YES／NO実験の統計処理

豊橋技術科学大学　梅村恭司

話の流れ

- <u>背景</u>
- 少ない数のとき，
- 多い数のとき，
- 統計検定
- 統計検定の閾値の根拠
- パーフェクトでないとき
- 知っておくべきこと

背景1

YES/NO実験とは，

実験結果が測定値ではなくて，YES/NOであるもの．

例：

アンケートによるシステム効果の評価

対戦結果の解釈

背景２

YES/NO実験とは，

実験結果が測定値ではなくて，YES/NOであるもの．

実験結果から結論を導くには，何回実験したらよいかについて考えてみたい。

話の流れ

- 背景
- 少ない数のとき，
- 多い数のとき，
- 統計検定
- 統計検定の閾値の根拠
- パーフェクトでないとき
- 知っておくべきこと

例

心理実験ををして，被験者４人が，このサンプルの２つでは４人とも色の差がみえたという回答を得ました.

したがって，このサンプルの間は人間にとって色の差がみえると考えられます．．．．

大丈夫でしょうか？考えてみてください。

１:大丈夫，２:だめ，３:よくわかない

似たような例

弟と将棋を４回指していて，４回とも僕が勝ちました.

したがって，私のほうが弟よりも将棋が強いと考えられます.

大丈夫でしょうか？考えてみてください。

１:大丈夫，２:だめ，３:よくわかない　→挙手

さらに似たような例

弟とジャンケンを4回して，４回とも僕が勝ちました.

したがって，私のほうが弟よりもジャンケンが強いと考えられます.

大丈夫でしょうか？考えてみてください。

１:大丈夫，２:だめ，３:よくわかない　→挙手

さらに似たような例

弟とジャンケンを４回して，４回とも僕が勝ちました．

したがって，私のほうが弟よりもジャンケンが強いと考えられます．

大丈夫でしょうか？考えてみてください。

１：大丈夫，２：だめ，３：よくわからない　→挙手

議論の時間

回答　1，2，3は，問題によってかわると考えるのでないかと思います。

どうしてかわると考えたのか考えましょう。

将棋の例

「将棋は，有意な実力差があると，勝ち負けはめったなことではひっくり返らないゲームです．また，たいていの場合，有意な実力差があります．」

という事前知識があるので，人は4回でも結論を導くのですが，

一般には，このような事前知識を納得してもらうこと，／証明することは難しいことと思いませんか？

話の流れ

- 背景
- 少ない数のとき，
- 多い数のとき，
- 統計検定
- 統計検定の閾値の根拠
- パーフェクトでないとき
- 知っておくべきこと

例

心理実験ををして，被験者１００人が，このサンプルの２つでは１００人とも色の差がみえたという回答を得ました.

(補足:この100人は人間全体から偏りなく選ばれたと考えてください。）

したがって，このサンプルの間は人間にとって色の差がみえると考えられます．．．．

大丈夫か．．．たぶん，そう

似たような例

弟と将棋を１００回指してして，１００回とも僕が勝ちました.

したがって，私のほうが弟よりも将棋が強いと考えられます.

大丈夫か．．．．．．たぶん，そう

さらに似たような例

弟とジャンケンを１００回して，１００回とも僕が勝ちました.

したがって，私のほうが弟よりもジャンケンが強いと考えられます.

大丈夫か．．．．．．たぶん，そう

４回と１００回の違いは何か

ジャンケンは，正常な条件ならば人の間に差がないゲームです．

しかしながら，１００回もつづけて，一方が勝つということが起きると，人間は片方がずるをしている（ゲームの前提がちがう）のではないかと疑います．

なぜでしょう．．．

１００回のジャンケン

「ジャンケンは，正常な条件ならば人の間に差がないゲームです．」

これが真実のとき，１００回やって，１００勝 0敗の起きる確率は？

それは具体的にいくつくらい？

１００回のジャンケン

「ジャンケンは，正常な条件ならば人の間に差がないゲームです．」

これが真実のとき，１００回やって，１００勝 0敗の起きる確率は？

それは具体的にいくつくらい？

答え：$2^{-100} \sim 10^{-30}$ 参考：100年は3153600000 秒

１００回のジャンケン

「ジャンケンは，正常な条件ならば人の間に差がないゲームです．」

という事前知識があっても，その事前知識が成立すると仮定すると，観測された事象が起きる確率が，生きている間に遭遇しないほど小さいので，人間は起きないはずのことが起きたと考え，事前知識が成立しない理由があると推論します．

話の流れ

- 背景
- 少ない数のとき,
- 多い数のとき,
- <u>統計検定</u>
- 統計検定の閾値の根拠
- パーフェクトでないとき
- 知っておくべきこと

問題設定ページワン

ジャンケンや将棋は事前の知識があると混乱するので，ページワン（トランプゲーム）のように，実力と運との両方があるゲームの勝敗を考えます．

状況：

弟とページワンの試合を8回やって，私は8回勝ちました．さて，私のほうがページワンがうまいでしょうか？

仮説をたてる

H_1:「弟より私のほうがページワンがうまい」ことを否定し，

H_0 :「弟より私がページワンがうまいとはいえない」という仮説をたてます。

「８回試合をして私は８回とも勝った」という結果

H_0が正しいとするとで，観測結果が起きる確率pは，最大が，実力が同じときで，そのとき0.0039(= 1 / 256)です．．．

用語

H_1:「弟より私のほうがページワンがうまい」

H_O :「弟より私がページワンがうまいとはいえない」

「８回試合をして私は８回とも勝った」という観測

H_Oが正しいと仮定したとき，観測結果が起きる確率pの最大値をp-value, p-valueでH_Oを疑うかどうかの基準値0.01を危険率α，H_Oを帰無仮説（きむかせつ Null hypothesis），H_1を対立仮説（Alternative hypothesis）と呼びます．

「有意」

H_1:「弟より私のほうがページワンがうまい」

H_0 :「弟より私がページワンがうまいとはいえない」という仮説

「８回試合をして私は８回とも勝った」という観測

H_0が正しいと仮定したとき，観測結果が起きる確率の最大値をpとし，危険率αと比較して，pがαより小さいとき，H_0は棄却されるという．また，H_1がα水準で有意であるという。

具体的には，0.01水準で，「弟より私がページワンがうまい」ということが有意であるという．

話の流れ

- 背景
- 少ない数のとき,
- 多い数のとき,
- 統計検定
- 統計検定の閾値の根拠
- パーフェクトでないとき
- 知っておくべきこと

0.01という基準は大きすぎるか？

100回に1回に結論が間違っていることが偶然に起きるえるのは，困るという立場があるのは当然のこと，，，

では，何回に1回で良いか考えてみてください。

0.01という基準は大きすぎるか？

100回に1回に結論が間違っていることが偶然に起きるえるのは，困るという立場があるのは当然のこと，，，

では，何回に1回で良いか考えてみてください。

一概には決められない。その結論の使い方（応用）によって，異なる。

情報を記録する応用

いま，ある情報を図書館に記録するのに足りるかどうかという状況を考える.

実際に効果があるかは疑わしいが，再試験することに耐える確率はいくつとするか？

確率は小さいほうがよいが，実際には実験のコストがあるので，社会合意できる値を決める.

その答え．α=0.01

可能性を示唆するときには，α=0.05

最低限の実験回数

- 結果が明白で，実験者には何度実験しても結論が変わらないとしても，結果が明白かどうかを査読者に納得させるのは難しい．むしろ，実験回数を増やしたほうが早道である．
- 結論を示唆するとき0.05： すくなくとも5回
- 結論を主張するとき0.01： 少なくとも7回
- 著者には将棋のように明らかでも，査読者は先入観なくジャンケンのように考えることがルール．

話の流れ

- 背景
- 少ない数のとき，
- 多い数のとき，
- 統計検定
- 統計検定の閾値の根拠
- パーフェクトでないとき
- 知っておくべきこと

パーフェクトでないとき

H_1:「弟より私のほうがページワンがうまい」

H_0 :「弟より私がページワンがうまいとはいえない」という仮説

「１３回試合をして私は１１勝２敗」という観測結果

H_0が成立すると観測より起こりにくいケースの「12勝１敗，１３勝０敗」がある。

パーフェクトでないとき

H_1:「弟より私のほうがページワンがうまい」

H_0 :「弟より私がページワンがうまいとはいえない」という仮説

「１３回試合をして私は１１勝２敗」という観測結果

観測結果の起きる確率だけでなくて，観測結果より，H_0が成立すると起こりにくいケースの「12勝１敗，１３勝0敗」の確率も合算して，H_0の危険率とする。

具体的な例

H_1:「弟より私のほうがページワンがうまい」

H_O:「弟より私がページワンがうまいとはいえない」という仮説

「１３回試合をして私は１１勝２敗」という観測結果

13勝0敗，12勝１敗，11勝2敗の確率も合計する．

$p = (1+13+13\times6)/8192$ で，これはおよそ0.011

0.01水準では有意でないが，0.05水準では有意

話の流れ

- 背景
- 少ない数のとき，
- 多い数のとき，
- 統計検定
- 統計検定の閾値の根拠
- パーフェクトでないとき
- <u>知っておくべきこと</u>

よくあるひどい間違い

H_1:「弟より私のほうがページワンがうまい」

H_0:「弟より私がページワンがうまいとはいえない」という仮説

仮説が棄却できなかったときにH_0が成立しているかのような主張をする.

なぜ，これが間違いだと思われるか

よくあるひどい間違い

H_1:「弟より私のほうがページワンがうまい」

H_O:「弟より私がページワンがうまいとはいえない」という仮説

仮説が棄却できなかったときにH_Oが成立しているかのような主張をする.

ただしくは，なにもいえないというべきもの。

まとめ

- 背景
- 少ない数のとき，（事前知識に依存）
- 多い数のとき，　（事前知識に依存しない）
- 統計検定（仮説，危険率）
- 統計検定の閾値の根拠（情報の登録の応用）
- パーフェクトでないとき（より差が大きいケース）
- 知っておくべきこと（棄却できないとき）