Proiect la materia „Probabilități & Statistică”

Jocul de cărți „Război”

Estimarea duratei unei partide folosind metoda simulării Monte Carlo

Militaru Ștefan-Octavian

Cucu Mihnea

Seria 25 – grupa 252

# Introducere

Pentru realizarea acestui proiect am ales ca temă jocul de cărți „Război”, un joc popular, mai ales în rândul copiilor, datorită regulamentului simplu și opțiunii de a juca cu un număr variabil de participanți. Alegerea acestui joc a fost datorată, în mare parte, faptului că acesta este un joc de noroc pur, ceea ce înseamnă că participanții nu pot influența cursul acestuia în niciun mod, fiind în totalitate constrânși de amestecarea aleatoare a pachetului.

Acest lucru este ideal pentru o simulare Monte-Carlo, scutindu-ne de implementarea unor algoritmi de decizie. Fluxul jocului este însă îndeajuns de complex încât calcularea unui timp de finalizare a unei partide este un exercițiu suficient de complex încât să justifice utilizarea unei simulări Monte-Carlo.

# Regulile Jocului

„Război” are, ca orice joc vechi tradițional, multe variante diferite de regulament, însă diferențele între acestea nu sunt așa semnificative ca la alte jocuri de cărți. Pentru implementarea proiectului noi am decis să alegem varianta în doi jucători fără utilizarea cărților „joker” (acestea fiind introduse pentru a le da participanților opțiunea de a le folosi strategic, eliminând componenta de noroc pur pe care dorim să o utilizăm).

Jocul începe prin amestecarea pachetului de cărți, acest pas fiind de departe cel mai important întrucât configurația de start a acestuia determină în totalitate câștigătorul și durata partidei. Pachetul este mai apoi împărțit in bucăți egale, în cazul nostru câte 26 de cărți pentru fiecare jucător, iar jocul propriu zis poate începe cu aceste pachete așezate cu fața în jos în dreptul participanților.

Structura jocului este simplă, acesta fiind împărțit în runde. La începutul fiecărei runde jucătorii extrag prima carte din pachetul lor și le plasează una lânga alta. Dacă valoarea afișată de una dintre cărți este mai mare, jucătorul care a plasat-o primește ambele cărți aflate în joc și le plasează în ordine la finalul pachetului său. În cazul cărților fără valori numerice precum Valetul (J), Regina (Q), Regele(K) și Asul(A), acestea au valorile 12, 13, 14, respectiv 11 sau 15 (depinzând de varianta aleasă).

În cazul unei egalități se declanșează „războiul”, originea denumirii întregului joc. Războiul pornește de la valoarea afișată pe cărțile aflate la mijloc, jucătorii fiind nevoiți să își pună în joc un număr de cărți egal cu aceasta. Dacă unul sau mai mulți jucători rămân fără cărți în timpul războiului, aceștia se opresc, cei care încă mai au cărți in pachet continuând până ajung la numărul țintă. În momentul în care se opresc toți jucătorii, se compară valorile ultimelor cărți puse în joc, iar cel cu valoarea cea mai mare primește toate cărțile din cadrul războiului. În cazul unei alte egalități se începe alt război cu valoarea afișată de cărțile care au declanșat impasul.

Jocul se termină în momentul în care unul din jucători rămâne fără cărți după o rundă, acesta fiind declarat pierzător. Dacă acesta rămâne fără cărți în timpul unui război pe care mai apoi îl câștigă după ce celălalt jucător termină de plasat toate cărțile, jocul poate continua. Dacă la finalul unui război ambii jucători rămân fără cărți și valoarea ultimelor cărți puse în joc este egală, partida se încheie la egalitate.

# Formularea Problemei

Întrucât prezicerea câștigătorului acestui joc nu este un exercițiu probabilistic foarte complicat, dorim să aflam informații despre durata unei astfel de partide. Pentru a translata problema în termeni ușor de înțeles pentru publicul general, am decis să măsurăm durata unui joc atât în runde cât și într-o aproximare grosieră temporală, obținută din numărul de runde calculat. Din experiența personală durata unei runde (plasarea cărților, analiza valorilor, incorporarea acestora în pachet) este în medie aproximativ 10 secunde, adăugând și timpul dintre runde pentru aranjarea cărților sau alte distrageri. Vom considera ca „rundă” și fiecare plasare de carte din cadrul unui război.

Întrebarea pe care ne-o punem, considerând toate informațiile de mai sus este: „Cât durează, în medie, o partidă de Război?”. Firește, nu ne vom axa în totalitate pe răspunsul la această problemă, dorind să expunem pe lângă acesta și alte informații statistice despre joc. Printre acestea se numără:

* „Care este probabilitatea ca un joc să dureze mai mult de o oră?”
* „La ce distribuție a timpilor de joc ne putem aștepta?”
* „Care este numărul mediu de războaie într-o partidă aleatoare? Dar într-una care durează peste o oră?”
* „Care este corelația între numărul de războaie și durata jocului?”
* „Există un prag minim sau maxim pentru durata jocului?”

# Rezolvarea Problemei

**Pasul 1: Pregătirea pachetului de cărți**

1. pachet = list(range(2, 15)) \* 4

2. np.random.shuffle(pachet)

3.

• Creăm un pachet complet de 52 de cărți (valori între 2 și 14, fiecare de 4 ori).

• Amestecăm pachetul folosind np.random.shuffle.

**Pasul 2: Distribuirea cărților**

1. jucator1 = deque(pachet[:26])

2. jucator2 = deque(pachet[26:])

3.

• Împărțim pachetul în două părți egale: 26 de cărți pentru fiecare jucător.

• Fiecare jucător își stochează cărțile într-un deque (listă dublu legată), care permite adăugarea și eliminarea eficientă a cărților.

**Pasul 3: Simularea rundelor**

1. runde = 0

2. razboaie = 0

3. while jucator1 and jucator2:

4.

• runde numără câte runde au loc.

• razboaie numără câte situații de “război” apar.

• Jocul continuă până când unul dintre jucători rămâne fără cărți.

**Cazurile dintr-o rundă obișnuită**

1. carte1 = jucator1.popleft()

2. carte2 = jucator2.popleft()

3.

4.

• Fiecare jucător scoate prima carte din pachet.

• **Dacă un jucător câștigă runda:**

1. if carte1 > carte2:

2. jucator1.extend([carte1, carte2])

3. elif carte2 > carte1:

4. jucator2.extend([carte2, carte1])

5.

6.

• Jucătorul care are cartea mai mare câștigă ambele cărți și le adaugă la coada sa.

**Războiul (situație de egalitate)**

1. else:

2. carti\_in\_joc = [carte1, carte2]

3. razboaie += 1

4. numar\_carti\_razboi = carte1

5.

6.

• Dacă cele două cărți sunt egale:

• Se inițiază un “război” (situație specială).

• Toate cărțile implicate sunt plasate temporar într-o listă carti\_in\_joc.

**Mecanicile războiului**

1. while True:

2. nr\_carti1 = len(jucator1)

3. nr\_carti2 = len(jucator2)

4. if nr\_carti1 + nr\_carti2 == 0:

5. break

6.

• Jucătorii scot un număr de cărți egal cu valoarea cărții inițiale (numar\_carti\_razboi).

• Dacă ambii jucători nu au suficiente cărți, războiul se încheie imediat.

**Rezultatul războiului**

1. if carte1 > carte2:

2. jucator1.extend(carti\_in\_joc)

3. break

4. elif carte2 > carte1:

5. jucator2.extend(carti\_in\_joc)

6. break

7. else:

8. razboaie += 1

9. numar\_carti\_razboi = carte1

10.

• Dacă o carte finală este mai mare, câștigătorul primește toate cărțile implicate.

• În caz de egalitate, războiul continuă cu mai multe cărți.

**Rezultatul funcției**

1. return (runde, razboaie)

2.

• Returnează numărul total de runde și numărul total de “războaie” pentru o simulare.

# Algoritmul de tip Monte Carlo

Algoritmii bazați pe simulări Monte Carlo sunt o clasă largă de algoritmi care se bazează, în general, pe simularea repetată a unor evenimente cu caracteristici aleatoare pentru a deduce din rezultatele acestoria informații statistice.

În cazul algoritmului nostru vom utiliza simulări de tipul Monte Carlo pentru a rula mii de jocuri de „război”, concluziile trase în urma demarării acestui proces fiind suficiente pentru a ne răspunde la întrebările formulate mai sus.

Pentru a răspunde la întrebarea principală ne vom folosi de vectorul „runde” care stochează numărul de runde din fiecare simulare a jocului. Media acestui vector este chiar cantitatea notată teoretic cu „Sn”.

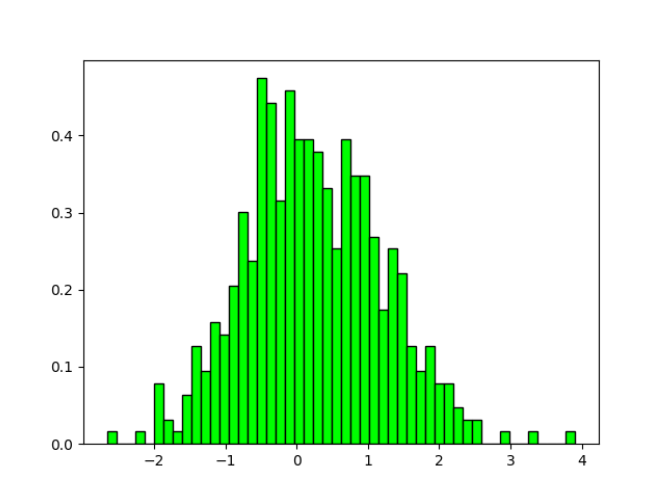
# Estimarea numarului necesar de simulari

Metoda Monte Carlo este bazată pe un raționament teoretic foarte riguros, iar formularea matematică a acesteia ne permite să obținem niște relații care pot estima acuratețea simulărilor generate de aceasta. Pentru a spori eficiența codului și a înțelege mai profund tipul acesta de soluționare a problemelor, vom analiza în acest capitol două teoreme care analizează acuratețea estimărilor.

## Teorema Limită Centrală

Teorema limită centrală pornește de la premisa că avem n variabile aleatoare identic și independent distribuite, cu media µ și varianța σ2 . Atunci putem modela , unde Sn este media obținută din cele n variabile, cu o variabilă aleatoare normală de medie 0 și varianță 1.

Pentru a verifica teorema pe exemplul nostru, este suficient să luăm un interval de valori pentru n, să facem simulările Monte Carlo pentru a obține o valoare Sn și, la final, să realizăm histograma datelor obținute prin formula . Pentru µ și σ vom lua valorile 120 și 84, obțuinute experimental.



Histograma rezultată arată clar distribuția normală – funcție utilizată în program: verificare\_tlc()

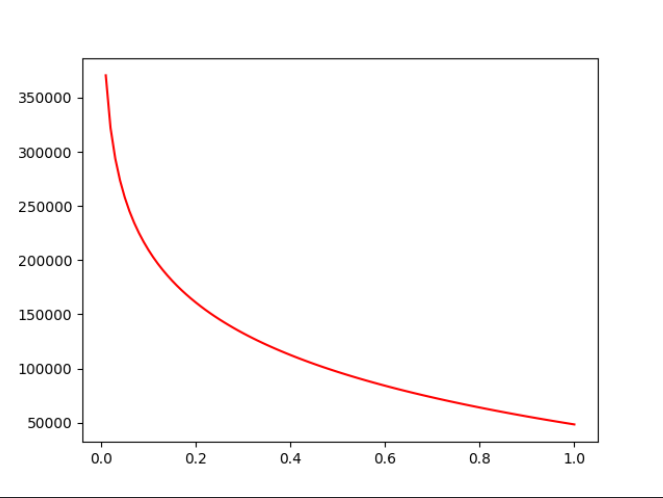
## Inegalitatea Chernoff-Hoeffling

Această formulă se aplică variabilelor aleatoare mărginite și oferă detalii despre discrepanța dintre aproximarea Monte Carlo și valoarea empirică a unei medii. Inegalitatea stipulează că pentru n variabile aleatoare identic și independent distribuite, mărginite inferior de o valoare a și superior de altă valoare b, probabilitatea ca diferența dintre Sn, media calculată a variabilelor, și µ valoarea ei empirică, să fie mai mare ca un număr oarecare ε este mai mică sau egală cu .

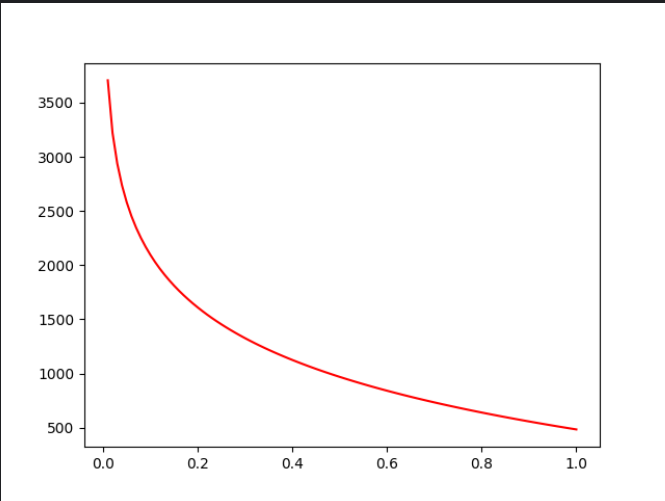
Variabila noastră este mărginită inferior de valoarea 26, însă nu are o limită superioară așa că pentru a putea aplica inegalitatea va trebui să selectăm doar rezultatele sub o anumită valoare aleasă. Vom alege ca prag superior 400 pentru că aceasta este o valoare care este rar depășită în simulările noastre.

Astfel, putem calcula că pentru a avea o probabilitate de sub 1% că diferența dintre media noastră și µ să depășească 1 (deci P = 1/100 și ε = 1) , va trebui să rulăm n = - ½ (400 -26)2 \* ln(1 / 200) simulări, adică 370.553,71. Magnitudinea numărului este datorată cazurilor improbabile de a simula un număr mare de runde, mărginirea acestuia fiind prea largă chiar și cu impunerea pragului superior. Funcția din program care modelează această valoare este chernoff\_hoeffling().

Întrucât valoarea lui n scade proporțional cu pătratul lui ε, dacă dorim o diferență de medii sub 10 trebuie să realizăm doar aproximativ 3500 de simulări.

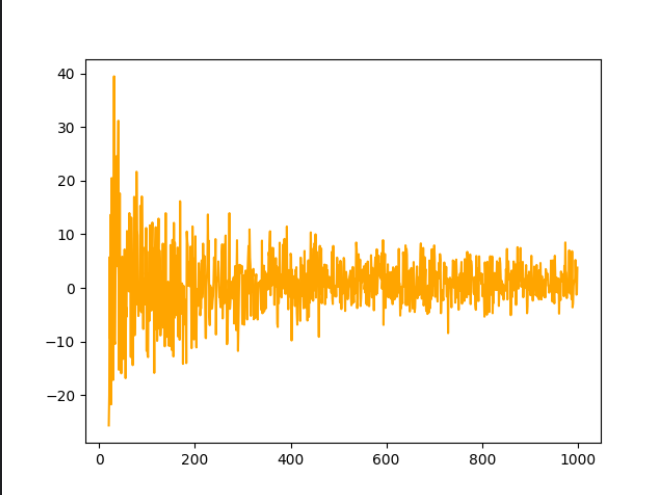


Evoluția numărului de simulări necesare (axa Y) pentru probabilitatea dorită (axa X) ca diferența de medii să fie sub ε = 1. funcție utilizată în program: chernoff\_hoeffling\_grafic()



Același grafic dar pentru ε = 10. Observați diferența dramatică în magnitudinea numărului de simulări necesare (axa Y)

În încheierea capitolului vom analiza evoluția diferenței dintre media calculată cu n variabile aleatoare și cea adevărată. Vom folosi functia simulari(n) și o vom rula pentru n de la 20 la 1000.



Vedem cum valorile converg foarte încet către medie. Functie: evolutie\_aproximari()

# Date obtinute

În acest capitol vom răspunde diverselor întrebări formulate în cadrul proiectului.

## Cât durează, în medie, o partidă de Război?

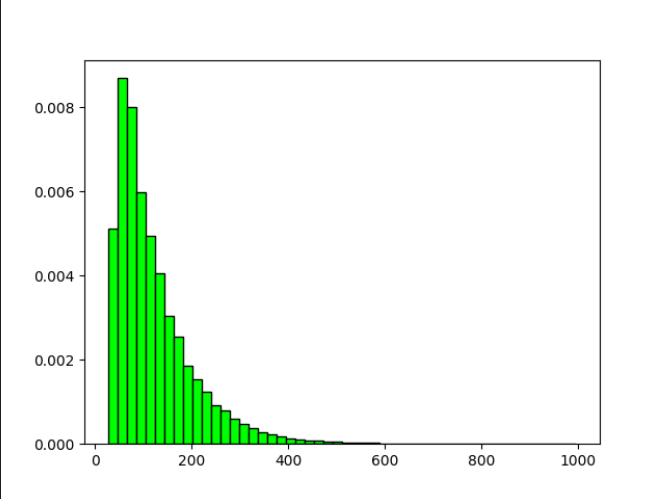
Utilizând media vectorului runde avem ca durată medie 120 de runde, adicâ translatat în valori temporale aproximativ 20 de minute pentru un joc.

## Care este probabilitatea ca un joc să dureze mai mult de o oră?

Analizând numărul de valori peste 360 (1 oră împărțită la 10 secunde) din vectorul de runde putem afla cu ușurință proporția de runde care durează peste o oră. Acestea sunt aproximativ 2% din jocuri.

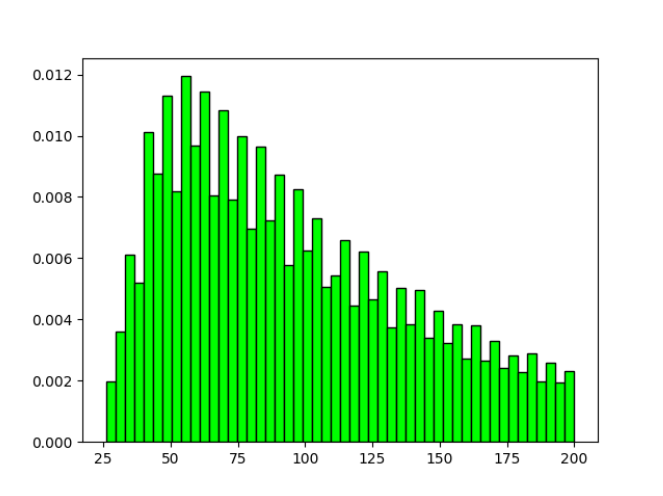
## La ce distribuție a timpilor de joc ne putem aștepta?

Cel mai ușor mod în care putem observa distribuția unui set de date este prin utilizarea unei histograme.



Distribuția totală a duratelor. Funcție utilizată: histo().

Putem extrage mai multe informații din acest grafic. În primul rând, vedem că există un prag minim evident, acela de 26 de runde, atins când unul dintre jucători are norocul de a extrage în fiecare rundă o carte mai valoroasă decât celalalt. Valorile probabilistice urcă apoi spre ~80, iar apoi se înregistrează o descreștere exponențială a probabilităților. Pentru a vedea mai bine zona de interes din jurul mediei putem realiza o nouă histogramă doar cu valorile de la 26 la 200.



Distribuția duratelor în jurul mediei. Funcție utilizată: histo\_redusa().

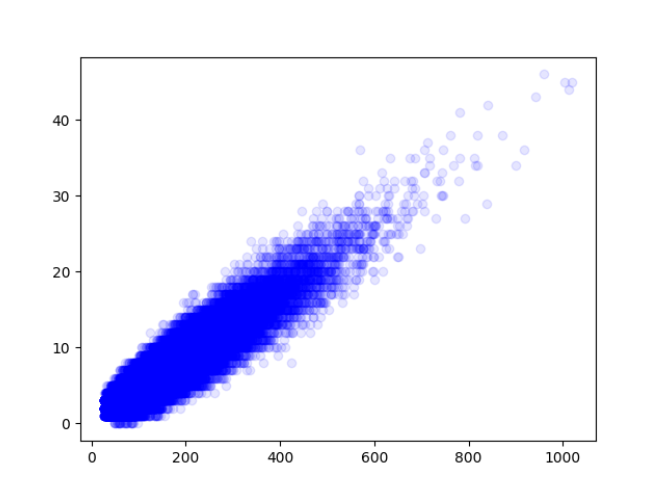
Avem prezentat „vârful” graficului și putem vedea mai clar cum acesta urcă cel mai mult înainte să ajungă la media de 120, valoarea acesteia fiind mărită artificial de rezultatele aleatoriu de mari care pot apărea mai rar. Vedem, deasemenea, o probabilitate puțin mai mare pentru numerele pare, întrucât un meci poate să se încheie într-un număr impar de runde doar dacă unul din jucători rămâne fară cărți în timpul unui „război”.

## Care este numărul mediu de războaie într-o partidă aleatoare? Dar într-una care durează peste o oră?

Rulând simularea Monte Carlo aflăm că o runda obisnuita are, in medie, 5.42 razboaie. Una mai lunga de o ora are 18.76. Creșterea abruptă ne arată că există o corelație semnificativă între durata unei partide și numărul de războaie prezente în cadrul acesteia, idee care va fi dezvoltată mai larg de următorul subpunct.

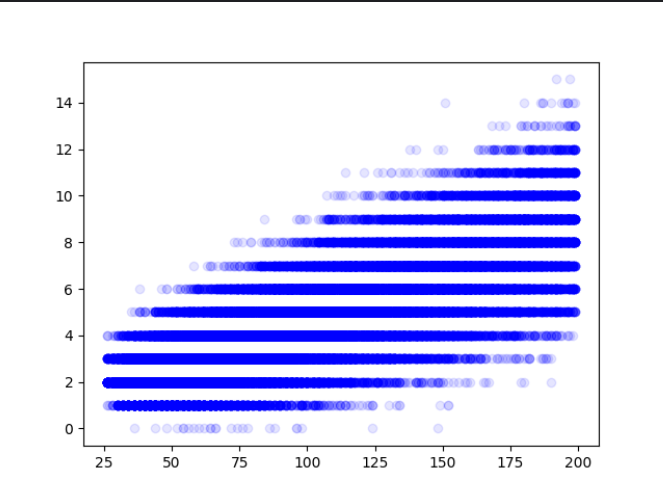
## Care este corelația între numărul de războaie și durata jocului?

Utilizând un grafic care corelează fiecare joc cu numărul de runde și de războaie din acesta putem deriva un grafic care să ne ajute să observăm dacă există sau nu o corelație între aceste 2 măsurători.



Pe axa X avem nr de runde si pe axa Y nr de războaie. Funcție utilizată: runde\_razboaie().

Lipsa punctelor în colțurile stânga sus și dreapta jos ne indică clar că nu există meciuri cu multe războaie și puține runde sau vice versa, arătând o corelație clară între cele 2 valori. Pentru a vedea mai bine putem reduce graficul pentru a scăpa de cele câteva experimente care trag graficul spre valori artificial de mari.



Graficul cu toate experimentele cu nr\_runde < 200. Funcție utilizată: runde\_razboaie().

Observăm că, deși încă este prezentă, corelația nu mai este așa evidentă precum părea în graficul complet. Întrucât magnitudinea numărului de războaie este mai mică ca cea a rundelor, datele par să se „stratifice” însă zonele albastre sunt doar zonele marcate de numerele întregi.