

コーヒーの消費量は生活にどのような影響を与えているのか

・序論

コーヒーを普段よく飲む者にとって、コーヒーの消費量がその他の生活に関わる要因（幸福度、ストレス、死亡率など）にどのような影響を与えているのかを知ることは非常に重大である。

そこで本稿では、国民1人あたりのコーヒー年間消費量が、その国の幸福度に与える影響を統計的検定（主に回帰分析）を用いながら分析していく。調査対象は、国際コーヒー機関（ICO）による「世界の1人あたりコーヒー消費量」資料に載っている国のうちEU域内の国を除いた17国である。

・方法

国民1人あたりのコーヒー年間消費量としては、ICO統計が実施した「世界の1人あたりコーヒー消費量」による資料を用いる。ただし、EU28国間の境界が年々低くなり国別の資料があまり意味をなさなくなった結果ICOは2014年からEU域内の国別の収集を取りやめEU全体の数字のみを把握することに決めている。そのため、EU28国は今回の調査対象から除外する。

その国の幸福度としては、SDSNが発表している『世界幸福度レポート』ⁱⁱのうち「キャントリルの階梯」ⁱⁱⁱに対する回答の平均値を用いる。

今回は、国民1人あたりのコーヒー年間消費量、その国の幸福度共に2016年度の統計データを用いる。

・データ

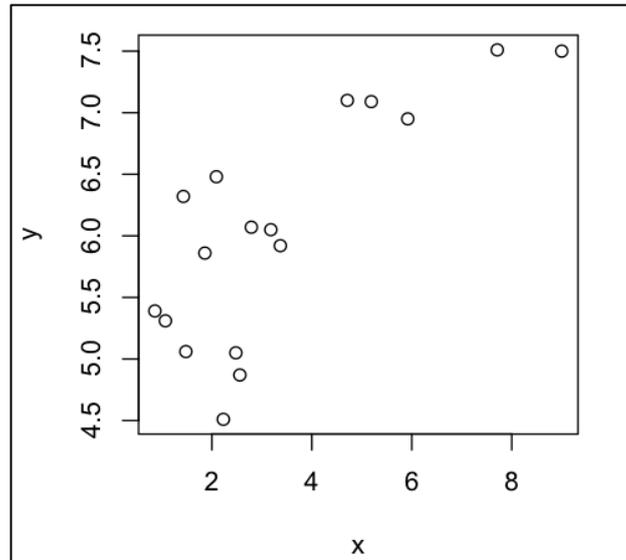
国名	1人あたりのコーヒー消費量 (kg/年)	その国の幸福度
ノルウェー	9	7.498
スイス	7.71	7.509
ブラジル	5.92	6.952
コスタリカ	5.19	7.087
アメリカ合衆国	4.71	7.104
日本	3.37	5.921
ベネズエラ	3.18	6.048
エルサルバドル	2.79	6.068
ホンジュラス	2.56	4.871
チュニジア	2.48	5.045
エチオピア	2.23	4.508
コロンビア	2.09	6.481
ロシア	1.86	5.856
ベトナム	1.48	5.061
グアテマラ	1.43	6.324
インドネシア	1.07	5.314
トルコ	0.86	5.389

・Rのコードと結果

▽回帰線を分析

散布図をプロット

```
> x = c(9.00,7.71,5.92,5.19,4.71,3.37,3.18,2.79,2.56,2.48,2.23,2.09,1.86,1.48,1.43,1.07,0.86)
> y = c(7.50,7.51,6.95,7.09,7.10,5.92,6.05,6.07,4.87,5.05,4.51,6.48,5.86,5.06,6.32,5.31,5.39)
> plot(x,y)
>
```



回帰線を求める

```
> result = lm(y~x)
> summary(result)
```

```
Call:
lm(formula = y ~ x)
```

```
Residuals:
    Min       1Q   Median       3Q      Max
-1.17522 -0.34648  0.07534  0.29289  0.89017
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.97332    0.25924   19.184 5.76e-12 ***
x             0.31924    0.06326    5.046 0.000145 ***
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 0.5937 on 15 degrees of freedom
Multiple R-squared:  0.6293, Adjusted R-squared:  0.6046
F-statistic: 25.47 on 1 and 15 DF, p-value: 0.0001448
```

残差の正規性と等分散性をチェック

```
> shapiro.test(resid(result))
```

```
Shapiro-Wilk normality test
```

```
data: resid(result)  
W = 0.96263, p-value = 0.6814
```

```
> library(car)
```

```
> ncvTest(result)
```

```
Non-constant Variance Score Test
```

```
Variance formula: ~ fitted.values
```

```
Chisquare = 1.043191 Df = 1 p = 0.3070804
```

まず、両者の散布図をプロットしたところ、両者の間には何かしら関係があるように見えた。そこで、国民1人あたりのコーヒー年間消費量を原因 (X)、その国の幸福度を結果(Y)とする回帰式を求めたところ、 $0.05 > p$ となり、 $Y = 4.97 + 0.32X$ という回帰線が引けることがわかった。

その後、残差の正規性と等分散性をシャピロ・ウェルク検定を用いて確認したところ、正規性、等分散性どちらも $p > 0.05$ となり残差には正規性も等分散性もあることがわかった。

これらの結果から、国民1人あたりのコーヒー年間消費量は、その国の幸福度に対して効いていると言える。

・結論と考察

今回の分析による結果では、国民1人あたりのコーヒー年間消費量は、その国の幸福度に対して効いている ($Y = 4.97 + 0.32X$) と結論付けられる。

しかし、データを取得できたサンプル数の少なさ、幸福度を測る基準の曖昧さ、またその他の要因 (GDP などといった第3の要因) が影響を与えている可能性を否定できないことから両者に因果関係があると言うのにはまだ不十分であると思われる。

また、今回取得できたデータはあくまで「国民」あたりのデータであり、個人間でのコーヒー消費量による幸福度の比較まで追求されていないところが今後の課題になると考えられる。

最後に、今回の分析経験を通じて学んだことは、目的に適したデータを集めることが非常に重要であり、なおかつ困難であるということである。国別のデータ等は既に公開されている場合が多いが、本来の研究したい要因以上に影響を与える要因が多く存在してしまうため参考にしづらいことが多く、またサンプル数も限りがあることが多い。そのため、基本的には自分でデータを集めることが正しい分析結果を導き出すためには重要なのではないかと私は考察する。

i http://coffee.ajca.or.jp/wp-content/uploads/2017/11/data09_201711.pdf 2018年6月24日にアクセス

ii <https://rocketnews24.com/2016/09/21/802324/> 2018年6月24日にアクセス

iii 「ハシゴを想像してみてください。ハシゴの各段には数字が振ってあり、ハシゴを上るにつれて数字は大きくなっていきます。最下段は“0”で、最上段は“10”です。最上段はあなたにとって“最高の人生”で、最下段は“最低の人生”です。今現在、あなたはハシゴの何段目に立っていると思いますか？」という内容の質問