

プロ野球の観客動員数の決定要因

赤木陽太，上野知和，吉田達哉

2022-11-09

はじめに

プロ野球のテレビ放送が減って，人気が衰えているのではないか．有名選手が大リーグに移籍して，日本の球団への関心がなくなっているのではないか．

などの，問題意識から，プロ野球の観客動員数の決定要因を考えてみたい．

先行研究について

日本のスポーツに関する観客動員数のこれまでの研究は，西山(2016)が2011年までの研究を簡単にまとめている．それによると，どれも観客数を被説明変数とした回帰分析を行っているが，要因について工夫が見られる．

すなわち，経済的要因(一人当たりの県民所得，人口)，観戦要因(休日の試合)，試合要因(ホームチームの順位，ビジターの順位，前年の順位)，人気要因，交流戦やクライマックスシリーズの影響がとりあげられている．西山(2016)は，2015年と2016年のセパ両リーグについて，とくに，球場の最大収容人数とイベントを要因として新たに加えている．その結果，イベントの効果が統計的に有意で大きかったと結論づけている．

続いて，鈴木(2018)は，2017年のデータをもちいて，年間勝利数，各スタジアムの収容人数，総年俵，プロ野球グッズストアにおける各球団のグッズの種類数を取りあげて分析している．その結果，年間の勝利数とスタジアムの収容人数が，観客動員数に統計的に有意な結果をもたらすことを見だしている．つまり，勝利数が多いほど観客数は多くなるが，それに見合ったスタジアムの大きさが必要であるという結論を導いている．

以上のことから，この研究においても，まずは経済的要因と，スタジアムの収容人数，勝利数を基本的な説明要因として考えて，これを手がかりに新しい要因についても考察していくことを目的としている．

データについて

データは各球団のホームページからの情報によっている。

もっとも重要な観客動員数の定義は、各球場・球団で発表されるデータのため厳密に正確さを追及することはできないが、おおむねトレンドは表しているものと考えている。

ただし、先行研究のデータが、2015年から2017年と比較的安定した時期のものであることにくらべて、われわれのデータは、2022年のものであることに注意したい。コロナ禍による観客制限が2020年2021年と行われたため、その直後のため自主的に控える時期も含まれていると考えられる。

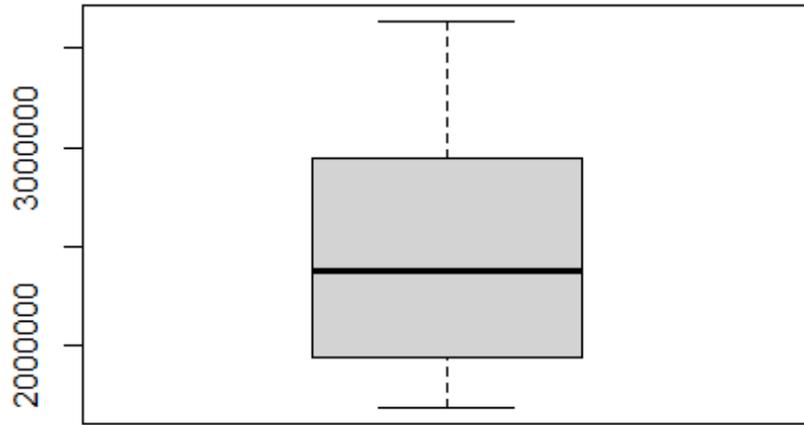
```
library(openxlsx)
x=read.xlsx("H:/seminar/2022/members/AkagiUenoYoshida/プロ野球の情報(1).xlsx",sheet=1)
```

データの記述統計量

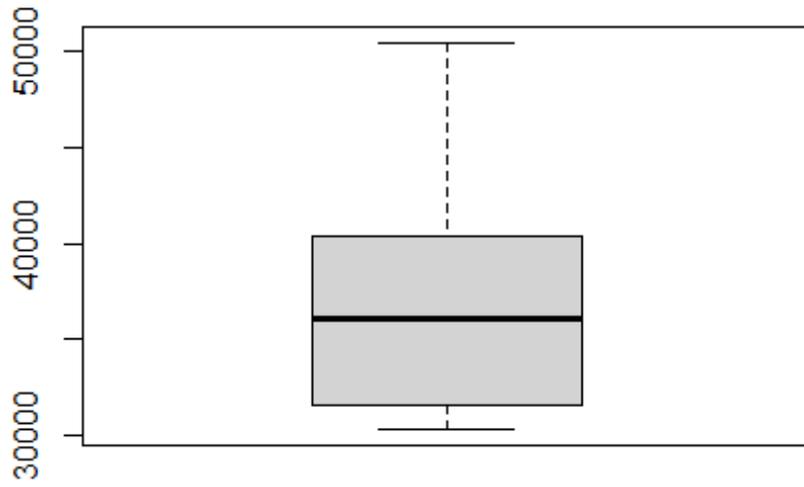
```
summary(x)
```

```
##          X1          入場者数          スタジアム収容人数
## Length:12          Min.   :1683700.00          Min.   :30308.0
## Class :character          1st Qu.:1960900.00          1st Qu.:31866.5
## Mode  :character          Median :2372450.00          Median :36116.0
##                               Mean   :2454441.67          Mean   :36930.5
##                               3rd Qu.:2860425.00          3rd Qu.:39454.5
##                               Max.   :3637000.00          Max.   :50454.0
##          地元人口          地元平均年収          スタジアム値段
## Min.   : 2302000.00          Min.   :4296600.00          Min.   :4100.00000
## 1st Qu.: 5202500.00          1st Qu.:4646050.00          1st Qu.:5175.00000
## Median : 6814500.00          Median :4846050.00          Median :5450.00000
## Mean   : 7439083.33          Mean   :4980291.67          Mean   :5608.33333
## 3rd Qu.: 8937750.00          3rd Qu.:5284175.00          3rd Qu.:5850.00000
## Max.   :14048000.00          Max.   :5849300.00          Max.   :8200.00000
##          勝利数
## Min.   :59.0000000
## 1st Qu.:67.5000000
## Median :69.0000000
## Mean   :70.1666667
## 3rd Qu.:73.7500000
## Max.   :80.0000000
```

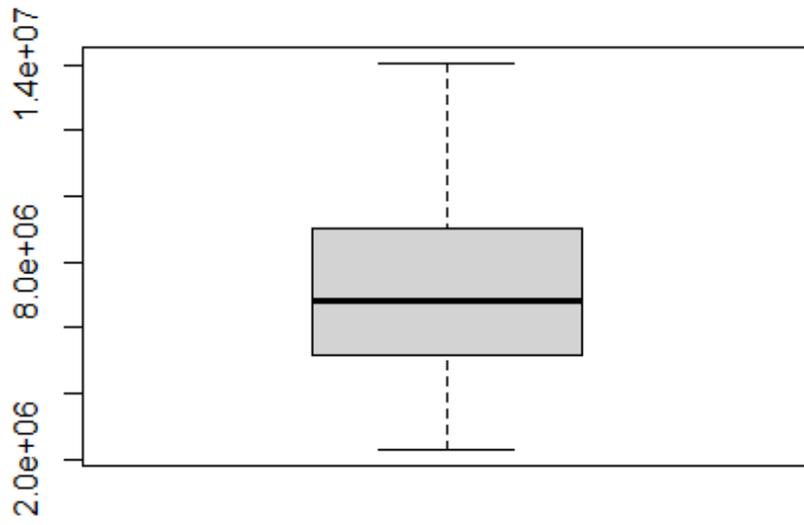
```
varname=colnames(x)
for( i in 2:7 ){boxplot(x[,i],xlab=varname[i])}
```



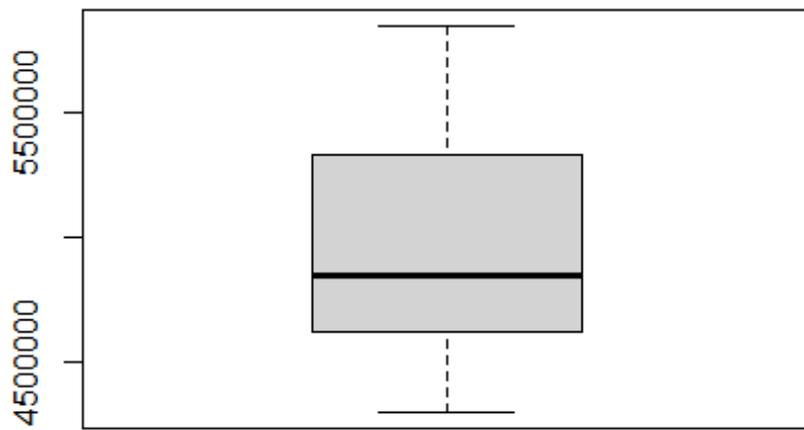
入場者数



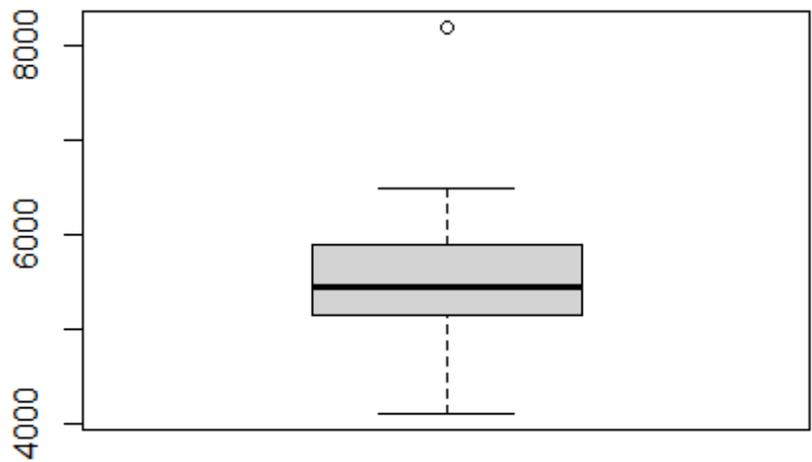
スタジアム収容人数



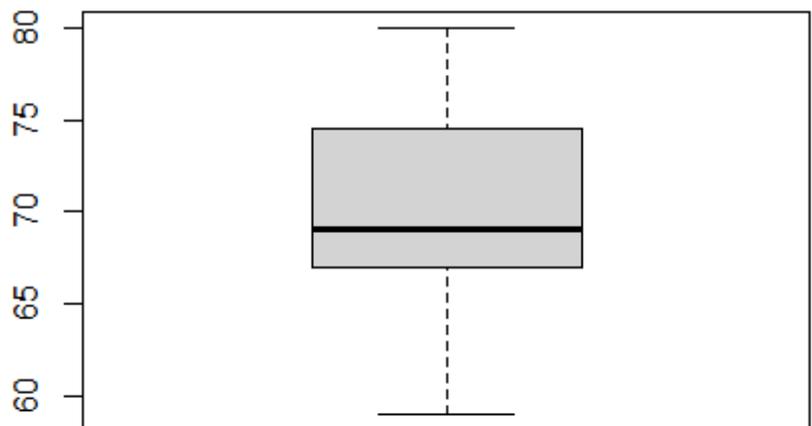
地元人口



地元平均年収



スタジアム値段



勝利数

観客入場者数については、偏った分布をしていないことがわかる。標本平均が標本メジアンよりも大きいので、入場者数の少ない球場がよりかたまっていることが見られるが、それほどひどく歪んでいるわけではないことが、ボックス・プロット図からもわかる。スタジアムの収容人数についても同様である。

偏っているのは、地元平均年収と勝利数である。年収についてはもともとの所得分布が偏っているので、当然である。勝利数は比較的偏らないものと考えられるが、一部の球団に勝利が集中している傾向があることがわかる。

以上のような記述統計の状況から、明確な外れ値は観察されていないこと、観客入場数についても偏りが大きくないことから、まず考えられるのは線形の回帰分析である。

分析手法について

データの記述統計の特徴や先行研究にしたがって、まずは重回帰分析を行って、傾向をとらえてみたい。

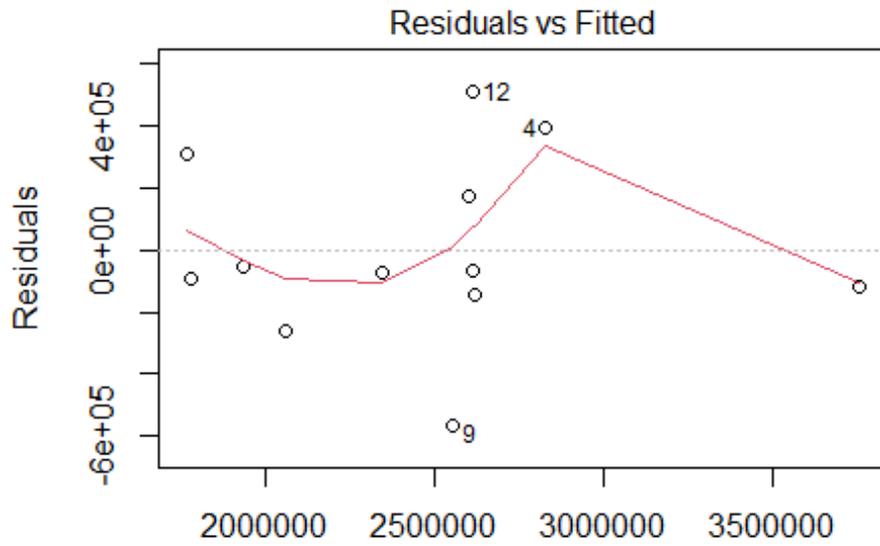
そのうえで、入場者数を収容人数で割った稼働率、スタジアムの寝代を平均年収で割った相対的なコストな説明要因として加えて分析してみたい。

```
reg1=lm(入場者数~スタジアム収容人数+地元人口+地元平均年収+スタジアム値段+勝利数,
data=x)
summary(reg1)

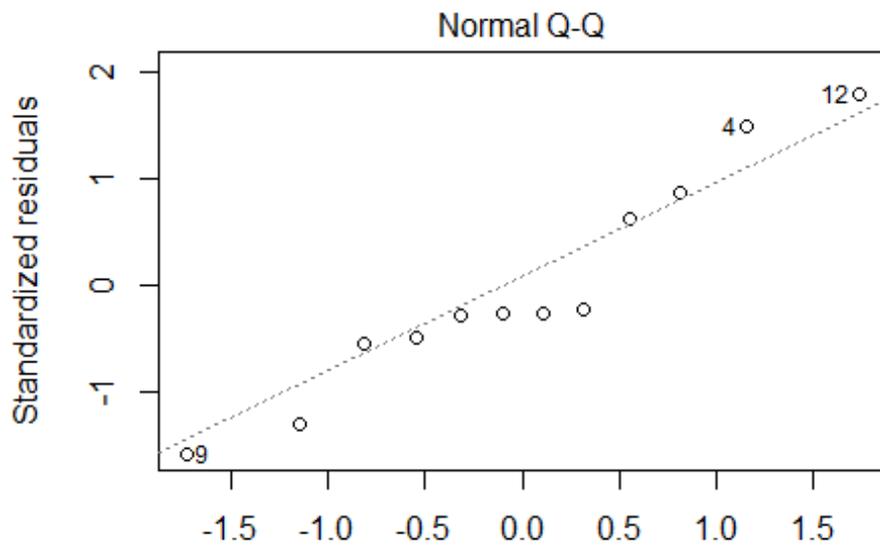
##
## Call:
## lm(formula = 入場者数 ~ スタジアム収容人数 + 地元人口 +
##     地元平均年収 + スタジアム値段 + 勝利数, data = x)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -562869.329 -125556.302  -70418.940  205797.498  507443.377
##
## Coefficients:
##              Estimate      Std. Error  t value Pr(>|t|)
## (Intercept) -6.41346664e+06  3.20538174e+06 -2.00084 0.092318 .
## スタジアム収容人数  6.70568065e+01  2.15898847e+01  3.10594 0.020957 *
## 地元人口      -2.22948672e-01  8.84845549e-02 -2.51963 0.045310 *
## 地元平均年収   1.48971602e+00  6.24172534e-01  2.38671 0.054269 .
## スタジアム値段 -1.19638647e+02  1.23775087e+02 -0.96658 0.371082
## 勝利数         1.85523107e+04  2.67829345e+04  0.69269 0.514405
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 403116.926 on 6 degrees of freedom
## Multiple R-squared:  0.773490213,    Adjusted R-squared:  0.584732057
## F-statistic: 4.09778433 on 5 and 6 DF,  p-value: 0.0577418882

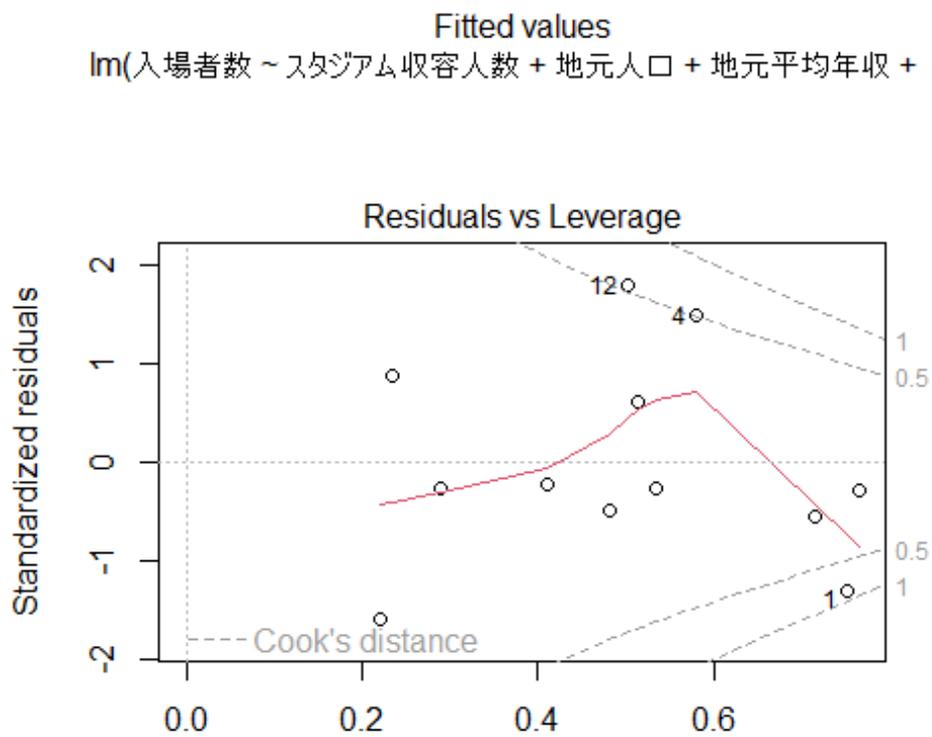
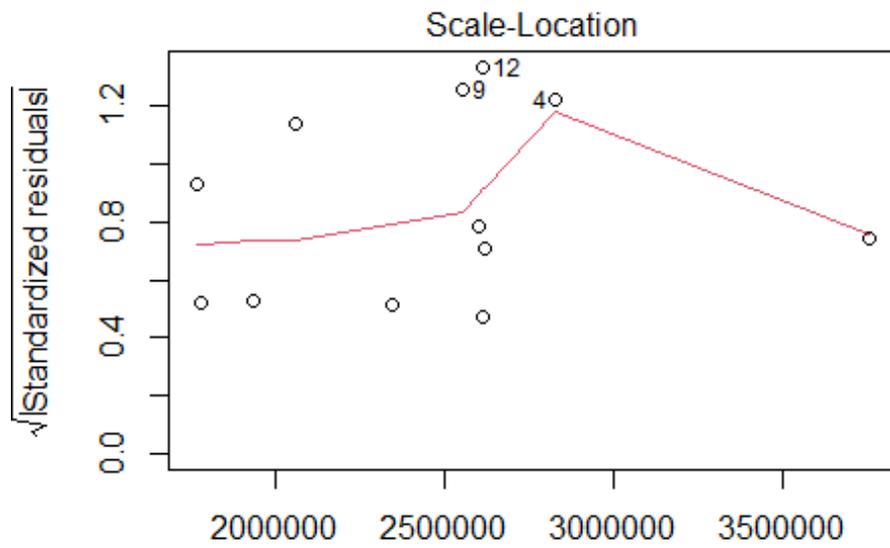
plot(reg1)
```



Fitted values
 $\text{lm}(\text{入場者数} \sim \text{スタジアム収容人数} + \text{地元人口} + \text{地元平均年収} + \dots)$



Theoretical Quantiles
 $\text{lm}(\text{入場者数} \sim \text{スタジアム収容人数} + \text{地元人口} + \text{地元平均年収} + \dots)$



Model: $\text{lm}(\text{入場者数} \sim \text{スタジアム収容人数} + \text{地元人口} + \text{地元平均年収} + \dots)$

```

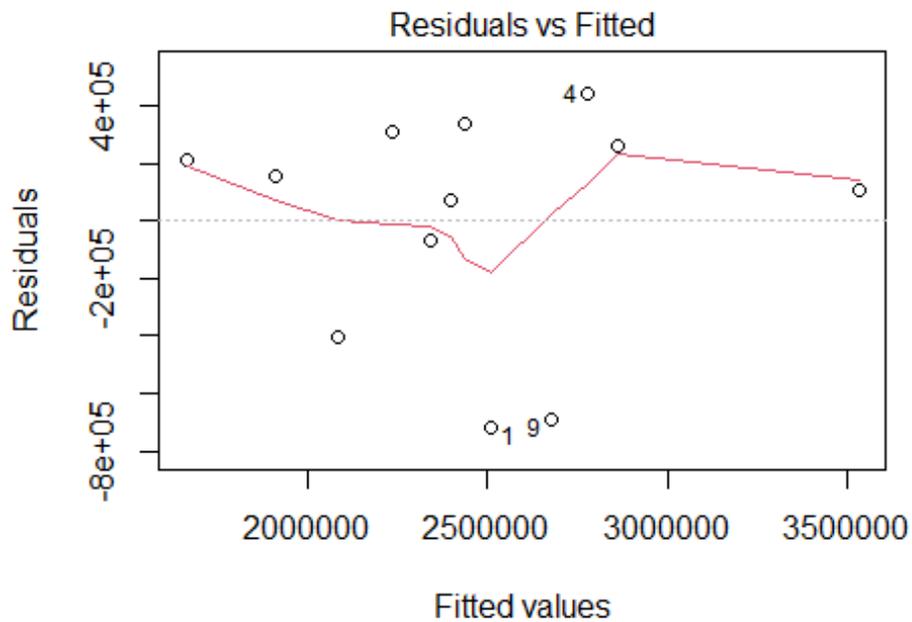
r_occupation=x$入場者数/x$スタジアム収容人数
r_cost=x$スタジアム値段/x$地元平均年収

reg2=lm(入場者数~スタジアム収容人数+地元人口+r_cost+勝利数,data=x)
summary(reg2)

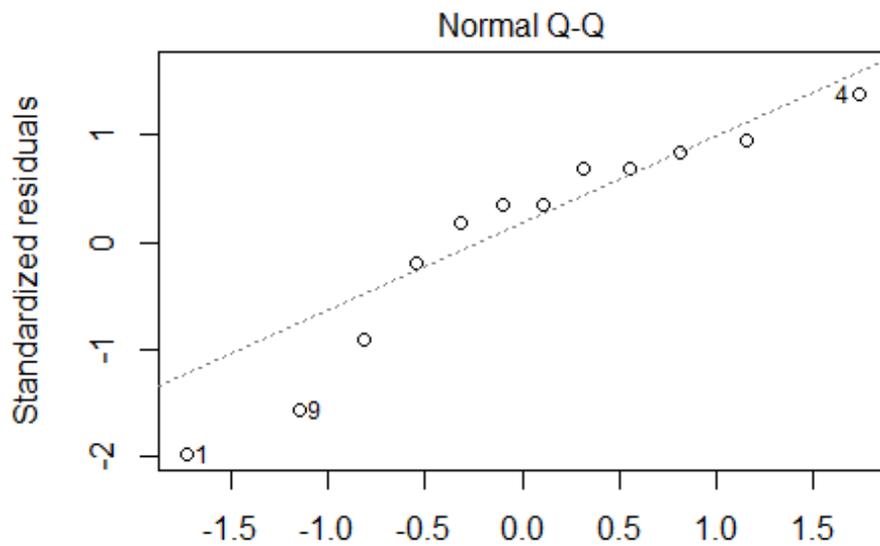
##
## Call:
## lm(formula = 入場者数 ~ スタジアム収容人数 + 地元人口 +
##     r_cost + 勝利数, data = x)
##
## Residuals:
##           Min             1Q         Median             3Q            Max
## -716748.085 -152000.697  128805.495  270678.266  440201.639
##
## Coefficients:
##              Estimate      Std. Error  t value Pr(>|t|)
## (Intercept) -5.73615212e+05  2.99524568e+06  -0.19151  0.853564
## スタジアム収容人数  7.05878280e+01  2.64194612e+01   2.67181  0.031916 *
## 地元人口      -6.02164758e-02  5.52602204e-02  -1.08969  0.311935
## r_cost       -8.24415599e+08  6.75990172e+08  -1.21957  0.262122
## 勝利数       2.58178532e+04  3.25522963e+04   0.79312  0.453747
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 493941.538 on 7 degrees of freedom
## Multiple R-squared:  0.603244648,    Adjusted R-squared:  0.376527304
## F-statistic: 2.66077856 on 4 and 7 DF,  p-value: 0.122399466

plot(reg2)

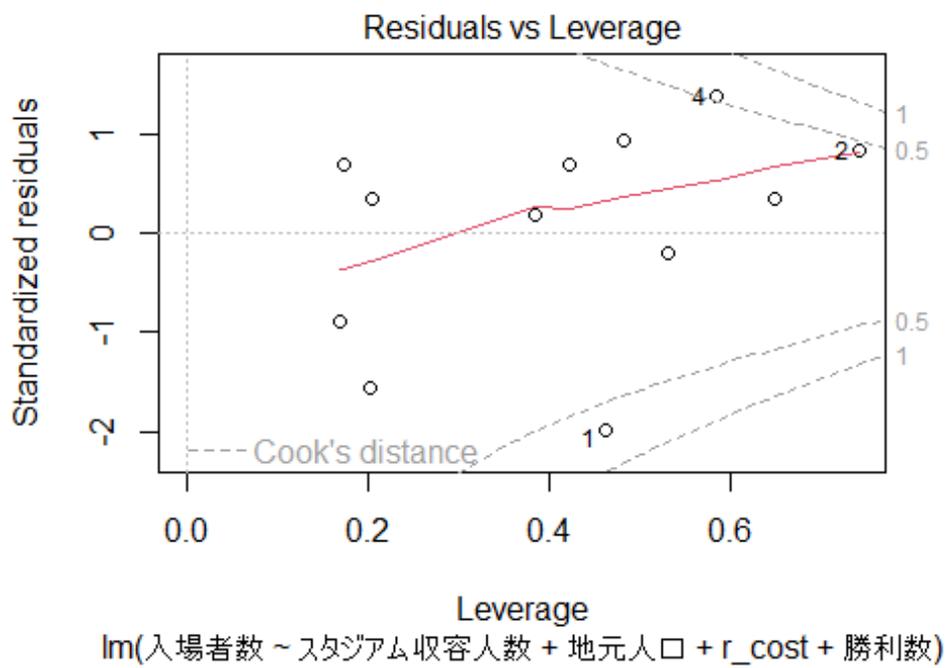
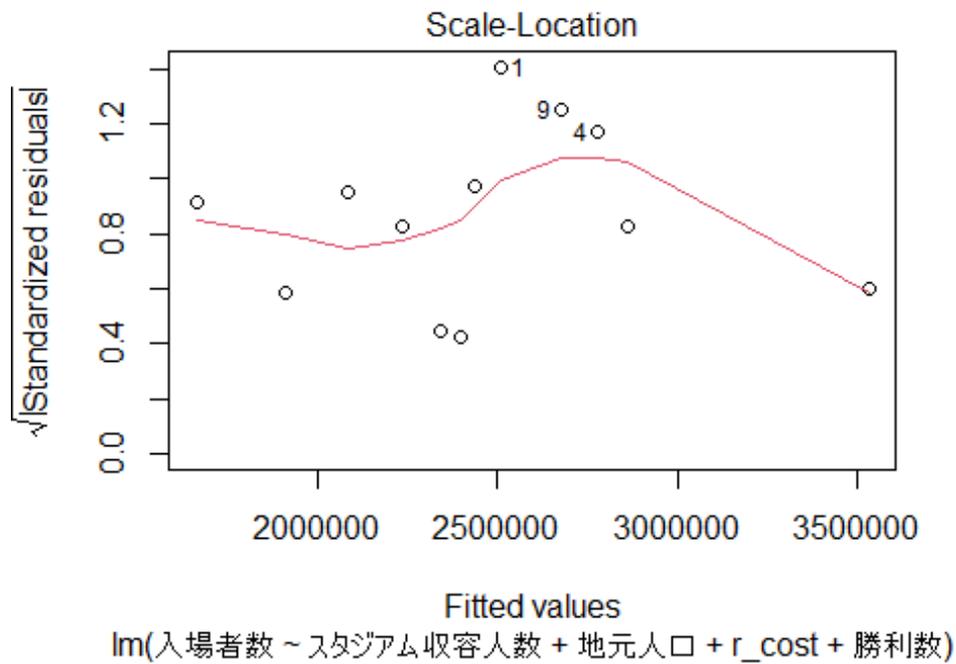
```



lm(入場者数 ~ スタジアム収容人数 + 地元人口 + r_cost + 勝利数)



lm(入場者数 ~ スタジアム収容人数 + 地元人口 + r_cost + 勝利数)



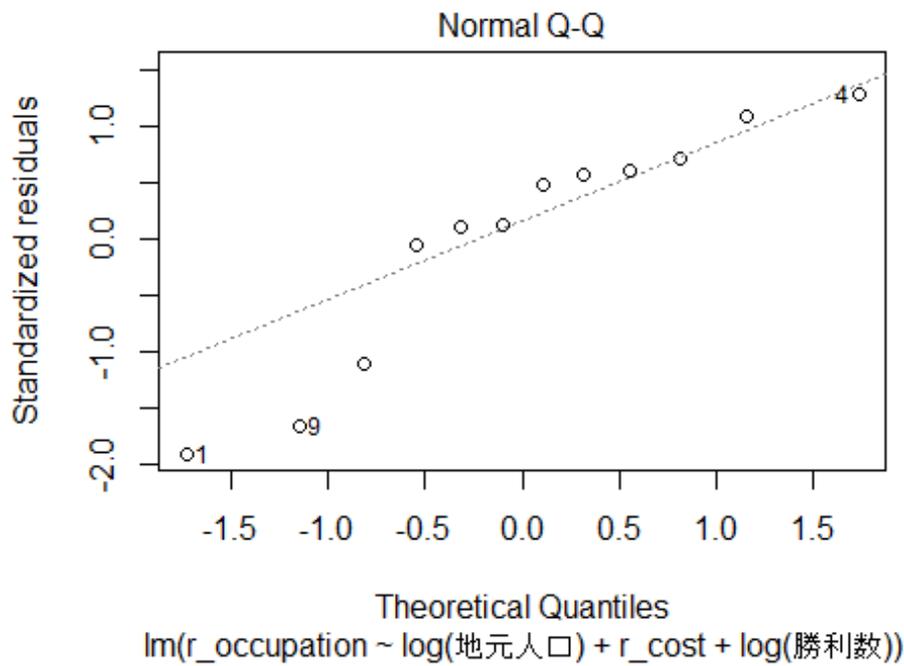
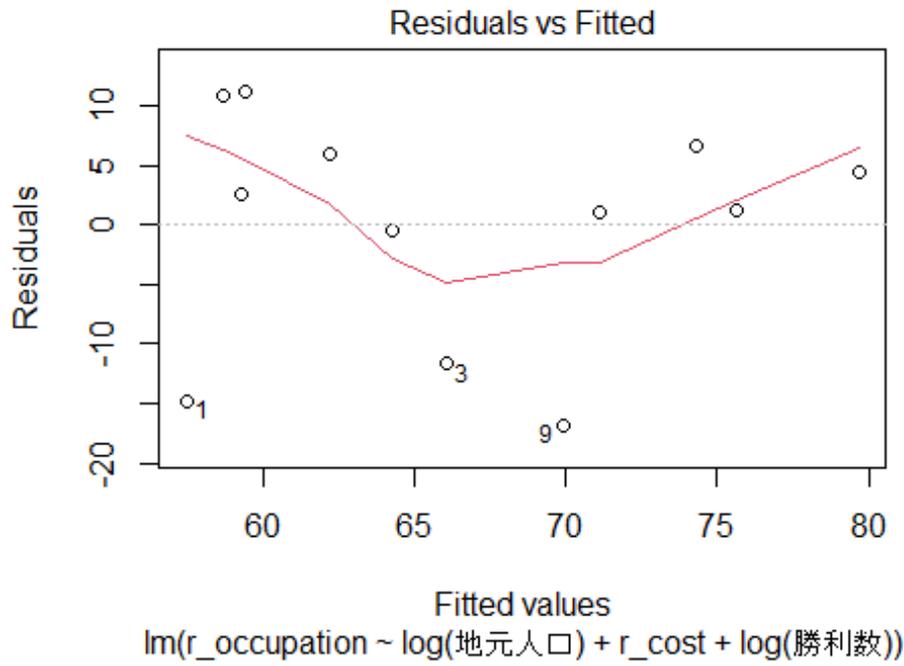
```

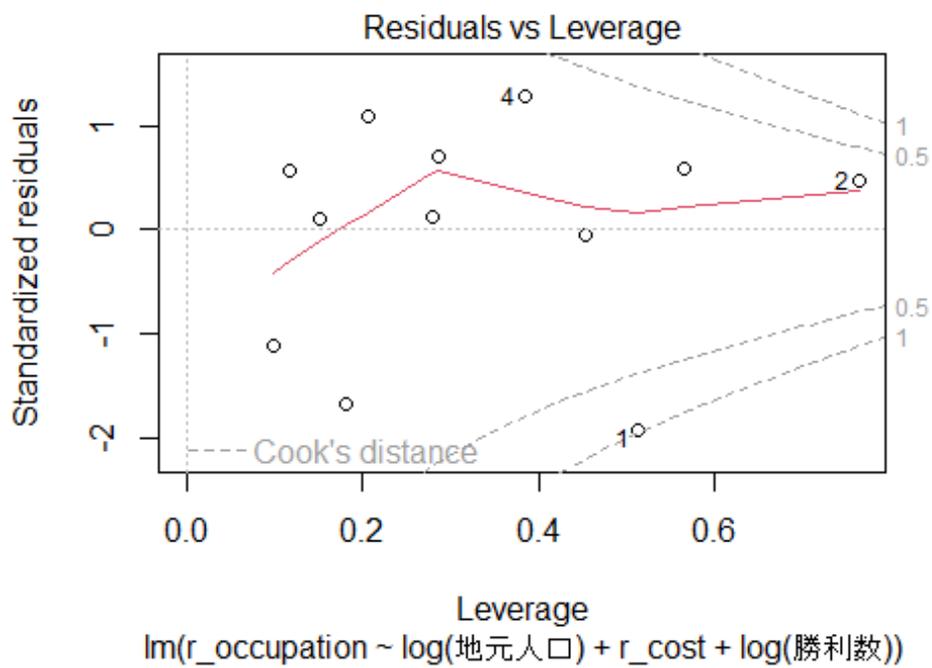
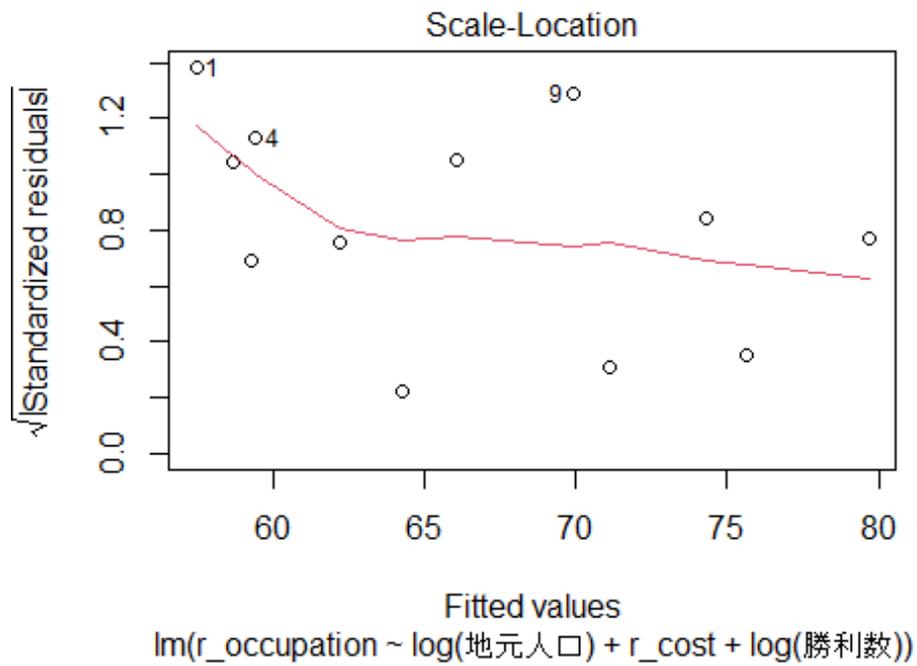
reg3=lm(r_occupation~log(地元人口)+r_cost+log(勝利数),data=x)
summary(reg3)

##
## Call:
## lm(formula = r_occupation ~ log(地元人口) + r_cost + log(勝利数),
##     data = x)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -16.78316125  -3.22107976   1.88821661   6.11919904  11.14793563
##
## Coefficients:
##              Estimate      Std. Error  t value Pr(>|t|)
## (Intercept)    200.35504715    220.92569895   0.90689 0.390963
## log(地元人口)  -16.95861614     9.52557221  -1.78033 0.112894
## r_cost        -33746.14193431  16946.57641957 -1.99133 0.081605 .
## log(勝利数)     40.27591722     45.23469940   0.89038 0.399248
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 11.1066846 on 8 degrees of freedom
## Multiple R-squared:  0.384940044,    Adjusted R-squared:  0.15429256
## F-statistic: 1.66895401 on 3 and 8 DF,  p-value: 0.249862381

plot(reg3)

```





分析の結果

サンプルサイズが 12 と小さいためはっきりした結論は言えないが、グラフの赤い線が水平であることが望ましいが、ところどころ折れ曲がっているため、不均一分散がある可能性が示唆される。残差が正規分布しているかどうかは、Normal Q-Q プロットの点線上に○点が乗っていることが必要であるが、そのようにはなっていない。

分析手法を重回帰から変更する必要があるといえる。つまり、得られた推定パラメータの標準偏差がバイアスしていて、その結果得られる t-値が t-分布していない可能性がある。統計的な有意性に疑問が残るということである。

その点を留保したうえで、あえて解釈をすすめると、観客入場者数に収容人数は統計的に有意な結果で計測されている。地元人口と地元平均年収も最初の分析では有意な結果であるが、負の係数が得られた地元の人口については、人口が多いと入場者数は少なくなるという期待しない結果である。地元の年収が高い場合や、勝利数が多い場合には、入場者数は増える傾向にあるといえる。さらに、スタジアムの価格が高いと、入場者数減る。という、おおむね期待通りの推定結果が得られている。

つぎにコスト面でより鮮明にするかと思ひ、地元の所得でスタジアムの価格を割った r_cost (相対費用)の係数については、負の値が得られているものの、統計的に有意ではないという結果となった。

最後に、稼働率を決める変数としては、 r_cost は負の有意な結果になったので、入場者数を収容人数で割った場合には、地元民に対して観戦の相対コストが高いと稼働率は下がることが計算されている。

おわりにかえて

ここでの結果は、まだはじめばかりであるので、今後はより説得的な変数を探すと同時に、分析手法の拡充を考えていきたい。

参考文献

- 西山雄大(2016)「プロ野球の観客動員数に関する分析」早見ゼミ卒業論文第9期。
- 鈴木夏海(2018)「日本プロ野球において観客動員数を増加させるためには何が有効か」早見ゼミ卒業論文第11期。
- 日本野球機構オフィシャルサイト (<http://npb.jp/>)

- NPB オフィシャルオンラインショップ
(<https://shop.npb.or.jp/npbshop/index.html>)

球団ホームページ

- オリックス・バファローズ <http://www.buffaloes.co.jp/>
- 埼玉西武ライオンズ <http://www.seibulions.jp/>
- 千葉ロッテマリーンズ <http://www.marines.co.jp/>
- 中日ドラゴンズ <http://dragons.jp/>
- 東北楽天ゴールデンイーグルス <http://www.rakuteneagles.jp/>
- 阪神タイガース <http://hanshintigers.jp/>
- 広島東洋カープ <http://www.carp.co.jp/>
- 福岡ソフトバンクホークス <http://www.softbankhawks.co.jp/>
- 北海道日本ハムファイターズ <http://www.fighters.co.jp/>
- 横浜 DeNA ベイスターズ <http://www.baystars.co.jp/>
- 読売ジャイアンツ <http://www.giants.jp/>
- 東京ヤクルトスワローズ <http://www.yakultswallows.co.jp/>