##projekt PYTHON

-- UWAGA: przeczytać pliki ze slacka nt. generatorów

Sprawdza: wiedza statystyczna + wiedza algorytmiczna

Timeline: 3-4 zajęć projektowych

\*Bez przerostu formy nad treścią. Prezentacja to pokazanie wyników działania kodu. Nie trzeba

\* robić ogromnej otoczki typu kolor wykresu itd.

Zadanie w projekcie:

1. implementacja metody: generatora ciagu liczb

2. ciag spelnia testy statystyczne

3. zilustrowanie tych dane - wykresy

Komputer działa deterministycznie, nie potrafi wygenerować losowych danych.

Generatory ustalają zmienne na podstawie wyznaczników, które wydają nam się losowe.

Tematy:

1. LCG (liniowy generator kongruentny) - RUNDOOM - metoda deterministyczna

2. LFG (generator Fibonacciego) - VENUS - metoda deterministyczna

3. generator liczb z rozkładu liniowego/normalnego (2 metody) - APACZE

4. generator metodą eliminacji - trzeba zdefiniować metodę za pomocą funkcji - kilka rozkładów (3-4) do uzgodnienia z grupą - ESPERANTO

Krzysiek:

są dwie metody ale tylko metoda Boxa-Muellera jest do zaimplementowania oraz metoda z centralnego twierdzenia granicznego

w drugim dokumencie jest dokladnie opisane jak je zrobic.

FUnkcje sinus cosinus sa wbudowane w scipy albo w numpy.

Najpierw wybieramy liczbe losowa (w numpy). Rozklad jednostajny jest warunkiem poczatkowym.Str. 18 drugiego pliku

Krzysiek przesle strone gdzie jest to opisane algorytmicznie.

Wyjsciowa funkcja (standardowo jest 0,1) - zadana średnia i zadane odhylenie standardowe - trzeba napisac funkcje.

ALGORYTM - plik w ktorym to jest opisane - str. 16.

Nastepnie analiza tych n + statystyka czy spelnia test czy nie

Losujemy liczby, warto zobaczyc to ile razy + robimy test statystyczny

potem drugi raz i znowu test (oczywiscie w petli tzw. metoda monte carlo - dla n=7 losowalismy to 1000 razy i 50% nie przeszlo testu statystycznego; losujemy 30 liczb losowych z kazdego algorytmu i sprawdzamy czy przechodza test statystyczny)

1. Napisac 2 funkcje do losowania z rozkladow + porownanie testow statystycznych

2. wykresy.

3. Do funkcji warto zrobic scenariusze np. ze wprowadzamy parametr, ale gdy wprowadzimy zly to pojawi sie jakis warning.

4. Porownanie tego z normalna funkcja numpy.

Dla rozkladu normalnego sa dostepne w numpy

test kołmogorowa smirnoffa???

test Shafiro (Scipy) jest info w necie o tym jak interpretowac wyniki tego testu

TEST METODA MONTECARLO (19.01.2019)

Losujemy 100 liczb i dla nich robimy te testy.

jesli powtorzymy to 10000 razy to wiemy ze dla ciagu tych liczb mamy w 80% pozytywny wynik testu, a dla 1000 tylko w 90%.

To sie robi dla metody o centralnym twierdzeniu granicznym

dla metody boxa muellera

najlepiej funkcja (argument, liczba iteracji oraz funkcja losujaca tzn. albo numpy, albo boxmueller, albo ctg)

Rezultat - ladna tabelka ktora porownuje rezultalty tych tabel

CENTRALNE TWIERDZENIE GRANICZNE

Central limit theorem – 20.01.2019

Statystyka prawdopodobieństwa/ probability statistics

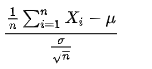
*Pisalismy kod do CTG.*

*Ostatecznie Krzysiek go napisał.*

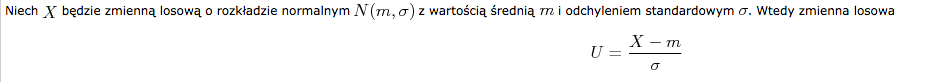
*Zajawka teoretyczna:* Rozkład normalny, inaczej zwany rozkładem Gaussa, krzywą Gaussa jest najważniejszym rozkładem teoretycznym prawdopodobieństwa w statystyce. Rozkład normalny jest też najbardziej intuicyjnym rozkładem statystycznym. W wielkim skrócie opisuje on sytuacje w świecie, gdzie większość przypadków jest bliska średniemu wynikowi, a im dany wynik bardziej odchyla się od średniej tym jest mniej reprezentowany. Najwięcej jest przypadków blisko przeciętnej. Im dalej oddalamy się od średniego wyniku, tym przypadków jest mniej. Można to z łatwością odnieść do rzeczywistych sytuacji [więcej tutaj](http://www.naukowiec.org/wiedza/statystyka/rozklad-normalny-rozklad-gaussa_710.html) po polsku.

1. Zmienna losowa X ma rozklad normalny N(mi, sigma) [mi = srednia w populacji generalnej; sigma = odchylenie standardowe] N oznacza r.normalny
2. Należy obliczyć prawdopodobieństwo, ze zmienna x przyjmie wartości w jakims przedziale (miedzy np. -1 do 1)
3. Aby to obliczyć należy przeprowadzić standardyzacje zmiennej N(mi,sigma) na taki inny układ Z (0,1) <czyli taki gdzie srednia wynosi 0, a odchylenie standardowe 1). Ten rozklad zawsze będzie normalny

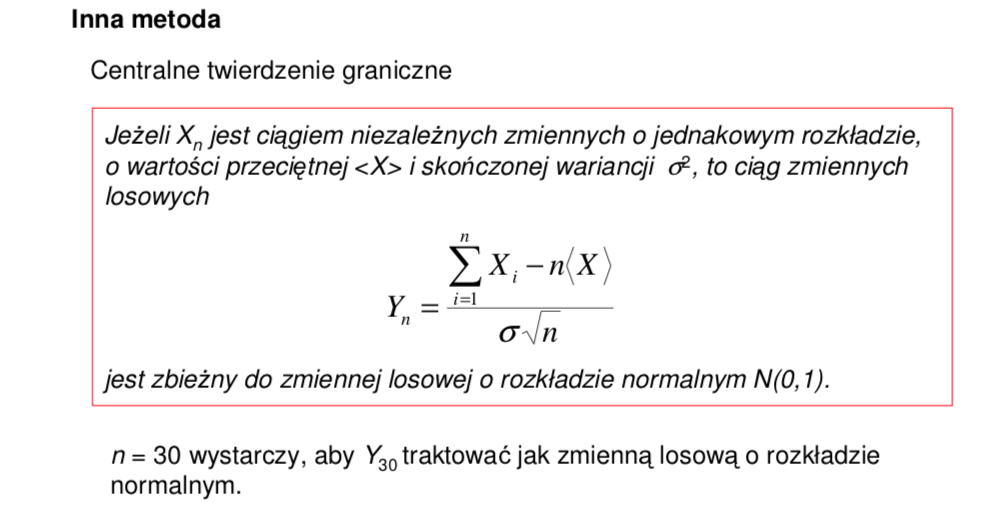
WZOR 1 :

(source: [Wikipedia.pl)](https://pl.wikipedia.org/wiki/Centralne_twierdzenie_graniczne)

Wzor 2 ([matematyka.pl](https://www.matematyka.pl/291136.htm?mobile=on))



Wzor 3: [(skrypt link)](http://www.if.pw.edu.pl/~agatka/numeryczne/wyklad_12.pdf)



**Wyjasnienie zmiennych:**

**N =**

**n= ilość licz liczb wygenerowanych**

**x = niezalezna zmienna losowa o jednakowym rozkładzie**

**xi (1,2,3,4,) = kolejne niezależne zmienne losowe**

**y = rezultat funkcji f(x)**

**mi = srednia**

**sigma = odchylenie standardowe (sigma ^2 to wariancja)**

**i = liczba iteracji (obrotow w pętli)**

import numpy as np

import matplotlib.pyplot as plt

x\_n = 30 ## N (duże N) ze slajdu

n = 10000 ## n - liczba liczb do wygenerowania

Y = []

for k in range(n):

X = np.random.uniform(-1,1 , size=x\_n)

y = np.mean(X-np.mean(X) )/(np.std(X)/ np.sqrt(x\_n))

Y.append(y)

plt.hist(Y)

plt.show()

Linki:

1. Wyklad: Tensorflow playground (niezwiazany z tematem) <https://playground.tensorflow.org>
2. Wyklad: Precision and recall (dobry przykład z kotem i psem, tez wykład) <https://en.wikipedia.org/wiki/Precision_and_recall>
3. Mnie ten filmik pomogl zrozumieć ten problem: <https://www.youtube.com/watch?v=Oh9iLuiJuJI>

TEST Shapiro Wilka

1. [Link po polsku](http://www.naukowiec.org/wiedza/statystyka/test-shapiro-wilka_758.html)
2. [Wzor](http://www.naukowiec.org/wzory/statystyka/test-shapiro-wilka_333.html)
3. Tablica rozkładu W dla testu [link](http://www.naukowiec.org/tablice/statystyka/rozklad-w-test-shapiro-wilka-_335.html)

TEST KOŁMOGOROWA – SMIRNOFA

1. [Link po polsku](http://www.naukowiec.org/wiedza/statystyka/test-kolmogorowa-smirnowa_752.html)