

# Raport 5

Magdalena Potok

2023-05-20

## Wstęp

Poniższy raport przedstawia porównanie dokładności wyników przy konstruowaniu przedziałów ufności za pomocą metody Agrestiego-Coulla oraz metody klasycznej. Symulacje zostały przeprowadzone dla różnych rozmiarów prób oraz trzech różnych wartości prawdopodobieństwa w rozkładzie Bernoulliego. W każdym z przypadków poziom ufności wynosi 95%.

Metoda klasyczna zakłada że rozkład danych jest zbliżony do rozkładu normalnego. Wykorzystuje ona estymator, który jest obliczany na podstawie średniej oraz odchylenia standardowego. Przedział ufności konstruuje za pomocą błędu standardowego oraz wartości krytycznej, jej wielkość zależy od dobranego poziomu ufności (w tym raporcie 95%). Przedział jest postaci "średnia próbki  $\pm$  wartość krytyczna \* SE".

Metoda Agrestiego-Coulla względnie lepsze dopasowanie przedziału ufności dla skrajnych wartości prawdopodobieństwa. Korzysta ona z korekty Laplace'a oraz dodatkowej stałej, aby poprawić to oszacowanie. Przedział jest obliczany za pomocą estymacji proporcji i błędu standardowego.

**p = 0.9**

### Metoda Agrestiego-Coulla

##	Pocz_przedział	Kon_przedział	Dług_przedział	Estymator	Wielk_próby	Praw_wart
## 1	0.65	1.08	0.43	0.86	10	0.9
## 2	0.62	0.97	0.35	0.79	20	0.9
## 3	0.77	1.00	0.23	0.88	30	0.9
## 4	0.86	1.01	0.15	0.93	40	0.9
## 5	0.83	0.99	0.16	0.91	50	0.9
## 6	0.75	0.94	0.18	0.84	60	0.9
## 7	0.69	0.88	0.19	0.78	70	0.9
## 8	0.81	0.95	0.14	0.88	80	0.9
## 9	0.84	0.97	0.12	0.90	90	0.9
## 10	0.87	0.98	0.10	0.92	100	0.9

### Metoda klasyczna

##	Pocz_przedział	Kon_przedział	Dług_przedział	Estymator	Wielk_próby	Praw_wart
## 1	1.00	1.00	0.00	1.00	10	0.9
## 2	0.69	1.01	0.31	0.85	20	0.9
## 3	0.84	1.02	0.18	0.93	30	0.9
## 4	0.93	1.02	0.10	0.98	40	0.9
## 5	0.87	1.01	0.13	0.94	50	0.9
## 6	0.78	0.95	0.17	0.87	60	0.9
## 7	0.71	0.89	0.19	0.80	70	0.9
## 8	0.83	0.97	0.13	0.90	80	0.9
## 9	0.87	0.98	0.11	0.92	90	0.9
## 10	0.89	0.99	0.09	0.94	100	0.9

Częstość zawierania się prawdziwej wartości w przedziale ufności za pomocą metody Agrestiego-Coulla:

```
## [1] 90
```

Częstość zawierania się prawdziwej wartości w przedziale ufności za pomocą metody klasycznej:

```
## [1] 70
```

Analizując powyższe wyniki można zauważyć pewnie różnice. W metodzie Agrestiego-Coulla widać, że przedziały ufności są nieco szersze, niż w metodzie klasycznej oraz częściej zawierają one rzeczywistą wartość estymatora. To wynika z faktu, że ta metoda uwzględnia lepsze dopasowania przedziałów ufności dla skrajnych wartości.

**p = 0.5**

#### Metoda Agrestiego-Coulla

##	Pocz_przedział	Kon_przedział	Dług_przedział	Estymator	Wielk_próby	Praw_wart
## 1	0.19	0.81	0.62	0.50	10	0.5
## 2	0.24	0.68	0.44	0.46	20	0.5
## 3	0.32	0.68	0.36	0.50	30	0.5
## 4	0.24	0.54	0.30	0.39	40	0.5
## 5	0.32	0.60	0.28	0.46	50	0.5
## 6	0.36	0.61	0.25	0.48	60	0.5
## 7	0.41	0.64	0.23	0.53	70	0.5
## 8	0.38	0.60	0.22	0.49	80	0.5
## 9	0.39	0.59	0.21	0.49	90	0.5
## 10	0.39	0.59	0.20	0.49	100	0.5

#### Metoda klasyczna

##	Pocz_przedział	Kon_przedział	Dług_przedział	Estymator	Wielk_próby	Praw_wart
## 1	0.19	0.81	0.62	0.50	10	0.5
## 2	0.23	0.67	0.44	0.45	20	0.5
## 3	0.32	0.68	0.36	0.50	30	0.5
## 4	0.22	0.53	0.30	0.38	40	0.5
## 5	0.32	0.60	0.28	0.46	50	0.5
## 6	0.36	0.61	0.25	0.48	60	0.5
## 7	0.41	0.65	0.23	0.53	70	0.5
## 8	0.38	0.60	0.22	0.49	80	0.5
## 9	0.39	0.59	0.21	0.49	90	0.5
## 10	0.39	0.59	0.20	0.49	100	0.5

Częstość zawierania się prawdziwej wartości w przedziale ufności za pomocą metody Agrestiego-Coulla:

```
## [1] 100
```

Częstość zawierania się prawdziwej wartości w przedziale ufności za pomocą metody klasycznej:

```
## [1] 100
```

Dla  $p = 0.5$  można zauważyć, że obie metody mają w miarę podobne wyniki. Obie metody oszacowują wartość środkową między 0 a 1. Przedziały ufności obejmują tę wartość w średnio podobnym stopniu, co oznacza, że obie metody dają podobną pewność co do prawdziwej wartości estymatora. Nadal można zauważyć pewne minimalne różnice w przedziałach ufności, jednak nie wpływa to na różnice w szerokościach przedziałów i obie metody mają podobne długości przedziałów.

$p = 0.1$

#### Metoda Agrestiego-Coulla

##	Pocz_przedział	Kon_przedział	Dług_przedział	Estymator	Wielk_próby	Praw_wart
## 1	0.00	0.56	0.56	0.28	10	0.1
## 2	0.03	0.38	0.35	0.21	20	0.1
## 3	0.02	0.27	0.25	0.15	30	0.1
## 4	0.06	0.30	0.24	0.18	40	0.1
## 5	0.06	0.27	0.21	0.17	50	0.1
## 6	0.04	0.21	0.17	0.12	60	0.1
## 7	0.06	0.23	0.17	0.15	70	0.1
## 8	0.01	0.11	0.10	0.06	80	0.1
## 9	0.03	0.16	0.12	0.10	90	0.1
## 10	0.05	0.17	0.12	0.11	100	0.1

#### Metoda klasyczna

##	Pocz_przedział	Kon_przedział	Dług_przedział	Estymator	Wielk_próby	Praw_wart
## 1	0.00	0.45	0.45	0.20	10	0.1
## 2	0.00	0.31	0.31	0.15	20	0.1
## 3	0.00	0.21	0.21	0.10	30	0.1
## 4	0.04	0.26	0.22	0.15	40	0.1
## 5	0.04	0.24	0.19	0.14	50	0.1
## 6	0.02	0.18	0.15	0.10	60	0.1
## 7	0.05	0.21	0.16	0.13	70	0.1
## 8	0.00	0.08	0.08	0.04	80	0.1
## 9	0.02	0.13	0.11	0.08	90	0.1
## 10	0.03	0.15	0.11	0.09	100	0.1

Częstość zawierania się prawdziwej wartości w przedziale ufności za pomocą metody Agrestiego-Coulla:

```
## [1] 100
```

Częstość zawierania się prawdziwej wartości w przedziale ufności za pomocą metody klasycznej:

```
## [1] 90
```

Metoda Agrestiego-Coulla dla  $p = 0.1$  ponownie daje szersze przedziały ufności niż metoda klasyczna. Prawdopodobieństwo zawierania rzeczywistej wartości jest podobne, jednak, ponieważ A-C przedział jest szerszy, więc dla tej metody to prawdopodobieństwo jest większe - oznacza to, że ta metoda daje większą pewność w oszacowaniu, ale kosztem większego marginesu błędu.

## Podsumowanie

Metoda klasyczna oraz metoda Agrestiego-Coulla są porównywalne, gdy prawdopodobieństwo w rozkładzie Bernoulliego jest bliskie 0.5. W tym przypadku skonstruowane przedziały były najbardziej do siebie zbliżone i najczęściej zawierały rzeczywistą wartość estymatora. Dlatego właśnie dla tego prawdopodobieństwa obie metody mogą być stosowane i dadzą podobne wyniki. Dla prawdopodobieństw skrajnych (takich jak 0.1 lub 0.9) metoda Agrestiego-Coulla będzie tą bardziej odpowiednią. Jest tak ponieważ uwzględnia ona lepsze dopasowanie przedziału ufności poprzez zastosowanie korekty Laplace'a. Co prawda w tej metodzie przedziały były szersze, ale znacznie częściej zawierały rzeczywistą wartość. A więc dla prawdopodobieństw bliskich 0 lub 1 metoda Agrestiego-Coulla jest bardziej dokładna. Metoda klasyczna nie nadaje się do zastosowania gdy rozkład danych nie jest zbliżony do rozkładu normalnego. Jest tak ponieważ ta metoda opiera się właśnie na takich założeniach i w ten sposób konstruuje swój przedział ufności. Gdy to założenie nie będzie spełnione ta metoda może dawać błędne wyniki i powinno się wybrać jakąś inną metodę (np. Agrestiego-Coulla). Jednak gdy rozkład danych będzie nadzwyczajnie nietypowy, bądź próbka będzie bardzo mała (nie wystarczająca ilość obserwacji do wiarygodnego oszacowania parametru), to obie metody mogą dawać błędne wyniki i może być konieczne zastosowanie bardziej zaawansowanych technik statystycznych dla danego rozkładu.