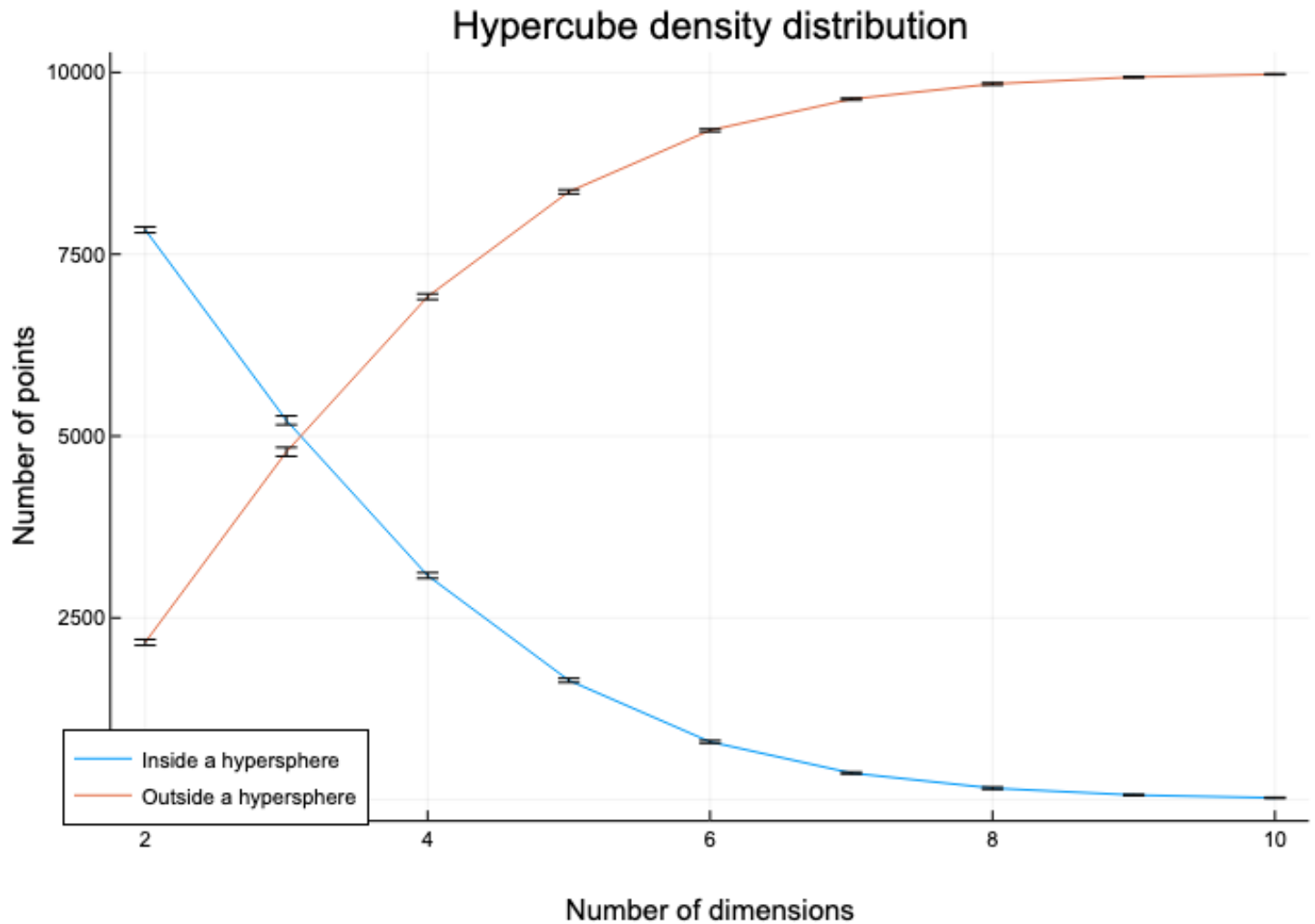


Kłątwa wymiaru

Mamy hiperkulę o promieniu równym 1 wpisaną w hipersześcian o krawędziach długości 2. Zapełniamy hipersześcian losowymi punktami o równomiernym rozkładzie.

Punkt 1

Mierzmy rozkład punktów wewnątrz wielowymiarowego sześcianu.

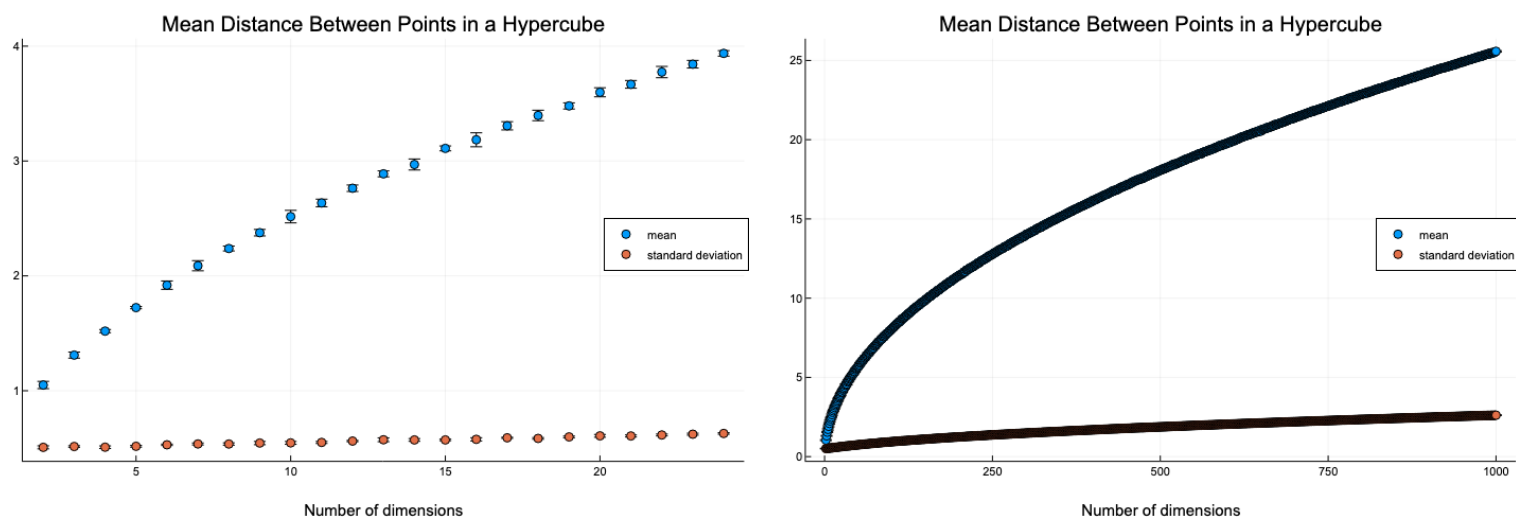


Ustaliłem liczbę punktów na 10^4 , dla każdego wymiaru doświadczenie powtarzałem 10 razy aby zwiększyć wiarygodność wyników. Wykres zawiera uśrednione wyniki wraz z odchyleniem standardowym.

Okazuje się, że wraz ze wzrostem ilości wymiarów drastycznie maleje liczba punktów wewnątrz sfery - co oznacza, że narożniki są znacznie bardziej pojemne niż środek sześcianu (hiperkula wpisana w ten sześcian).

Punkt 2

Tutaj będziemy badać odległości między punktami w hipersześcianie.

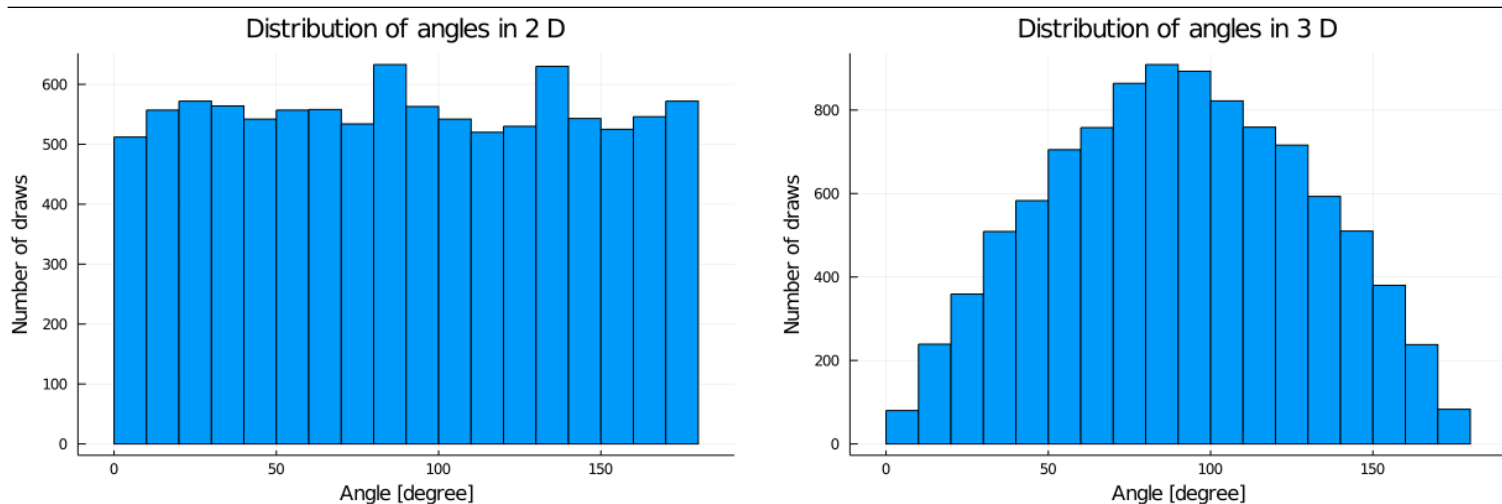


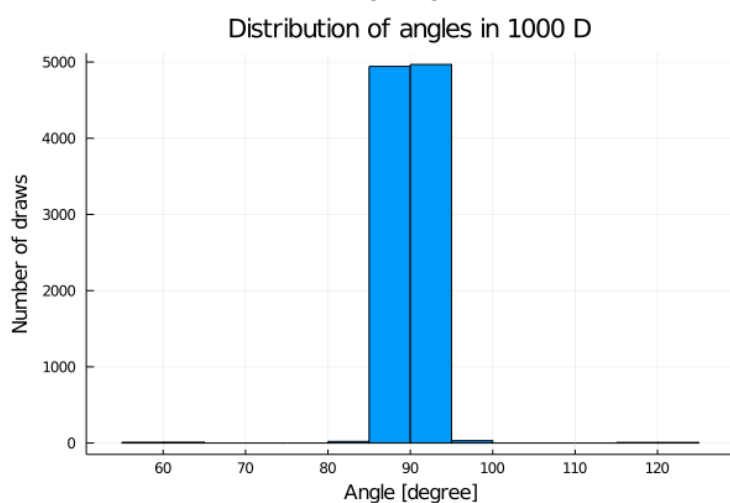
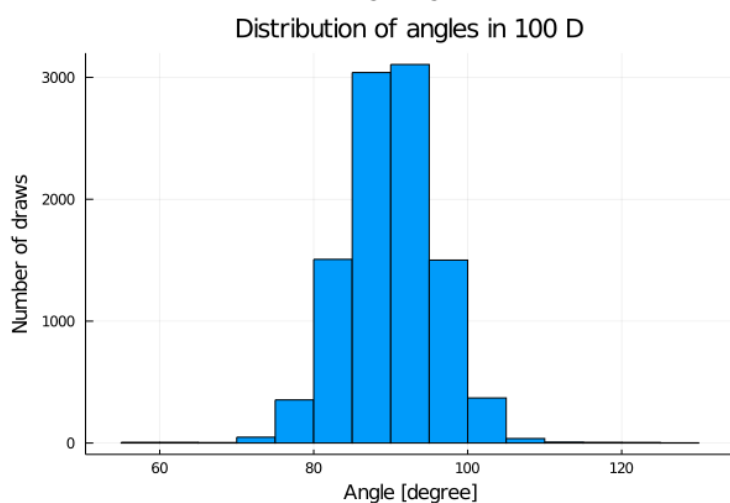
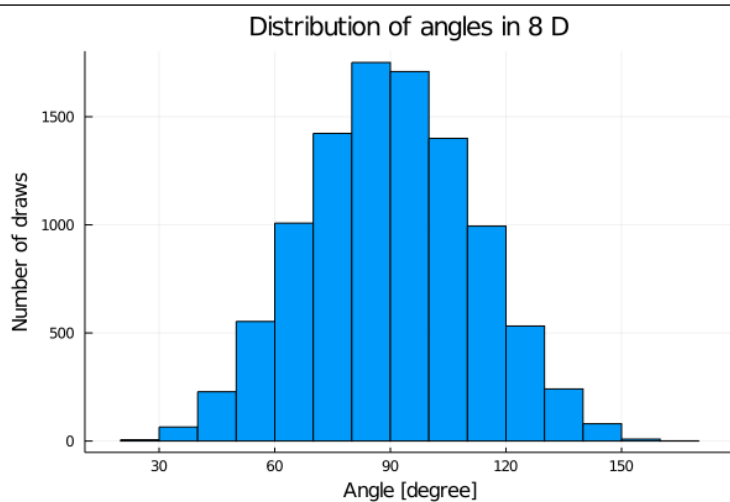
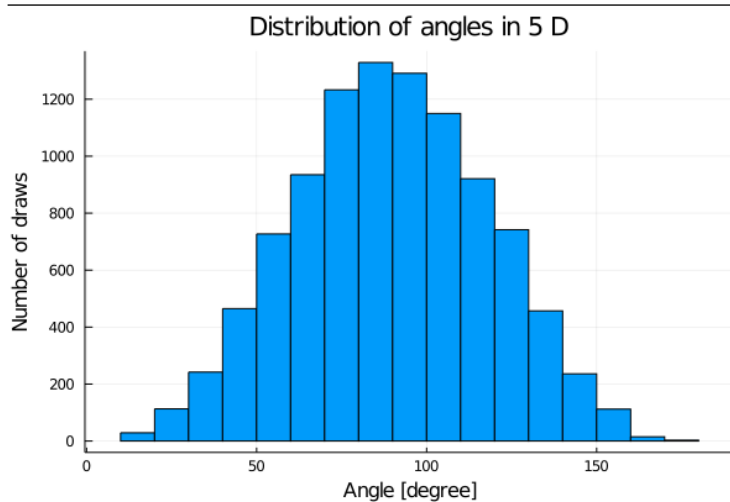
Każde doświadczenie powtórzyłem 5 razy, losowałem po 100 punktów i liczyłem odległości dla każdego z każdym - na osi Y mamy serię uśrednionych wartości odchylenia standardowego i uśrednioną serię średniej odległości między punktami dla konkretnego wymiaru.

Widać, że wraz ze wzrostem liczby wymiarów średnia odległość rośnie w skali logarytmicznej (lub $\sqrt{(x)}$). Myślę, że można o tym wyniku myśleć, że wraz ze wzrostem liczby wymiarów przestrzeni robi się coraz rzadsza - tzn. mamy stałą liczbę punktów, które mają coraz więcej miejsc do wyboru (bo każdy wymiar to dodatkowa swoboda) i jeśli są one losowo rozmieszczone na ograniczonej przestrzeni to ta średnia odległość będzie rosła.

Punkt 3

Tutaj badamy rozkład kątów między wektorami utworzonymi z wylosowanych punktów wewnątrz sześcianu.





Losowałem 1000 punktów i do uzyskania poszczególnego histogramu robiłem 10 tysięcy losowań (2 wektory, po 2 punkty na wektor) z których obliczałem kąt z iloczynu wektorowego.

$$\phi(v_1, v_2) = \arccos\left(\frac{v_1 \cdot v_2}{\|v_1\| \cdot \|v_2\|}\right)$$

Obserwujemy, że wraz ze wzrostem wymiarów kąty między wektorami są coraz ciaśniej rołożone między kątem 90°.

Mozemy więc wywnioskować, że mierzenie podobieństwa między wektorami za pomocą wzoru wyżej (kąta bądź wartości cosinusa) nie będzie się dobrze sprawdzać.