

```
In [8]: import scipy
import pandas as pd
from scipy.stats import chi2_contingency

def rv_from_uni(p): return int(uni() < p)
```

ЗАДАНИЕ

- Написать программы, реализующие рассмотренные методы имитации случайных событий.
- Провести статистические исследования полученных результатов..
- Проверить гипотезы о соответствии полученных результатов требуемым.

```
In [9]: from numpy import random

seed = 1
random.seed(1)
uni = random.uniform
n = 10000
```

1) Имитация сложного события(события независимы)

Даны две независимые дискретные случайные величины(события) $A, B, A \sim Bern(p_A), B \sim Bern(p_B)$

```
In [32]: P_A = .1
P_B = .5
```

Функция генерации события

```
In [33]: def complex_event(p): return rv_from_uni(p)
```

Генерация данных

```
In [34]: A = [complex_event(P_A) for _ in range(n)]
B = [complex_event(P_B) for _ in range(n)]
df = pd.DataFrame({'A': A, 'B': B})
crosstab = pd.crosstab(index=df['A'], columns=df['B'])
crosstab
```

Out[34]:

	B	0	1
A			
0	4450	4534	
1	518	498	

Проверка гипотезы

Для проверки гипотезы воспользуемся χ^2 тестом на независимость для категориальных данных. В качестве реализации используем $\chi^2_{contingency}$ из пакета *scipy*. В качетсве H_0 гипотезы будем считать зависимость случайны величин

```
In [36]: chi2, p, dof, ex = chi2_contingency(crosstab)
print(f'p-value={p}')
```

p-value=0.39859797396360463

Т.к. p-value достаточно велико, мы можем отвергнуть гипотезу о зависимости СВ.

2) Имитация сложного события, состоящего из зависимых событий.

Даны зависимые дискретные случайные величины(события) $A, B, A \sim Bern(p_A), B \sim Bern(p_B), P(A|B) = p_{A|B}, P(B|A) = p_{B|A}$

```
In [189]: P_A = .5
P_B_A = 0.3
```

```
In [190]: def complex_dependent_event():
A = rv_from_uni(P_A)
B = rv_from_uni(P_B_A) if A else rv_from_uni(1 - P_B_A)
return [A, B]
```

```
In [191]: data = np.array([complex_dependent_event() for _ in range(n)])
df = pd.DataFrame({'A': data[:,0], 'B': data[:,1]})
crosstab = pd.crosstab(index=df['A'], columns=df['B'])
crosstab
```

Out[191]:

	B	0	1
A			
0	1478	3486	
1	3512	1524	

Проверка гипотезы

Для проверки гипотезы воспользуемся χ^2 тестом на независимость для категориальных данных. В качестве реализации используем $\chi^2_{contingency}$ из пакета *scipy*. В качетсве H_0 гипотезы будем считать зависимость случайны величин

```
In [192]: chi2, p, dof, ex = chi2_contingency(crosstab)
print(f'p-value={p}')
```

p-value=0.0

p -value стремится к нулю, что позволяет нам отвергнуть нулевую гипотезу. Зависимость СВ доказана

3) Имитация событий, составляющих полную группу

Даны три независимые дискретные случайные величины(события) $A, B, C, A \sim Bern(p_A), B \sim Bern(p_B), C \sim Bern(p_C)$

```
In [193]: P_A = .3
P_B = .3
P_C = .4
```

```
In [194]: #A, B, C = 0, 1, 2
def triple_complex_event():
t = uni()
return 0 if t < P_A else 1 if P_A < t < P_A + P_B else 2
```

```
In [196]: data = [triple_complex_event() for _ in range(10)]
df = pd.DataFrame({'state': data})
crosstab = df.groupby('state').size()
crosstab
```

Out[196]:

state	
0	3
1	1
2	6
dtype:	int64

```
In [197]: chi2, p, dof, ex = chi2_contingency(crosstab)
print(f'p-value={p}')
```

p-value=1.0

p -value стремится к 1, что говорит в пользу нулевой гипотезы о независимости СВ

```
In [ ]:
```