

検定と信頼区間: 一般論

- 黒木玄
- 2022-05-31~2022-05-31

このノートでは[Julia言語 \(https://julialang.org/\)](https://julialang.org/)を使用している:

- [Julia言語のインストールの仕方の一例 \(https://nbviewer.org/github/genkuroki/msfd28/blob/master/install.ipynb\)](https://nbviewer.org/github/genkuroki/msfd28/blob/master/install.ipynb)

自明な誤りを見つけたら、自分で訂正して読んで欲しい。大文字と小文字の混同や書き直しが不完全な場合や符号のミスは非常によくある。

このノートに書いてある式を文字通りにそのまま読んで正しいと思ってしまうとひどい目に会う可能性が高い。しかし、数が使われている文献には大抵の場合に文字通りに読むと間違っている式や主張が書いてあるので、内容を理解した上で訂正しながら読んで利用しなければいけない。実践的に数学を使う状況では他人が書いた式をそのまま信じていけない。

このノートの内容よりもさらに詳しいノートを自分で作ると勉強になるだろう。膨大な時間を取られることになるが、このノートの内容に関係することで飯を食っていく可能性がある人にはそのためにかけた時間は無駄にならないと思われる。

目次

- ▼ [1 P値と検定と信頼区間の関係](#)
 - [1.1 P値およびP値関数の定義](#)
 - [1.2 P値を使った検定](#)
 - [1.3 P値関数を使った信頼区間](#)
 - [1.4 P値関数が「よい」かどうかの判断基準達](#)
 - [1.5 信頼区間と検定の表裏一体性](#)
- ▼ [2 二項分布モデルの場合](#)
 - [2.1 Clopper-Pearsonの信頼区間](#)
 - [2.2 Waldの信頼区間](#)
 - [2.3 Wilsonの信頼区間](#)
 - [2.4 Sterneの信頼区間](#)
- [3 よくある誤解](#)

In [1]:

```
1 ENV["LINES"], ENV["COLUMNS"] = 100, 100
2 using BenchmarkTools
3 using Distributions
4 using LinearAlgebra
5 using Printf
6 using QuadGK
7 using Random
8 Random.seed!(4649373)
9 using Roots
10 using SpecialFunctions
11 using StaticArrays
12 using StatsBase
13 using StatsFuns
14 using StatsPlots
15 default(fmt = :png, titlefontsize = 10, size = (400, 250))
16 using SymPy
```

```
In [2]: 1 # Override the Base.show definition of SymPy.jl:
2 # https://github.com/JuliaPy/SymPy.jl/blob/29c5bfd1d10ac53014fa7fef468bc8deccadc2fc/src/types.
3
4 @eval SymPy function Base.show(io::IO, ::MIME"text/latex", x::SymbolicObject)
5     print(io, as_markdown("\displaystyle " * sympy.latex(x, mode="plain", fold_short_frac=false)
6 end
7 @eval SymPy function Base.show(io::IO, ::MIME"text/latex", x::AbstractArray{Sym})
8     function toeqnarray(x::Vector{Sym})
9         a = join(["\displaystyle " * sympy.latex(x[i]) for i in 1:length(x)], "\\")
10        "\"\\left[ \\begin{array}{r}$a\\end{array} \\right]\""
11    end
12    function toeqnarray(x::AbstractArray{Sym,2})
13        sz = size(x)
14        a = join([join("\displaystyle " .* map(sympy.latex, x[i,:]), "&") for i in 1:sz[1]],
15        "\"\\left[ \\begin{array}{c} " * repeat("r",sz[2]) * "}" * a * "\\end{array}\\right]\""
16    end
17    print(io, as_markdown(toeqnarray(x)))
18 end
```

```
In [3]: 1 safemul(x, y) = x == 0 ? x : x*y
2 safediv(x, y) = x == 0 ? x : x/y
3
4 x ≲ y = x < y || x ≈ y
5
6 mypdf(dist, x) = pdf(dist, x)
7 mypdf(dist::DiscreteUnivariateDistribution, x) = pdf(dist, round(Int, x))
8
9 distname(dist::Distribution) = replace(string(dist), r"{.*}" => "")
10 myskewness(dist) = skewness(dist)
11 mykurtosis(dist) = kurtosis(dist)
12 function standardized_moment(dist::ContinuousUnivariateDistribution, m)
13     μ, σ = mean(dist), std(dist)
14     quadgk(x -> (x - μ)^m * pdf(dist, x), extrema(dist)...)[1] / σ^m
15 end
16 myskewness(dist::MixtureModel{Univariate, Continuous}) = standardized_moment(dist, 3)
17 mykurtosis(dist::MixtureModel{Univariate, Continuous}) = standardized_moment(dist, 4) - 3
```

Out[3]: mykurtosis (generic function with 2 methods)

1 P値と検定と信頼区間の関係

1.1 P値およびP値関数の定義

1.2 P値を使った検定

1.3 P値関数を使った信頼区間

1.4 P値関数が「よい」かどうかの判断基準達

1.5 信頼区間と検定の表裏一体性

2 二項分布モデルの場合

2.1 Clopper-Pearsonの信頼区間

データ「 n 回中 k 回成功」について、「データの数値以上に極端な」の意味を「 k 以上または k 以下の」とした場合。

2.2 Waldの信頼区間

データ「 n 回中 k 回成功」について、「データの数値以上に極端な」の意味を「 k 以上または k 以下の」とした場合。ただし、二項分布の中心極限定理(正規分布近似)と分散が少しことなる正規分布による近似を使う。

2.3 Wilsonの信頼区間

データ「 n 回中 k 回成功」について, 「データの数値以上に極端な」の意味を「 k 以上または k 以下の」とした場合. ただし, 二項分布の中心極限定理(正規分布近似)のみを使う.

2.4 Sterneの信頼区間

「データの数値以上に極端な」の意味を「生成される確率がデータの数値以下の」とした場合.

3 よくある誤解

$\sqrt{x}/\sqrt{2}$

In []:

1	
---	--