

Napisać losowy generator danych działający zgodnie z rozkładem reprezentowanym przez daną sieć Bayes'a.

Sieć powinna opisywać zależności między zmiennymi losowymi (binarne: 0/1)

Na wejściu dany jest opis grafu połączeń/zależności między zmiennymi oraz tabele prawdopodobieństw warunkowych.

Problem:

	Studia		Samotność
T	0.64	T	0.15
F	0.36	F	0.85

Studia	Samotność	Zajadanie smutków
T	T	0.85
T	F	0.3
F	T	0.65
F	F	0.15

Zajadanie smutków	Depresja
T	0.7
F	0.2

Testy:

Liczba próbek = 100

```
From data:
{'studies': 0.59, 'loneliness': 0.18, 'overeating': 0.37, 'depression': 0.4}
From distribution:
{'studies': 0.64, 'loneliness': 0.15, 'overeating': 0.3258, 'depression': 0.3629}
Difference:
{'studies': 0.050000000000000044, 'loneliness': 0.03, 'overeating': 0.04420000000000002, 'depression': 0.03710000000000002}
```

```
P(depression|overeating) from data:
0.7000000000000001
P(depression|overeating) from distribution:
0.6284375861118764
P(depression|overeating) difference
0.07156241388812368
```

Liczba próbek = 1000

```
From data:
{'studies': 0.662, 'loneliness': 0.153, 'overeating': 0.351, 'depression': 0.388}
From distribution:
{'studies': 0.64, 'loneliness': 0.15, 'overeating': 0.3258, 'depression': 0.3629}
Difference:
{'studies': 0.02200000000000002, 'loneliness': 0.003000000000000027, 'overeating': 0.0252, 'depression': 0.02510000000000001}
```

```
P(depression|overeating) from data:
0.6417525773195876
P(depression|overeating) from distribution:
0.6284375861118764
P(depression|overeating) difference
0.01331499120771118
```

Liczba próbek = 10000

```
From data:
{'studies': 0.6299, 'loneliness': 0.1477, 'overeating': 0.3269, 'depression': 0.3644}
From distribution:
{'studies': 0.64, 'loneliness': 0.15, 'overeating': 0.32580000000000003, 'depression': 0.3629}
Difference:
{'studies': 0.010099999999999998, 'loneliness': 0.0022999999999999965, 'overeating': 0.0010999999999999899, 'depression': 0.0015000000000000013}
```

```
P(depression|overeating) from data:
0.6267837541163557
P(depression|overeating) from distribution:
0.6284375861118766
P(depression|overeating) difference
0.001653831995520938
```

Sprawdziłem poprawność danych poprzez oszacowanie prawdopodobieństw $P(\text{Studia})$, $P(\text{Samotność})$, $P(\text{Zajadanie smutków})$, $P(\text{Depresja})$, $P(\text{Depresja} | \text{Zajadanie smutków})$ i porównanie ich z tymi z rozkładu.

Powyższe testy wskazują, że dane są zgodne z rozkładem.