

# Statystyka - Projekt

Mateusz Marczyk & Karol Sewiło

2023-01-26

## Wstęp

Celem projektu jest porównanie działania dwóch algorytmów minimalizacji stochastycznej. Do porównania wybraлиśmy **algorytm poszukiwania przypadkowego (PRS)** oraz **algorytm genetyczny (GA)**.

## Zmienne globalne

```
budget <- 1000
cost <- 100
alfa <- 0.95
```

## Implementację wybranych algorytmów

### Algorytm Pure Random Search (PRS)

```
PRS <- function(lower_bound, upper_bound, f,n) {
  if(length(lower_bound) != length(upper_bound)) stop("Bounds must be equal dimensions")
  p <- integer(length(lower_bound))
  fmin <- 1e7
  for (x in 1:n) {
    for(i in seq_along(p)) {
      p[i] <- runif(1, lower_bound[i], upper_bound[i])
    }
    fmin <- min(fmin, f(p))
  }
  return(fmin)
}
```

### Algorytm Genetyczny (GA)

```
GA <- function(f, n){
  result <- ga(
    type = "real-valued",
    fitness = function(x) (-1)^f(x),
    lower = getLowerBoxConstraints(f),
    upper = getUpperBoxConstraints(f),
    maxiter = 40,
    popsize = 25)
  return(result$fitnessValue)
}
```

### Funkcja obliczania przedziału ufności

```
CONF_INTRV <- function(vector, alfa)
  return((mean(c(mean(vector),quantile(vector, 1-(alfa/2)))*sd(vector)/sqrt(length(vector)), mean(vector)*quantile(
    vector, 1-(alfa/2))*sd(vector)/sqrt(length(vector)))))
```

## Wybrane funkcje

Do testów wybraлиśmy funkcję **Ackley'a** oraz **Rastrigin'a**. Obie są skalarnymi wielomodalnymi oraz akceptują dane w różnych wymiarach.

## Opracowanie danych

Na początku analizować będziemy wyniki obu algorytmów dla funkcji Ackley'a w odpowiednio: 2, 10 oraz 20 wymiarach. Następnie przeprowadzimy analogiczną analizę dla funkcji Rastrigin'a.

## Hipoteza zerowa

- Algorytmy nie różnią się efektywnością szukania minimum funkcji.

## Hipotezy alternatywne

- Algorytm PRS jest efektywniejszy w szukaniu minimum funkcji.
- Algorytm genetyczny jest efektywniejszy w szukaniu minimum funkcji.

## Funkcja Ackley'a

Wzór funkcji:

$$f(\mathbf{x}) = -a \cdot \exp\left(-\frac{1}{4} \sqrt{\sum_{i=1}^d x_i^2}\right) - \exp\left(\frac{1}{4} \sum_{i=1}^d \cos(2\pi x_i)\right) + a + \exp(1)$$

Minimum globalne:

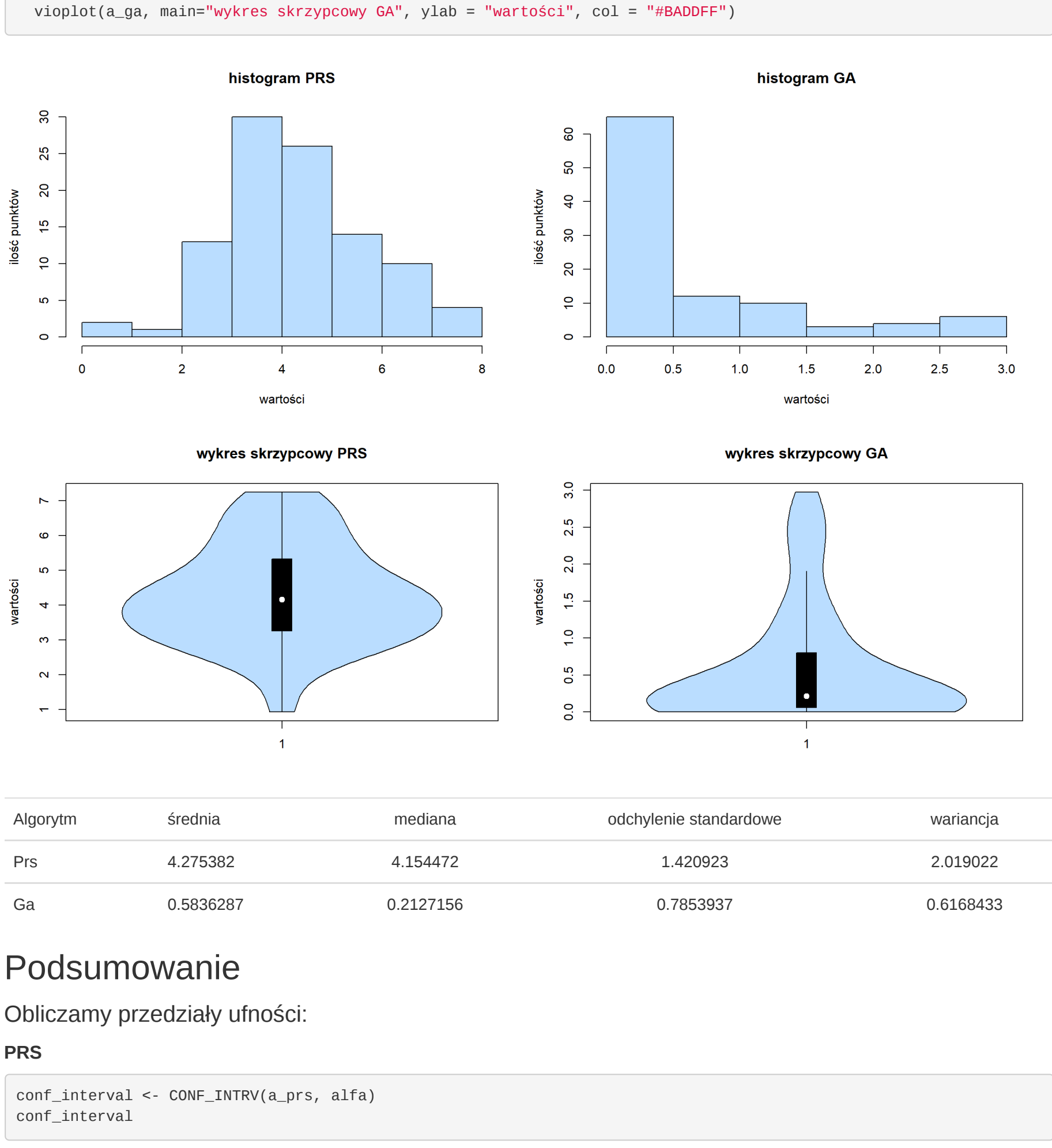
$$f(\mathbf{x}^*) = 0, \text{ at } \mathbf{x}^* = (0, \dots, 0)$$

Zalecana dziedziła poszukiwań:

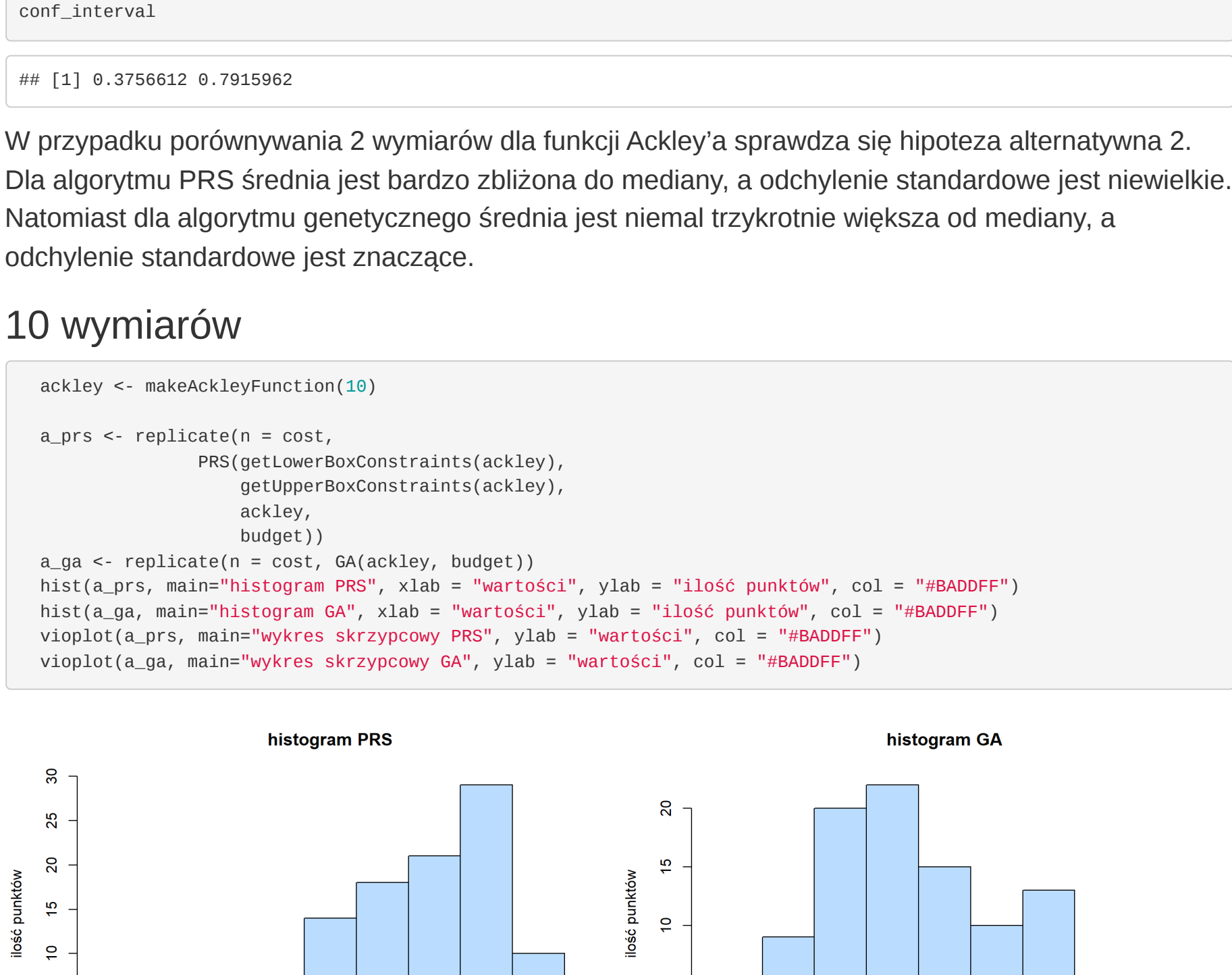
$$x_i \in [-32.768, 32.768], \forall i \leq d, \text{ gdzie } d \text{ to wymiar przestrzeni}$$

Domyślne stałe:

$$a = 20, b = 0.2, c = 2\pi$$



## 2 wymiary



Algorytm	średnia	mediana	odchylenie standardowe	wariancja
Prs	4.275382	4.154472	1.420923	2.019022
Ga	0.5836287	0.2127156	0.7853937	0.6168433

## Podsumowanie

Obliczamy przedziały ufności:

```
PRS
conf_interval <- CONF_INTRV(a_prs, alfa)
conf_interval
```

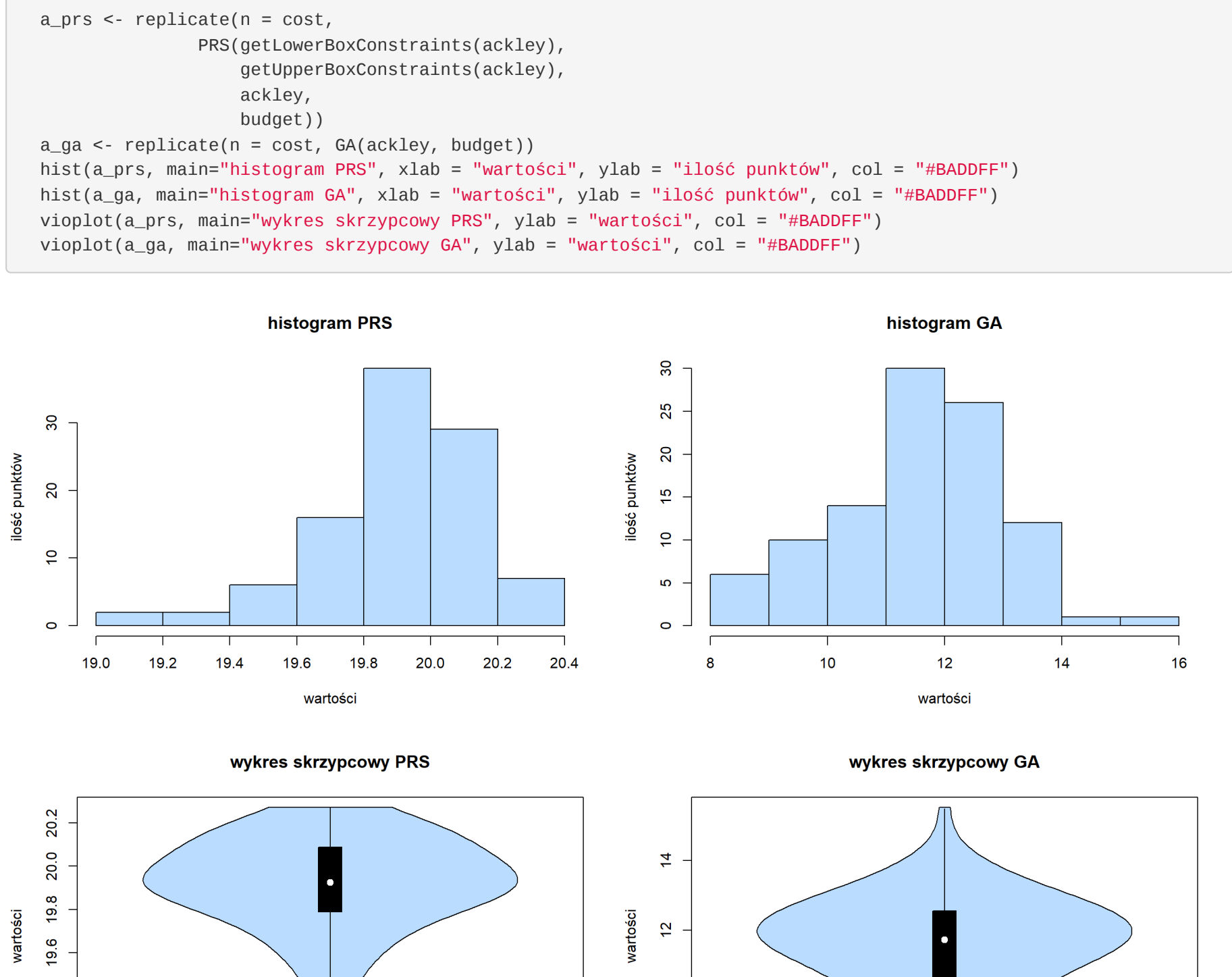
```
GA
conf_interval <- CONF_INTRV(a_ga, alfa)
conf_interval
```

```
## [1] 3.274548 5.276216
```

```
## [1] 0.3756512 0.7915962
```

W przypadku porównywania 2 wymiarów dla funkcji Ackley'a sprawdzą się hipotezy alternatywne 2. Dla algorytmu PRS średnia jest bardzo zbliżona do mediany, a odchylenie standardowe jest niewielkie. Natomiast dla algorytmu genetycznego średnia jest niemal trzykrotnie większa od mediany, a odchylenie standardowe jest znaczące.

## 10 wymiarów



Algorytm	średnia	mediana	odchylenie standardowe	wariancja
Prs	18.0935143	18.2096402	0.7707545	0.5940626
Ga	7.991149	7.778349	1.900478	3.611818

## Podsumowanie

Obliczamy przedziały ufności:

```
PRS
conf_interval <- CONF_INTRV(a_prs, alfa)
conf_interval
```

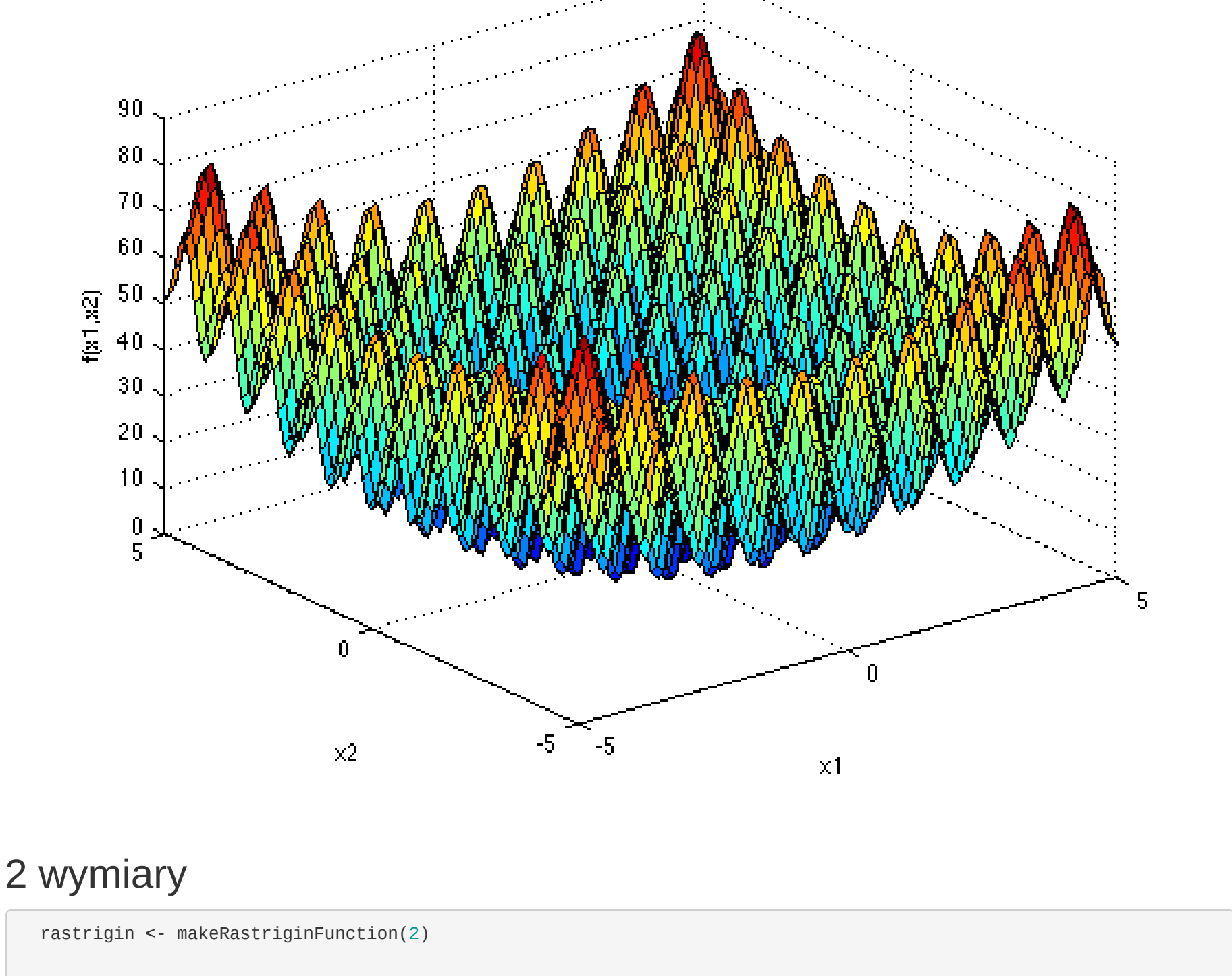
```
## [1] 16.62239 19.56473
```

```
GA
conf_interval <- CONF_INTRV(a_ga, alfa)
conf_interval
```

```
## [1] 5.796658 10.185640
```

W przypadku porównywania 10 wymiarów dla funkcji Ackley'a sprawdzą się hipotezy alternatywne 2. Dla algorytmu PRS średnia jest prawie identyczna jak mediana, a odchylenie standardowe jest niewielkie. Dla algorytmu genetycznego średnia jest bardzo zbliżona do mediany, a odchylenie standardowe jest stosunkowo niewielkie.

## 20 wymiarów



Algorytm	średnia	mediana	odchylenie standardowe	wariancja
Prs	19.90033005	19.92528026	0.24156182	0.05835211
Ga	11.549499	11.712324	1.454063	2.114299

## Podsumowanie

Obliczamy przedziały ufności:

```
PRS
conf_interval <- CONF_INTRV(a_prs, alfa)
conf_interval
```

```
## [1] 19.41127 20.38940
```

```
GA
conf_interval <- CONF_INTRV(a_ga, alfa)
conf_interval
```

```
## [1] 9.519652 13.579346
```

W przypadku porównywania 20 wymiarów dla funkcji Ackley'a sprawdzą się hipotezy alternatywne 2. Dla algorytmu PRS średnia i mediana są porównywalne, a odchylenie standardowe jest niewielkie. Natomiast dla algorytmu genetycznego średnia jest bardzo zbliżona do mediany, a odchylenie standardowe jest znaczące.

## Funkcja Rastrigin'a

Wzór funkcji:

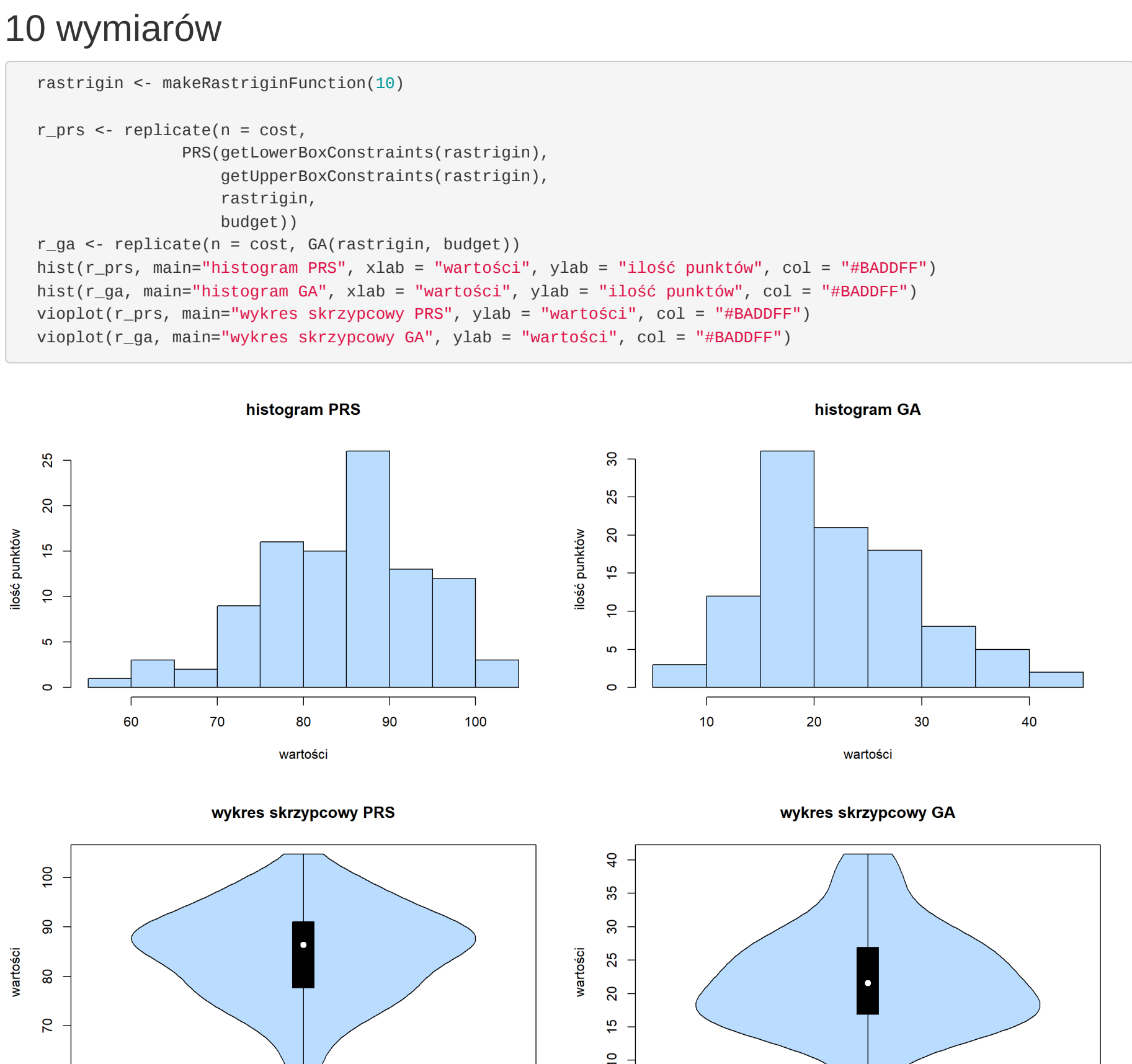
$$f(\mathbf{x}) = 10d + \sum_{i=1}^d (x_i^2 - 10 \cos(2\pi x_i))$$

Minimum globalne:

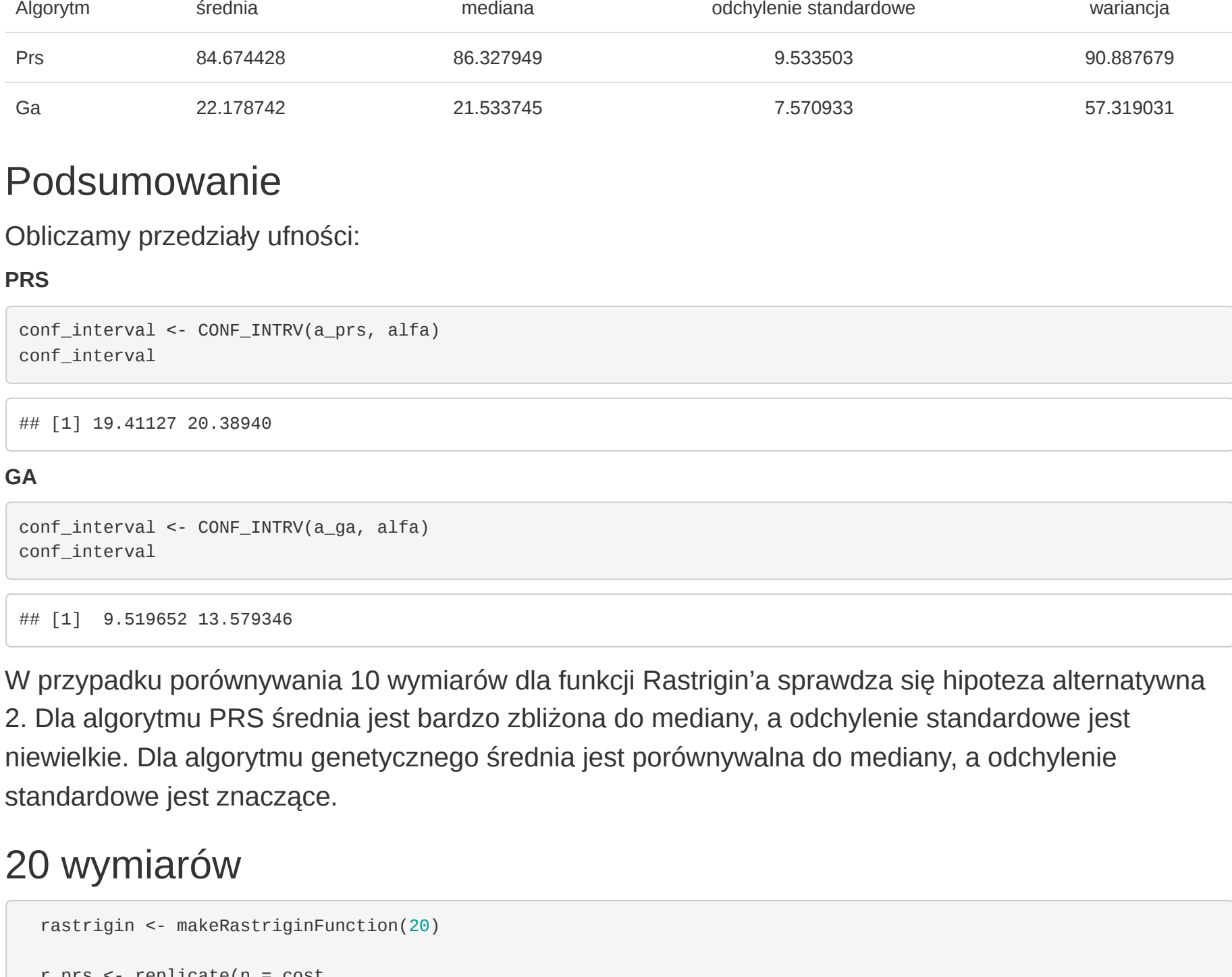
$$f(\mathbf{x}^*) = 0, \text{ at } \mathbf{x}^* = (0, \dots, 0)$$

Zalecana dziedziła poszukiwań:

$$x_i \in [-5.12, 5.12], \forall i \leq d, \text{ gdzie } d \text{ to wymiar przestrzeni}$$



## 2 wymiary



Algorytm	średnia	mediana	odchylenie standardowe	wariancja
Prs	1.6849430	1.5276102	0.8991705	0.8085075
Ga	0.41478896	0.08661706	0.53380372	0.28494641

## Podsumowanie

Obliczamy przedziały ufności:

```
PRS
conf_interval <- CONF_INTRV(a_prs, alfa)
conf_interval
```

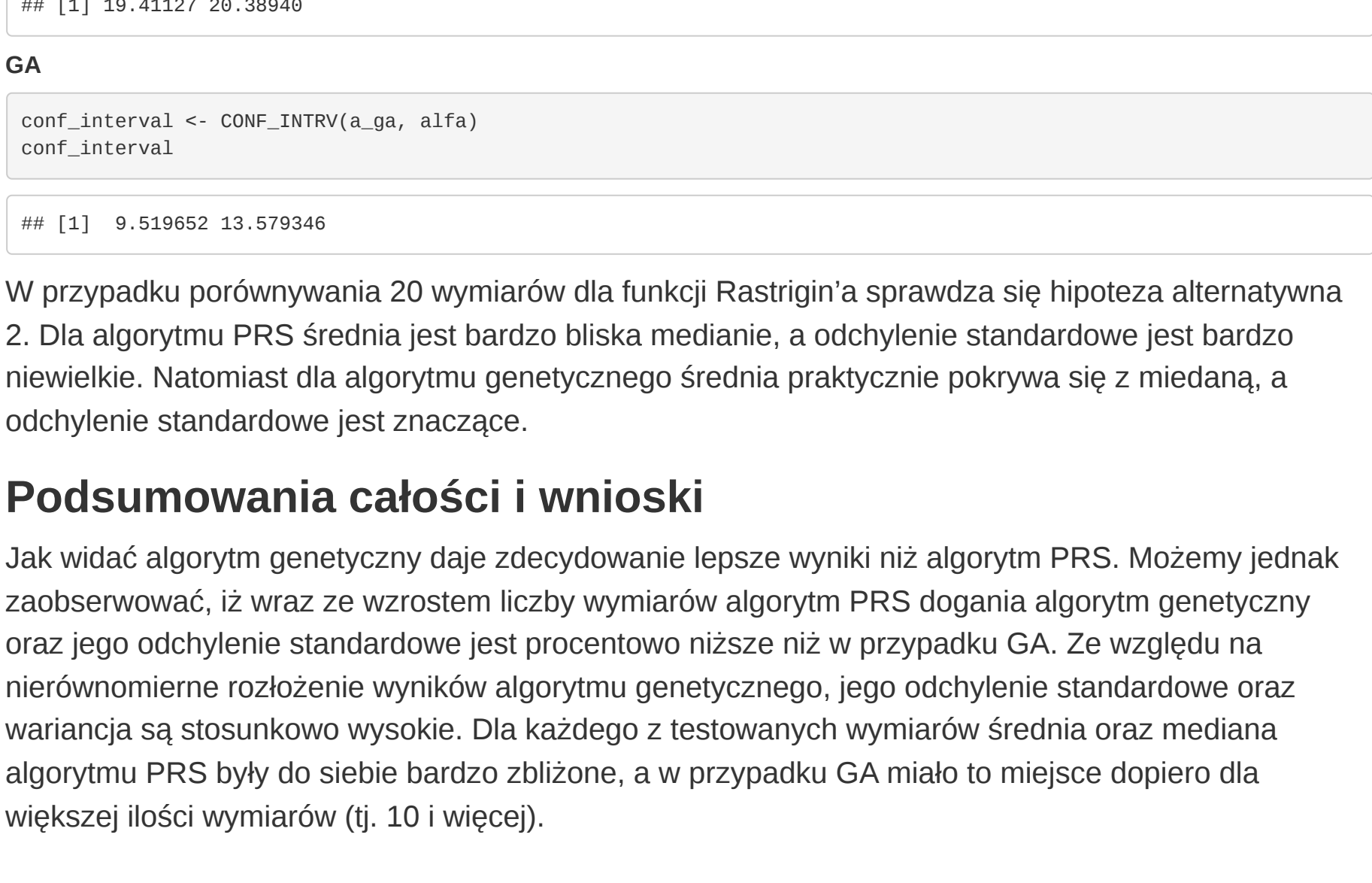
```
## [1] 19.41127 20.38940
```

```
GA
conf_interval <- CONF_INTRV(a_ga, alfa)
conf_interval
```

```
## [1] 9.519652 13.579346
```

W przypadku porównywania 10 wymiarów dla funkcji Rastrigin'a sprawdzą się hipotezy alternatywne 2. Dla algorytmu PRS średnia i mediana są porównywalne, a odchylenie standardowe jest niewielkie. Natomiast dla algorytmu genetycznego średnia jest bardzo zbliżona do mediany, a odchylenie standardowe jest znaczące.

## 20 wymiarów



Algorytm	średnia	mediana	odchylenie standardowe	wariancja
Prs	228.4766	230.6898	13.2615	175.8673
Ga	74.55114	74.01480	18.86128	355.74783

## Podsumowanie

Obliczamy przedziały ufności:

```
PRS
conf_interval <- CONF_INTRV(a_prs, alfa)
conf_interval
```

```
## [1] 19.41127 20.38940
```

```
GA
conf_interval <- CONF_INTRV(a_ga, alfa)
conf_interval
```

```
## [1] 9.519652 13.579346
```

W przypadku porównywania 20 wymiarów dla funkcji Rastrigin'a sprawdzą się hipotezy alternatywne 2. Dla algorytmu PRS średnia jest bardzo bliska medianie, a odchylenie standardowe jest bardzo niewielkie. Natomiast dla algorytmu genetycznego średnia jest praktycznie pokrywa się z medianą, a odchylenie standardowe jest znaczące.

## Podsumowania całości i wnioski

Jak widać algorytm genetyczny daje zdecydowanie lepsze wyniki niż algorytm PRS. Możemy jednak zaobserwować, iż wraz ze wzrostem liczby wymiarów algorytm PRS dogania algorytm genetyczny oraz jego odchylenie standardowe jest procentowo niższe niż w przypadku GA. Ze względu na nierównomierne rozłożenie wyników algorytmu genetycznego, jego odchylenie standardowe oraz wariancja są stosunkowo wysokie. Dla każdego z testowanych wymiarów średnia oraz mediana algorytmu PRS były do siebie bardzo zbliżone, a w przypadku GA miało to miejsce dopiero dla większej ilości wymiarów (tj. 10 i więcej).