Mineração Egoísta - P Value

Ivan Sendin

FACOM - Universidade Federal de Uberlândia
ivansendin@yahoo.com,sendin@ufu.br

27 de agosto de 2024

- Experimento: testar a eficácia de um adubo
  Poderia ser um remédio, vacina...qualquer coisa
- Esse adubo é melhor que o antigo
  ou sem adubo

- 5 plantas com adubo: C_i
- 5 plantas sem adubo: S_i
- Comparo a "produção"

$$C_i > S_i$$

- (Vamos assumir que nunca será igual)

- Vamos dizer que para $i = 2, 3$ e 5 temos

$$C_i > S_i$$

- Isso é bom ?
- Parece que sim!

- E se alguem argumentar que o adubo é inerte ... ele nao tem efeito algum e voce teve sorte?
- Parece uma afirmação razoável....

- São 5 casos unitários que podem valer 0 ou 1
- Entao eu conto quantas strings:00000,00001,00010,...11111
- 3 ou mais digitos 1
  "O adubo testado ganha"
- São 16 de 32
- Então, a probabilidade de obter este **tipo** de resultado é de 1/2 se vc não fizer nada

- E se $C_i > S_i$ para todos??
- Esse produto é 100% !
- $1/32 \approx 3\%$
- "Existe uma chance de 3% do adubo testado ganhar...mesmo sendo igual ao outro"

- Fisher/1925: Teste de significancia da hipótese nula
- Separar o ruído
- hipótese nula: não efeito, não relação,...
- "calcula a probabilidade(valor p) de encontrar um efeito igual ou maior que o observado com a hipótese nula verdadeira"
- O valor de corte é 0.05 é arbitrário
- Teste de Significância
  $\neq$ Efeito

- Multiplos testes...
- Faço um experimento com uma determinada quantidade de adubo...
- Faço um experimento com mais adubo...
- Faço um experimento com menos adubo...
- Faço um experimento com uma variedade diferente de semente...
- ....
- Faz sentido??

- Multiplos testes...
- Faço um experimento iniciando na lua cheia
- Faço um experimento iniciando em dia ímpar
- O funcionario que faz o experimento tem CPF número primo...

- Multiplos testes...
- Ignoro os experimentos que eu nao quero
- E publico resultado bom (para mim)
- Como existe o imponderável....**testes multiplos vão produzir algo de seu interesse...um dia**
- Em especial no mundo digital

- Multiplos testes...
- Em uma turma de 60 alunos, era provavel que alguem iria encontrar a mineração egoista...mesmo que ela não existisse

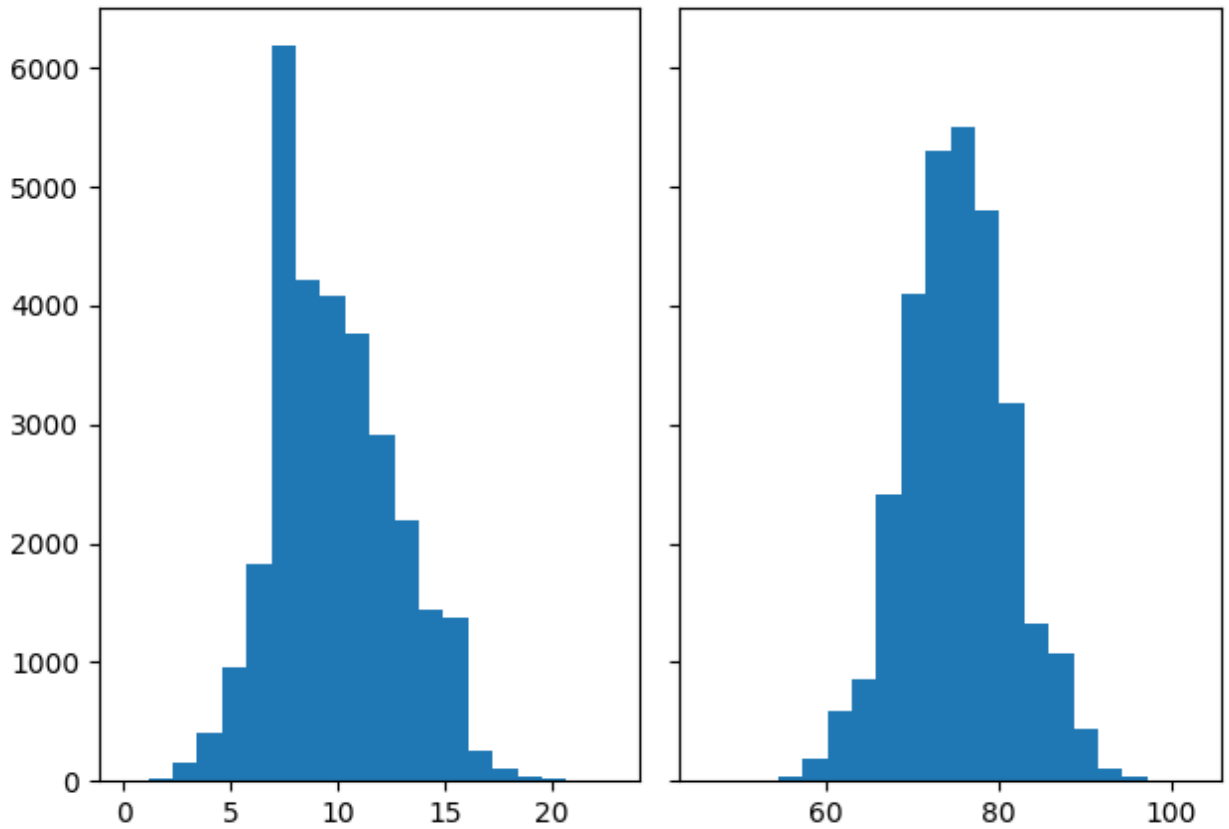
- The Extent and Consequences of P-Hacking in Science
- Why Most Published Research Findings Are False
- Uma Senhora tomando chá

- O mineração deveria ser *incentive compatible*
- O comportamento honesto deveria ser o caminho para maximizer o lucro
- Mineração Egoista
- Pistas na blockchain
- Mineração em seguida

- Dado um conjunto de blocos quereos saber se existe alguem fazendo mineração egoista
- Usando os testes, eu nunca vou “saber”

- É razoável dizer que quanto maior o poder...mais blocos...
- Mais minerações em seguida ocorrerão naturalmente
- Posso inferir o poder....
- Quem tem $x\%$ dos blocos tem $x\%$ do poder computacional

- Se alguem produz 10% dos blocos...quantas minerações em seq ele deve produzir ?
- Não tem um única resposta....é um **processo estocástico**
- Pode ser calculado
- ou...



Minerações em seq. para 10% e 20% - 30000 testes 1000 minerações

- Alguem com 10% dos blocos produziu 10 minerações em sequencia
  Parece razoável...voce acredita, não se espanta
- Alguem com 10% dos blocos produziu 19 minerações em sequencia
  Será ?? Tem algum truque!!
  Não que seja impossível!!

- Dado um vetor de minerações
- Esta "implicito" o poder de cada minerador
- Este vetor tem algumas minerações em sequencia...
- quando eu embaralho o vetor
- O poder do minerador é mantido
- E tenho uma saida de mineração em sequencia
- Repete...repete...

- Seja V um vetor ordenado com o numero de mineracoes em sequencia
  De um determinado m
  Obtido conforme a pagina anterior
- Digamos que são 100 experimentos
  $len(V) = 100$
- $\ldots, 10, 10, 11, 12, 14, 15]$
- Se nos dados reais existem 12 minerações em sequencia para m
- O $p - value$ é de 0.03
- Em 3% "das mineracoes geradas alatoriamente, com o poder aferido de m são tão ou mais extremas que o observado"

```
K=500
mes=1
podermes = calcPower(b,mes*blocosmes,(mes+1)*blocosmes)
for x in range(1000):
    print(x)
    random.seed(x)
    selfishsimulation,real = test(b[mes*blocosmes:(mes+1)*blocosmes],K)
    for minerador in podermes.keys():
        pwr = podermes[minerador]/blocosmes
        pv = 1-selfishsimulation.get(minerador,0)/K
        if pwr>0.1 and  pv<=0.05:
            print('\t',minerador, pv )
```

```
1
2
        37 0.010000000000000009
3
...
23
        37 0.038000000000000034
...
32
        37 0.0040000000000000036
33
        41 0.04600000000000004
34
35
        37 0.0040000000000000036
36
        37 0.028000000000000025
        41 0.038000000000000034
37
        32 0.0040000000000000036
38
39
...
```

```
988
989
990
          37 0.0020000000000000018
991
992
993
          32 0.008000000000000007
994
995
          41 0.0
996
997
```