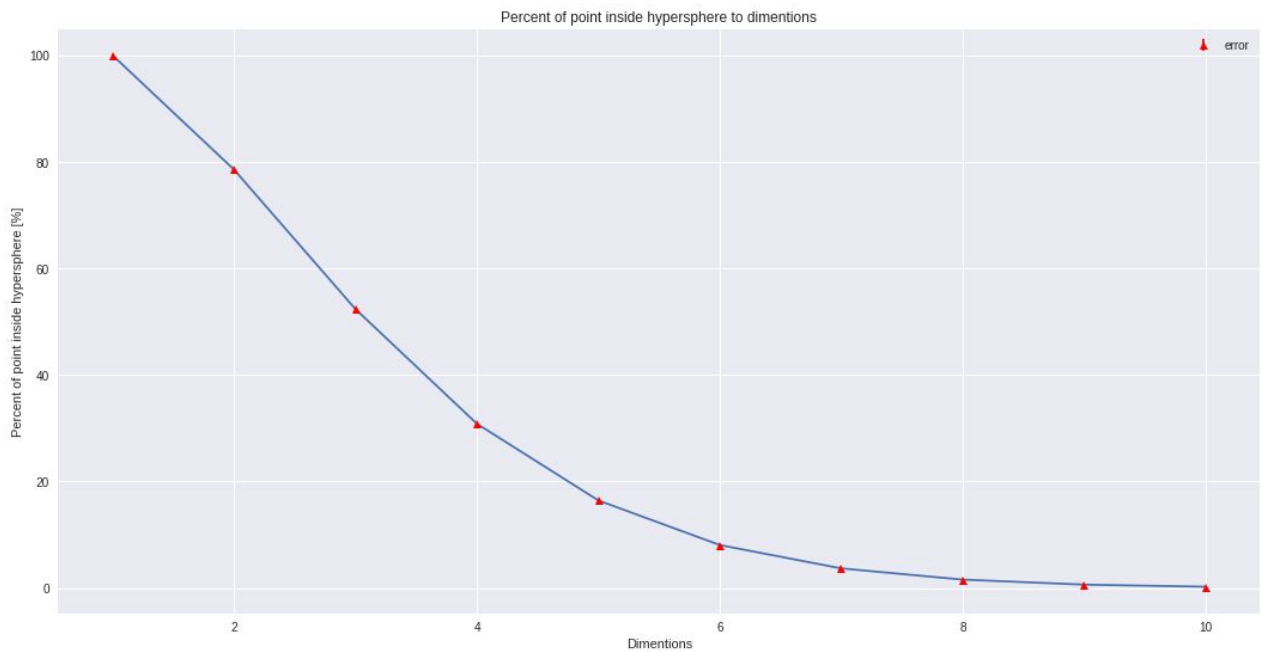


### Raport

#### Zad a.

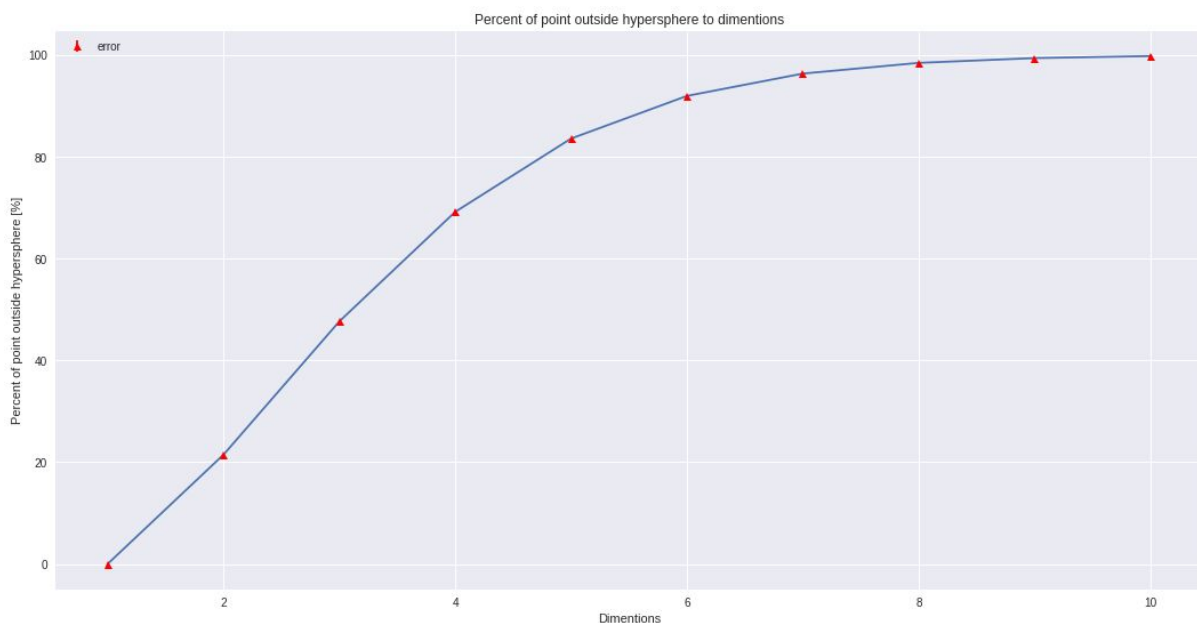
Wygenerowałam 200 000 punktów w hipersześcianie o krawędzi  $2X$ , którego środkiem jest środek układu współrzędnych. Następnie za pomocą metody monte carlo wyznaczyłam stosunek punktów znajdujących się w hiperkuli o promieniu  $X$ , i środku w środku układu współrzędnych.

Poniższy wykres przedstawia procent punktów zawierających się w hiperkuli do punktów wszystkich w hipersześcianie.



Czerwone strzałeczki oznaczają błąd, czyli odchylenie standardowe, które jest bardzo małe.

Poniższy wykres przedstawia procent punktów zawierających się poza hiperkulą do punktów wszystkich w hipersześcianie.

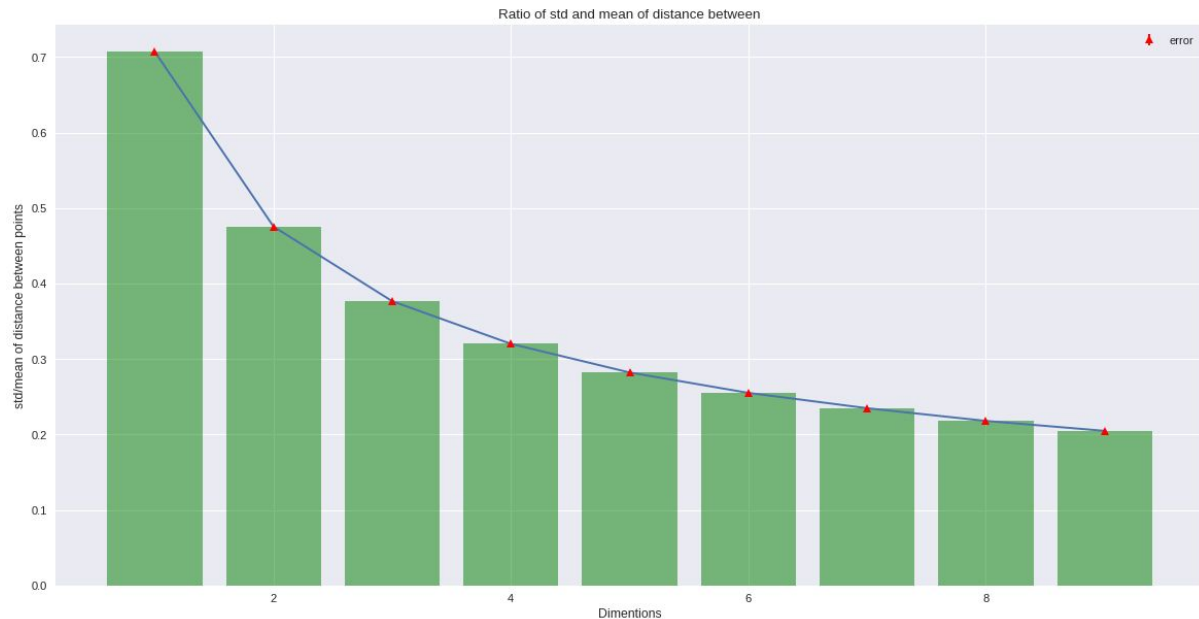


Jak widać na wykresach wraz ze wzrostem wymiaru liczba punktów w hiperkuli maleje, czyli można zauważyć, że są coraz dalej od środka kuli.

## Zad b.

Wygenerowałam 10 000 punktów w hipersześcianie o krawędzi 1. Policzyłam średnią i standardowe odchylenie odległości punktów w tym hipersześcianie i policzyłam ich stosunek.

Poniższy wykres przedstawia zależność stosunku odchylenia standardowego do średniej odległości punktów od wymiarowości hipersześcianu



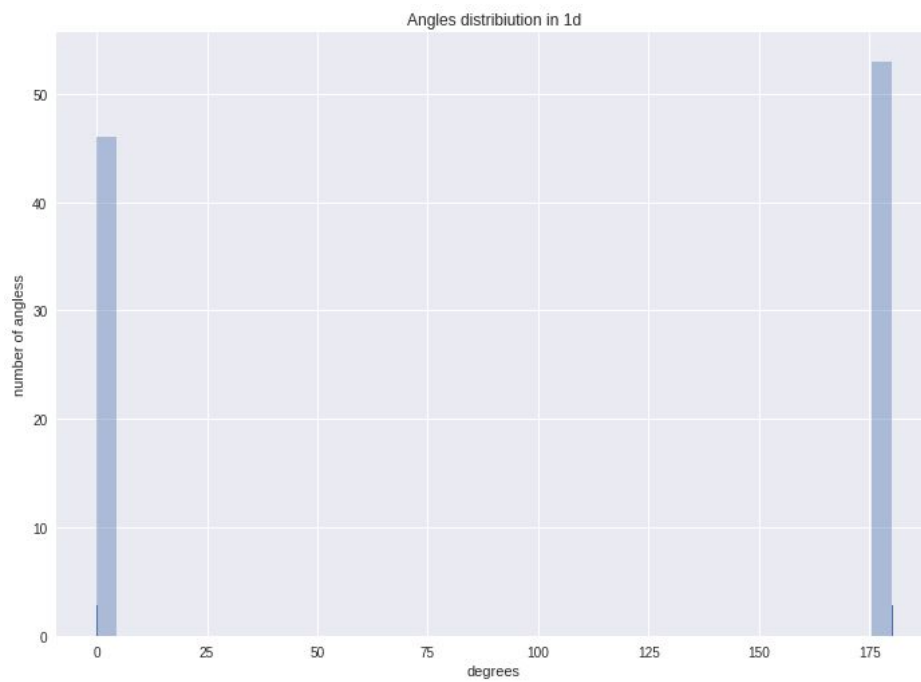
Stosunek odchylenia standardowego do średniej odległości między punktami w hipersześcianie maleje wraz ze wzrostem wymiaru.

## Zad c.

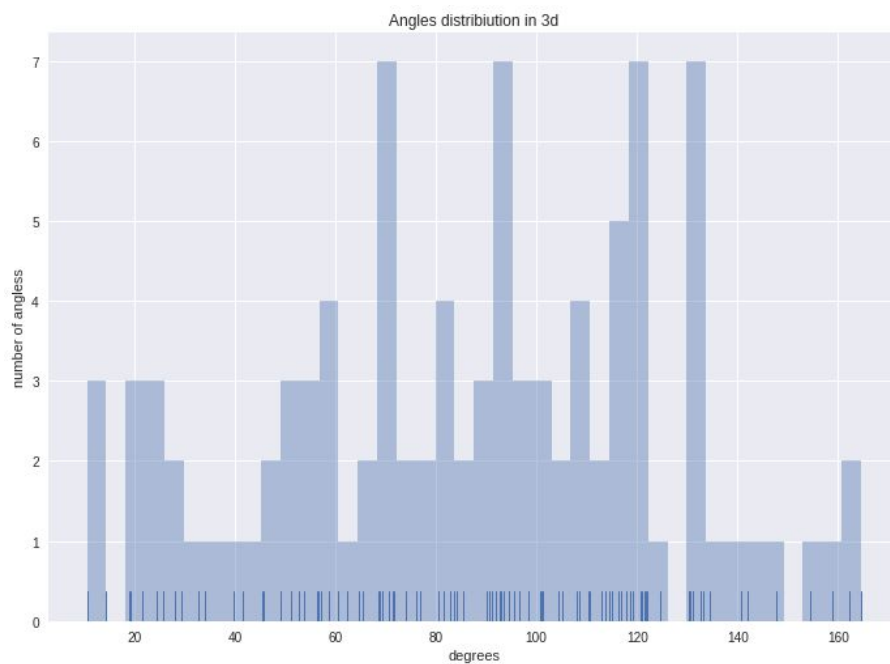
Wygenerowałam 200000 punktów w hipersześcianie o boku 1. Zmierzyłam kąty między 100 losowo wybranymi parami wektorów.

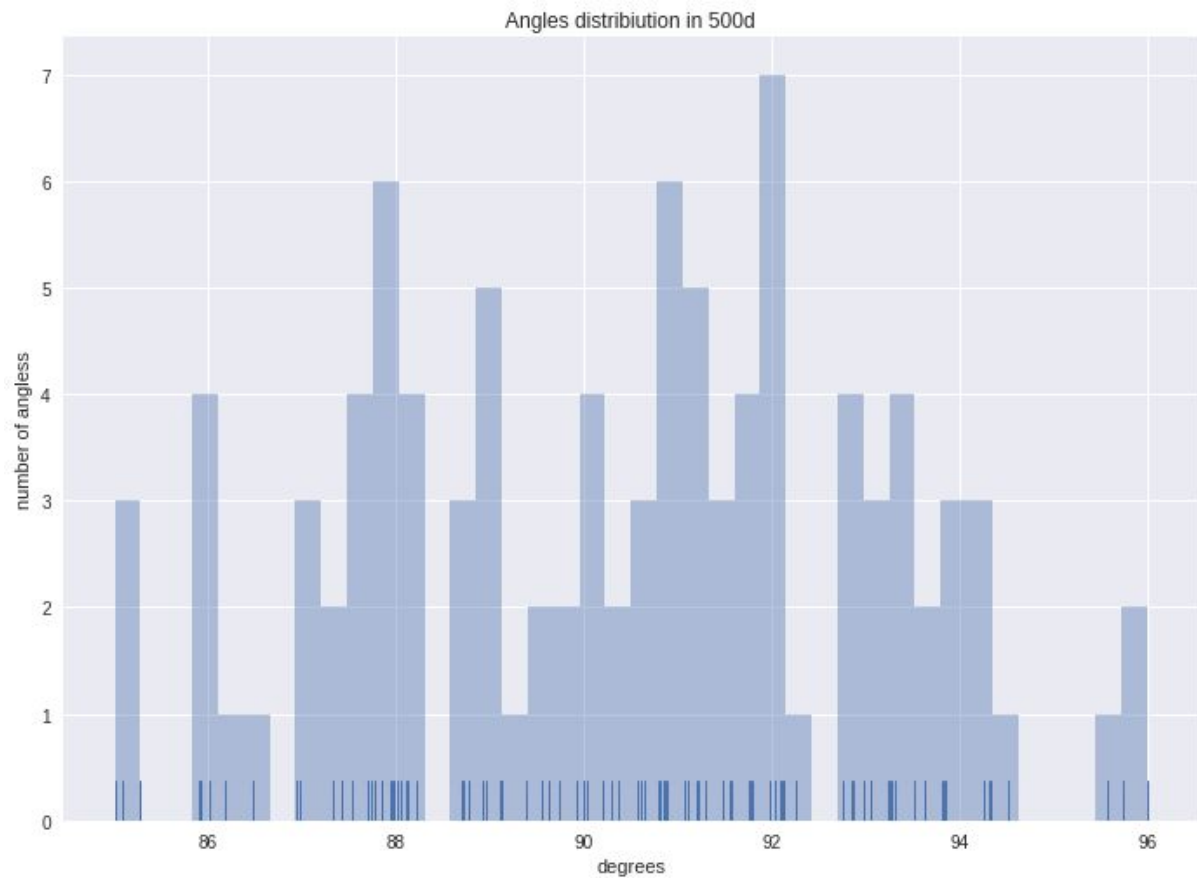
Wykresy przedstawiające dystrybucję kątów między wektorami w hipersześcianie:

Dla 1d:



Dla 3d





Widać, że w dużych wymiarach kąty dążą do  $90^\circ$ .

#### Wnioski:

- Wraz ze wzrostem wymiarowości punkty są dalej od siebie, przestrzeń bardziej pusta.
- Na podstawie zad a. w większych wymiarach widać, że dane są skoncentrowane w rogach i ciężiej je przez to sklasyfikować, niż gdyby były sobie wokół jakiegoś scentralizowanego punktu.
- Odległości od różnych punktów danych są praktycznie nierozróżnialne. Miara niepodobieństwa w większych wymiarach traci sens. Jeśli klasyfikator zależy od tej odległości to lepiej będzie działał w mniejszych wymiarach.
- Jeśli weźmiemy dużo feature'ów to będziemy musieli mieć bardzo dużą próbkę.