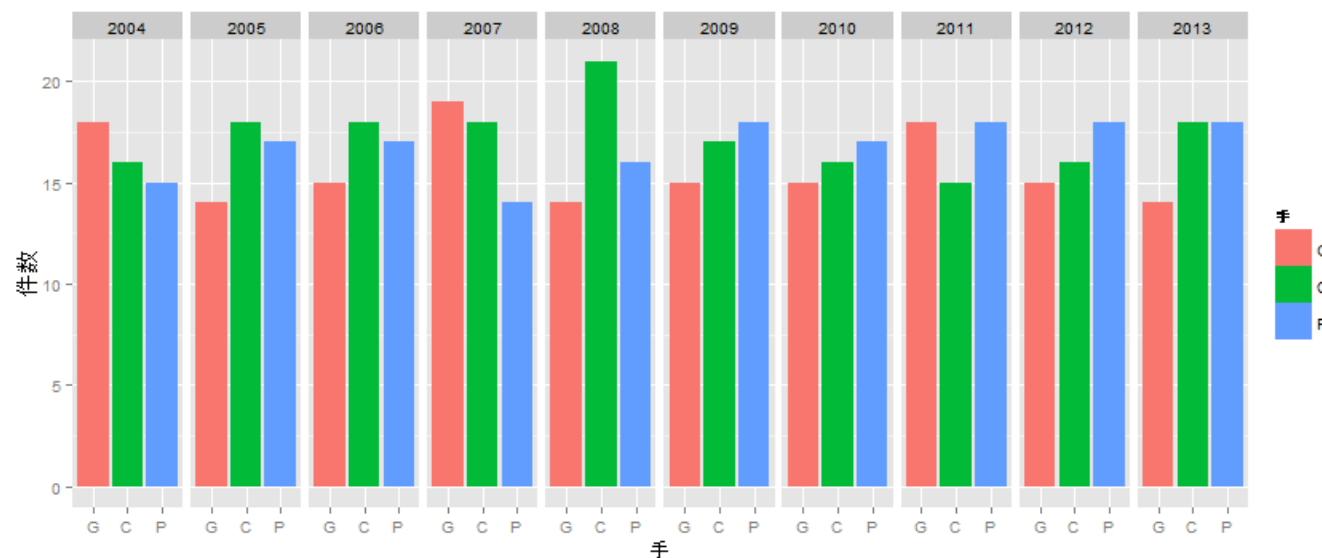
	kind	dt	cnt	idx
16	C	2009	17	2
17	G	2009	15	1
18	P	2009	18	3
19	C	2010	16	2
20	G	2010	15	1
21	P	2010	17	3
22	C	2011	15	2
23	G	2011	18	1
24	P	2011	18	3
25	C	2012	16	2
26	G	2012	15	1
27	P	2012	18	3
28	C	2013	18	2
29	G	2013	14	1
30	P	2013	18	3

## じゃんけんの年別種類別総数グラフ化の準備③

- ▶ 年別にグラフ化(facet\_grid関数を使用し横並びにする)

```
p <- ggplot(yeartbl, aes(x=reorder(kind, idx), y=cnt))
```

```
p + geom_bar(stat="identity", aes(fill=reorder(kind, idx))) + facet_grid(. ~ dt)+  
xlab("手") + ylab("件数") + labs(fill = "手")
```





# じゃんけんの年別種類別総数による分析

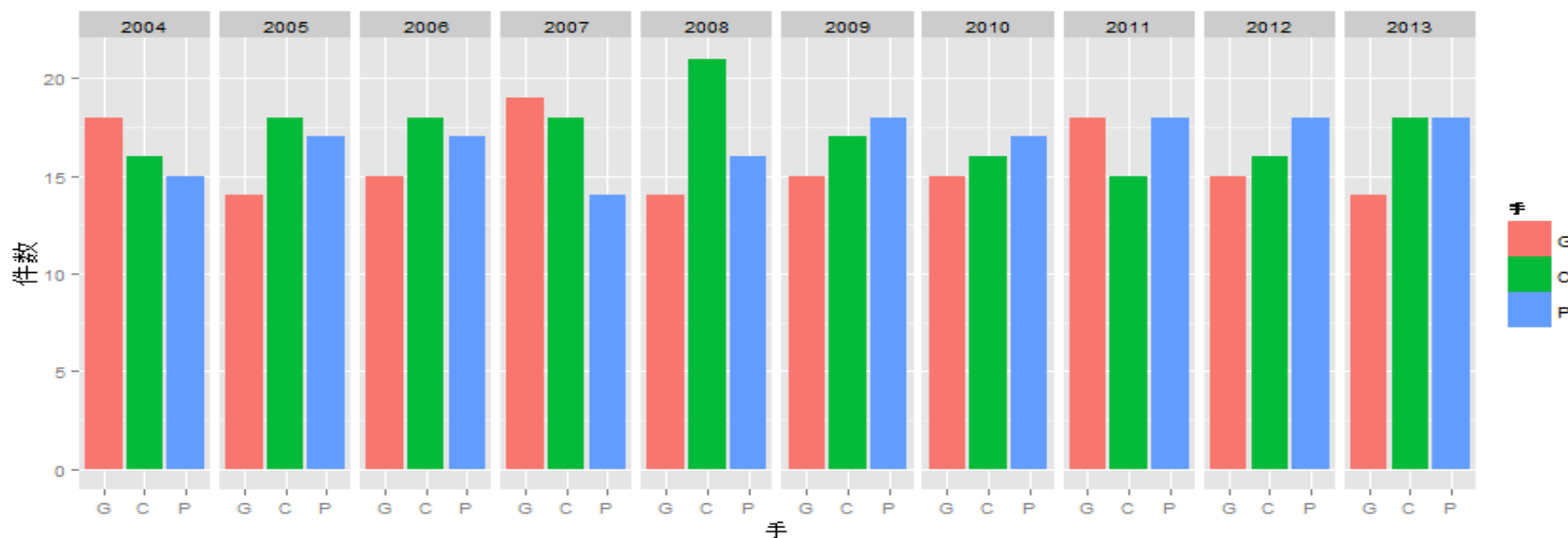
## ▶ 分析結果

直近10年分の年別種類別総数を見ると、2008年以外はだいたい均衡している。

グーの手は少ない年のほうが多く、3,4年ごとにバランスをとるように多くなっている。

このことから、2014年か2015年はグーが多くなる可能性が高い。

チョキとパーは均衡しており、パーを出すならチョキを出す方が勝率は高くなる。



## ■ じゃんけんの癖

### ▶ 10年分のデータからじゃんけんの癖を分析

- 出す手の出現率による分析

グー、チョキ、パー、あいこの出現率から分析する。

- 一手前との関連性による分析

自分が前に出した手と次に出す手のつながり

グーを出した後の次の手の出し方とチョキを出した後の出し方、パーを出した後の出し方を分析する。

- 三手の組み合わせによる分析

3回の手の続け方で癖を分析する。

人は同じ物を連続して出しにくいようで「グー」を出した次は「パー」か「チョキ」を出しがちです。

- 同手の間隔と順番による分析

同手が出る間隔に法則性があるのか分析する。

また、前回の同手と関連性があるのか分析する。

# 出す手の出現率による分析

- ▶ 10年分のデータからグー、チョキ、パー、あいこを出す確率  
美林大学の芳沢光雄教授による「[ジャンケンに関する研究結果](#)」によると  
グー（35.0%）、パー（33.3%）、チョキ（31.7%）とグーが最も多い。  
また、あいこ（22.8%）は出にくい結果になっている。

では、サザエさんのじゃんけんではどうなのか分析する。

【総数】

	kind	cnt
1	G	157
2	C	173
3	P	168

【あいこ】

	kind	cnt
1	G	31
2	C	25
3	P	34

チョキ(34.7%)、パー(33.7%)、グー(31.5%)とチョキが最も多く、上記とは違う結果となっている。あいこ（18.0%）については同様に出不にくい結果であった。

# 一手前との関連性による分析について

- ▶ 次の手は、前の手の勝ち手が負け手のどちらが多いのか  
2パターン(前手と当手)の組み合わせにして集計

	year	win	lose	draw	GP	CG	PC	GC	CP	PG	GG	CC	PP
1	2004	15	20	13	4	5	6	6	8	6	7	3	3
2	2005	21	17	11	6	7	8	6	7	4	3	4	4
3	2006	21	20	9	6	6	9	6	8	6	3	3	3
4	2007	19	18	14	5	8	6	7	6	5	6	5	3
5	2008	29	17	5	7	10	12	6	8	3	1	3	1
6	2009	8	31	11	2	3	3	11	11	9	3	3	5
7	2010	20	20	8	6	6	8	6	8	6	3	2	3
8	2011	21	21	9	7	9	5	9	5	7	2	1	6
9	2012	24	15	10	7	8	9	5	6	4	3	2	5
10	2013	22	18	10	4	7	11	6	9	3	4	1	5

# 一手前との関連性のグラフ化の準備

## ▶ ggplot2による円グラフ作成

```
#2パターンの列合計を求める  
sumptn <- colSums(dfptn[,-1])
```

#円グラフ 総合

```
allptn <- data.frame(kind=c("win", "lose", "draw"), val=c(sumptn["win"],sumptn["lose"],sumptn["draw"]))  
p <- ggplot(data.frame(allptn ), aes(x="", y = val, fill = kind))  
p + geom_bar(width = 1, stat = "identity" ) + coord_polar("y")+ xlab("") + ylab("総合") + labs(fill = "種類")
```

#円グラフ 勝手

```
winptn <- data.frame(kind=c("GP", "CG", "PC"), val=c(sumptn["GP"],sumptn["CG"],sumptn["PC"]))  
p <- ggplot(data.frame(winptn ), aes(x="", y = val, fill = kind))  
p + geom_bar(width = 1, stat = "identity" ) + coord_polar("y")+ xlab("") + ylab("勝手") + labs(fill = "種類")
```

#円グラフ 負手

```
loseptn <- data.frame(kind=c("GC", "CP", "PG"), val=c(sumptn["GC"],sumptn["CP"],sumptn["PG"]))  
p <- ggplot(data.frame(loseptn ), aes(x="", y = val, fill = kind))  
p + geom_bar(width = 1, stat = "identity" ) + coord_polar("y")+ xlab("") + ylab("負手") + labs(fill = "種類")
```

#円グラフ 同手

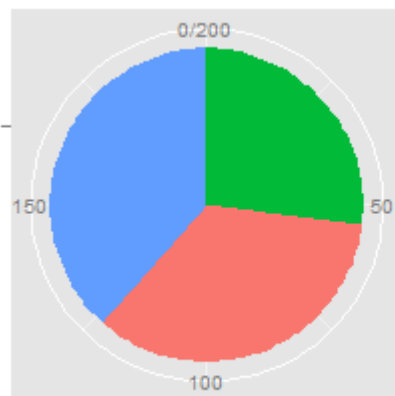
```
drawptn <- data.frame(kind=c("GG", "CC", "PP"), val=c(sumptn["GG"],sumptn["CC"],sumptn["PP"]))  
p <- ggplot(data.frame(drawptn ), aes(x="", y = val, fill = kind))  
p + geom_bar(width = 1, stat = "identity" ) + coord_polar("y")+ xlab("") + ylab("同手") + labs(fill = "種類")
```

# 一手前との関連性による分析

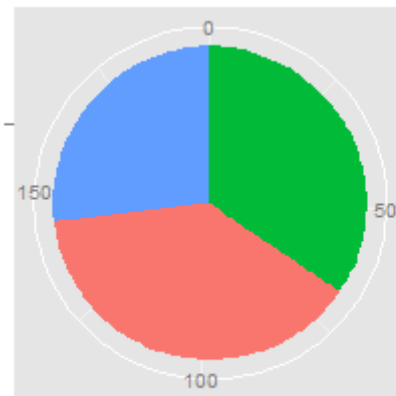
## ▶ 分析結果

一手前によって勝手(win)か負手(lose)にするかに差は見られない。また、勝手と負手の内訳でも差が見られない。同手(draw)が少ないのは、一般的な心理からか。同手の中でもチョキの割合は少ない、チョキの後はグーかパーに切り替えてくる。

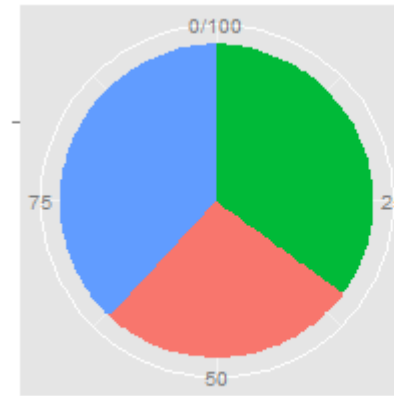
### 【内訳】



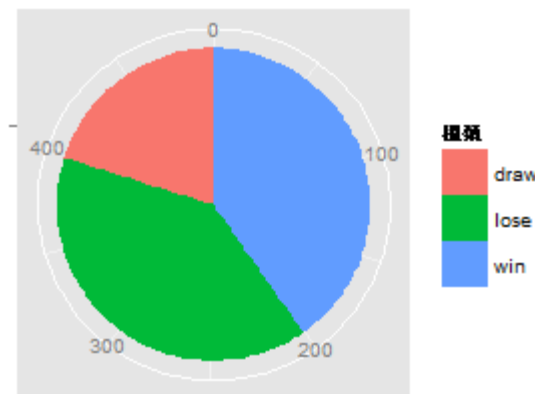
勝手



負手



同手



総合

# 三手の組み合わせによる分析について

- ▶ 二手前と一手前を含めた三手の組み合わせでどれが多いのか分析

3パターン(二手前と一手前と当手)の組み合わせにして集計

	year	GGG	GGC	GGP	GCG	GCC	GCP	GPG	GPC	GPP	CGG	CGC	CGP	CCG	CCC	CCP	CPG	CPC	CPP
1	2004	2	4	1	1	1	5	2	1	1	1	2	2	1	0	2	4	2	2
2	2005	0	2	1	2	1	2	1	4	1	1	1	5	3	0	1	1	4	2
3	2006	1	1	1	3	0	3	1	4	1	1	2	3	2	0	2	3	3	2
4	2007	0	5	1	2	2	4	2	1	2	4	1	3	3	0	1	2	3	1
5	2008	0	0	1	2	1	3	0	6	1	0	4	6	1	1	1	3	5	0
6	2009	0	3	0	1	2	7	0	0	2	0	3	0	0	0	3	6	2	3
7	2010	0	2	1	1	2	3	1	4	1	1	1	4	0	0	2	3	3	2
8	2011	0	2	0	4	1	4	2	2	3	0	2	7	0	0	1	2	0	3
9	2012	0	1	2	1	1	3	0	3	4	2	2	4	1	0	1	2	3	1
10	2013	0	2	2	1	0	5	0	3	1	3	2	2	0	0	1	3	2	4

PGG	PGC	PGP	PCG	PCC	PCP	PPG	PPC	PPP
4	1	1	3	1	2	0	3	0
2	2	0	2	4	3	2	1	1
1	3	2	1	3	4	2	1	0
2	2	1	3	2	1	1	2	0
1	2	0	7	2	3	0	1	0
3	4	2	3	0	1	3	2	0
2	3	1	5	0	3	2	1	0
2	5	0	5	0	0	3	3	0
1	2	1	6	1	2	2	3	0
1	2	0	5	1	3	0	4	0

# 三手の組み合わせによる分析

## ▶ トップ10

統計的には違う手の組み合わせが多い。

二手前と一手前が違う手になった場合、残りの手が出やすい。(例 GCならPを出す)

三手の中に同手がある場合、残りの手が出やすい。(例 GCGやGGCならPを出す)

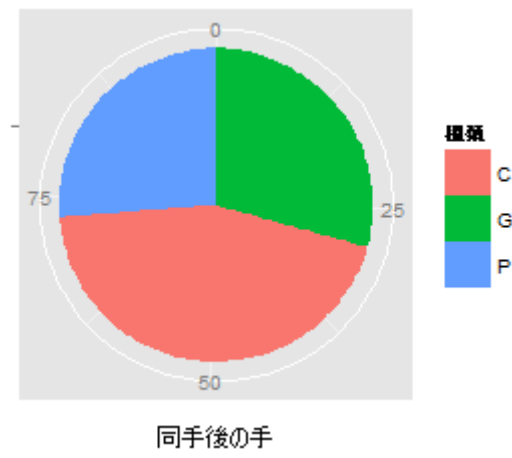
	row.names	1	2	3	4	5	6	7	8	9	10
1	kind	PCG	GCP	CGP	CPG	GPC	CPC	PGC	GGC	PCP	PPC
2	cnt	40	39	36	29	28	27	26	22	22	21

## ▶ 二手前と一手前が同手

二手前と一手前が同手の場合、チョキを出す割合が多い。(GGC,PPC,CCC)

	row.names	1	2	3	4	5	6	7	8	9
1	kind	GGC	PPC	CCP	PPG	CCG	GGP	GGG	CCC	PPP
2	cnt	22	21	15	15	11	10	3	1	1

	kind	val
1	G	29
2	C	44
3	P	26





# 同手の間隔と順番の分析

▶ 同手の間隔と順番に法則性があるのかを分析する  
右表の場合、グーの同手があり、1つ間隔を置いて  
チョキの同手となる。その後、3つ間隔を置いて  
パーの同手となっている。

同手になる間隔に法則性があるのか、また  
前回の同手の次の同手に法則性があるのかを  
分析する

488	1146	G	2013
489	1147	G	2013
490	1148	P	2013
491	1149	C	2013
492	1150	C	2013
493	1151	P	2013
494	1152	G	2013
495	1153	C	2013
496	1154	P	2013
497	1155	P	2013

# 同手の間隔と順番による分析

## ▶ どのくらいの間隔で同手となるのか

前回の同手があったから4回以内に同手になる割合が高い  
間隔なし(0回)で同手になることが多いのが意外であった。

	row.names	1	2	3	4	5	6	7	8	9	10	11	12	13
1	cnt	0	1	2	3	4	5	6	7	8	9	11	20	23
2	spccnt	17	20	9	10	4	5	3	3	5	1	1	1	1

## ▶ 同手の順番に法則性があるのか

前回の同手があった場合、前回の同手の勝手になる割合が高い

	win	lose	draw	GP	CG	PC	GC	CP	PG	GG	CC	PP
1	40	31	18	18	11	11	8	8	15	5	5	8

# じゃんけんの癖のまとめ

## ▶ 分析結果

### ● 迷ったらグーを出す

出現率ではチョキ(34.7%)、パー(33.7%)、グー(31.5%)とチョキが最も多い。

同手ではパーが最も多いが、次点のグーとの差は均衡している。

グーを優先し、チョキかパーで悩んだらチョキを出す。

### ● 前回と違う手を出す

同じ物を連続して出しにくいことから、前回と違う手を出す。

その際に、前手に対して勝手か負手を出すかを意識する必要はない。

### ● 三手とも違う手を出す

三手の組み合わせの分析から、二手前と一手前が分かれば残りの手が出てくる。

また、三手の中に同手がある場合にも 残りの手が出てくる。

三回連続で同手にすることは、ほぼ無いと思っていい。

### ● 同手を出した後にすぐに同手を出す

同手を出してきた場合、短い間隔で同手を出してくる。その際に前回の同手と違う手(勝手)を出す。

# ■ 人間乱数の分析

## ▶ 人間乱数の分析

人間によって次の手を決めているのか、コンピュータ等を使用して次の手を決めているのか分析する。

## ▶ 都市伝説 キター！！(by 都市伝説の女)

エンディングの最後で一家が山小屋に突っ込むシーン。このシーンで煙突の煙の状態を見ると、その後のじゃんけんに影響を及ぼすという。

- ・ 煙がわっか状 → グー
- ・ 煙がひも状 → チョキ
- ・ 煙がない → パー

という法則があるらしい。

下記サイトの検証結果からあくまで都市伝説である。

『サザエさんじゃんけん』には必勝法があるという噂がTwitterで広まる

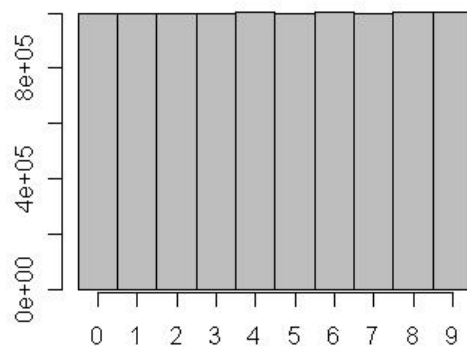
<http://www.yukawanet.com/archives/3978209.html>

# 人間乱数の分析の準備

## ► コンピュータの乱数の質

R言語はメルセンヌ・ツイスター法により一様乱数を生成している

「[Rによるデータの統計的取り扱い/一様乱数表](#)」のサイトによると  
大量の乱数(1000万個)を発生させたヒストグラムでほぼ一様な分布となります。



今回は15個の乱数を剰余で3種類にし、1年分の52個の乱数を生成しました。

```
hand <- c('G','C','P');  
rnd <- floor(runif(52, 0, 14));  
idx <- rnd[y] %% 3 + 1;  
kind <- hand[idx];
```

# 人間乱数の分析①

## ▶ あいこの出現率による分析

美林大学の芳沢光雄教授による「ジャンケンに関する研究結果」によると  
あいこ（22.8%）は出にくい結果になっている。

サザエさんのじゃんけんでは、あいこ（18.0%）でした。

コンピュータで生成したじゃんけんでは、498個中100個のあいこ(20.0 %)となり、あいこの出現率だけでは判断できない。

	kind	cnt
1	C	35
2	G	35
3	P	30

## 人間乱数の分析②

### ▶ 三手の組み合わせによる分析

サザエさんのじゃんけんでは、違う手の組み合わせが多かった。

三手とも同じ手が少ないことから、手を加えていると思われる。

	row.names	1	2	3	4	5	6	7	8	9	10
1	kind	PCG	CGP	GCP	GPC	CPG	CPC	PGC	GGC	PCP	PPC
2	cnt	49	46	39	37	29	27	26	22	22	21

GGG	CCC	PPP
3	1	1

コンピュータで生成したじゃんけんでは、同じ手が2つ入る組み合わせが多い。

今回の結果はあまり良くないですが、三手とも同じ手が多いのも乱数の特徴と思われる。

	row.names	1	2	3	4	5	6	7	8	9	10
1	kind	GCP	CGC	CPC	CCC	CCG	PCP	GCC	CGG	CPG	PCG
2	cnt	28	28	25	24	23	22	21	21	21	21

GGG	CCC	PPP
12	24	6

# 人間乱数の分析③

## ▶ 同手の間隔による分析

サザエさんのじゃんけんでは、意外と前回の同手から直ぐに同手になる割り合いが高かった。

	row.names	1	2	3	4	5	6	7	8	9	10	11	12	13
1	cnt	0	1	2	3	4	5	6	7	8	9	11	20	23
2	spccnt	17	20	9	10	4	5	3	3	5	1	1	1	1

コンピュータで生成したじゃんけんでも同様に、前回の同手から直ぐに同手になりやすい。  
0回が最も多くなってしまうのは、コンピュータによる乱数ならではと思われる。

	row.names	1	2	3	4	5	6	7	8	9	10
1	cnt	0	1	2	3	4	5	7	9	13	15
2	spccnt	33	17	11	7	9	5	3	3	1	1



# 人間乱数の分析のまとめ

## ▶ 分析結果

コンピュータによる乱数でもグー、チョキ、パー、あいこの出現率だけなら、人間による手の出現率と比べても偏りに違いはなかった。

ただ、3種類と少ないとどうしても同じ手が連続で出やすくなってしまう。

サザエさんのじゃんけんでは、コンピュータやサイコロなどを用いた乱数で手を決めたとしても、最終的には人が関わって調整していることがわかる。

今回は15種類の乱数を剰余により3分割する簡易な方法であったが、同じ手が出にくいようにアルゴリズムを調整することで、人間乱数っぽく出来るだろう。

## ▶ 別の方の調査結果

研究タイトル「偏りを知る」 ～じゃんけんなどにおける統計的推定～

<https://ssh.jst.go.jp/research/show/404>

## ■ 次の手の予測と勝敗結果

▶ じゃんけんの癖から次の手の勝手を予測する

- ・ チョキが多いので、グー > チョキ > パーの優先順位とする
- ・ 前回と違う手を出すので、上記の優先順位で勝手を選ぶ
- ・ 二手前と一手前が違う手なら、残りの手を出すので勝手を選ぶ
- ・ 三手の中に同手がある場合、残りの手を出すので勝手を選ぶ
- ・ 二手前と一手前が同じ手なら、勝手を出すので負手を選ぶ

上記を基にプログラムを組み過去データと検証する。かなりベタベタです。

※同手の間隔処理については面倒なのでやめる。

# 次の手の予測プログラム

```
if(fstkind == 'G') guess <- 'C';
```

```
if(fstkind == 'C') guess <- 'G';
```

```
if(fstkind == 'P') guess <- 'G';
```

#2手前が在る場合

```
if(sndkind != ""){
```

```
  if(sndkind != fstkind){
```

```
    #残りの手が出ると予想するので残りの手の勝手にする
```

```
    ptn <- paste(fstkind,sndkind,sep="");
```

```
    if(ptn == 'GC' || ptn == 'CG') guess <- 'C';
```

```
    if(ptn == 'CP' || ptn == 'PC') guess <- 'P';
```

```
    if(ptn == 'PG' || ptn == 'GP') guess <- 'G';
```

```
  }
```

```
else{
```

```
  #同一なら勝手と予想するので負手にする
```

```
  if(fstkind == 'G') guess <- 'C';
```

```
  if(fstkind == 'C') guess <- 'P';
```

```
  if(fstkind == 'P') guess <- 'G';
```

```
}
```

# 次の手の予測プログラム

#3手前が在る場合

```
if(thrkind != ""){  
  #違う組み合わせ 残りの手が出ると予想するので残りの手の勝手にする  
  ptn <- paste(fstkind,sndkind,sep="");  
  if(ptn == 'GC' || ptn == 'CG') guess <- 'C';  
  ptn <- paste(fstkind,sndkind,thrkind,sep="");  
  if(ptn == 'GCG' || ptn == 'CGC') guess <- 'C';  
  if(ptn == 'CPC' || ptn == 'PCP') guess <- 'P';  
  if(ptn == 'PGP' || ptn == 'GPG') guess <- 'G';  
  if(ptn == 'GGC' || ptn == 'CCG' || ptn == 'GCC' || ptn == 'CGG') guess <- 'C';  
  if(ptn == 'CCP' || ptn == 'PPC' || ptn == 'PCC' || ptn == 'CPP') guess <- 'P';  
  if(ptn == 'PPG' || ptn == 'GGP' || ptn == 'GPP' || ptn == 'PGG') guess <- 'G';  
}  
}
```

# 予測した手との勝敗結果

2004年～2013年のじゃんけんデータを基に予測した手との勝率

勝率の計算方法 勝ち数÷(勝ち数+負け数) ※引き分けを除外

年	勝ち	負け	引き分け	勝率
2004	24	10	15	0.71
2005	21	15	13	0.58
2006	23	13	14	0.64
2007	22	17	12	0.56
2008	29	13	9	0.69
2009	31	6	13	0.84
2010	28	7	13	0.80
2011	31	8	12	0.79
2012	27	12	10	0.69
2013	25	11	12	0.69

平均の勝率  
0.7と高確率  
となりました。

オモコロの下記サイトの記録「2011年の51戦28勝12敗11分」には報いました。

[サザエさんと一年中じゃんけんしてみた 男の戦い](#)

## まとめ

人間が最終的に調整しているので案外予測しやすくなったのかなと思います。  
きっと、アシスタントディレクターとディレクターやプロデューサーの間では  
こんなやり取りをして調整しているかも。

A D 「今週のじゃんけんの手は何にします？」

P D 「チョキでいいんじゃない」

A D 「先々週と先週もチョキでしたけど、どうしますか？」

P D 「そっか、じゃーグーで。よろしく！」

A D 「今週のじゃんけんの手は何にします？」

P D 「グーでいいんじゃない」

A D 「今月は、パーとグーばかりですよ」

P D 「そっか、じゃーチョキだね。よろしく！」

# たわごと

今回、R言語を使ってデータ分析する上で統計解析ツール(標準偏差や主成分分析やカイ二乗検定など)を使いこなせれば良かったんですが、私の知識不足もあって中学生レベルの解析方法となってしまうしました。

前半はR言語でこういうことが出来ますよ、グラフをきれいに  
出せますよって感じで書きましたが、中盤以降はソースリストを  
掲載するのが資料上難しくなったので、その過程は飛ばしました。  
今回のRスクリプトは公開しているので見てさわってもらえばいい  
でしょう。

ご清聴ありがとうございました！