Proiect Probabilitati si Statistica

Capitanu andreea

Florescu Bogdan-ilie

fologea valentin-alexandru

ivascu ioan-andrei

2024

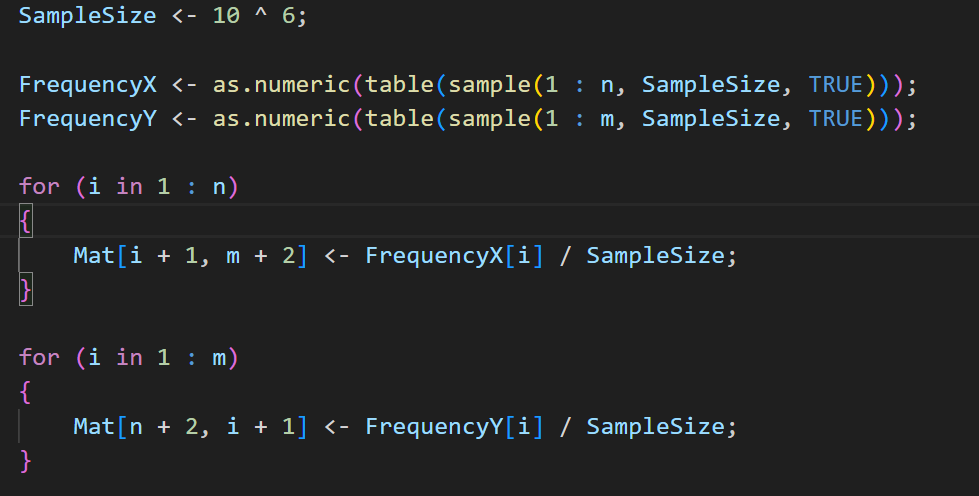
# Exercitiul 1

## Subpunctul a)

Construiţi o funcţie frepcomgen care primeşte ca parametri m şi n şi care generează un tabel cu repartiţia comună a v.a. X şi Y incompletă, dar ȋntr-o formă ȋn care poate fi completată ulterior.

Pentru a rezolva acest exercitiu am dat valori lui X de la 1 la n si lui Y de la 1 la m, in felul acesta ne asiguram ca nu avem valori duplicate.

Pentru a afla probabilitatile lui X si Y am generat 10^6 valori in intervalul [1, n] respectiv [1, m], am calculat frecvențele pentru fiecare valoare a lui X și Y utilizând funcția table și am împărțit rezultatele la 10^6 pentru a obține probabilitățile. Pentru a afla P(X = i, Y = j) am folosit P(X = i) \* P(Y = j). De mentionat ca functia noastra doar variabile aleatoare independente.



In acest mod ne-am asigurat ca repartitia comuna este valida, iar apoi am inceput sa stergem valori din repartitia comuna, astfel incat algoritmul de la subpunctul b sa le poata rezolva.

Daca pe o linie sau pe o coloana stim elementul toate elementele inafara de unul atunci ne putem folosi de suma probabilitatilor pentru a afla ultimul element.

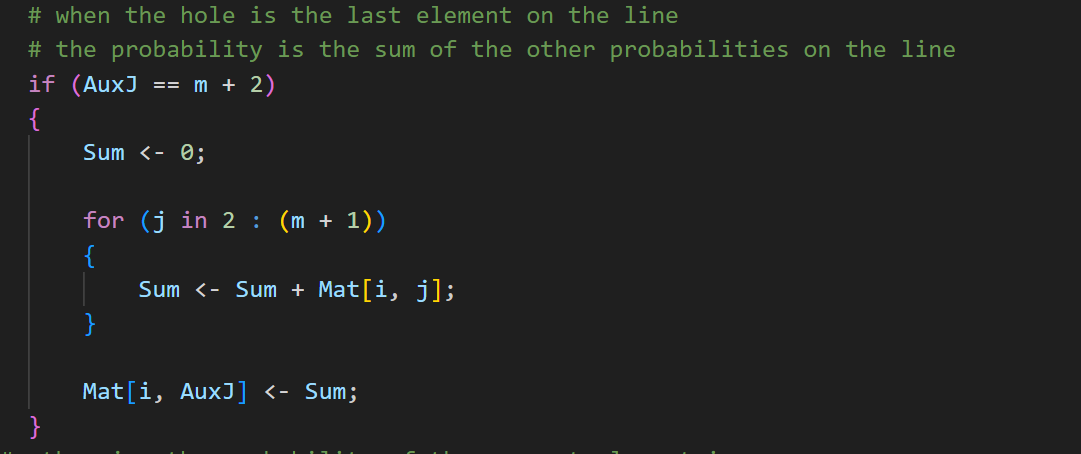
Asadar, pentru a sterge elemente din matrice si ca sa stim sigur ca matricea poate fi rezolvata, cautam linii sau coloane complete si stergem un element aleator de pe acestea. Facem acest lucru pana nu mai exista nicio linie sau coloana completa.

## Subpunctul b)

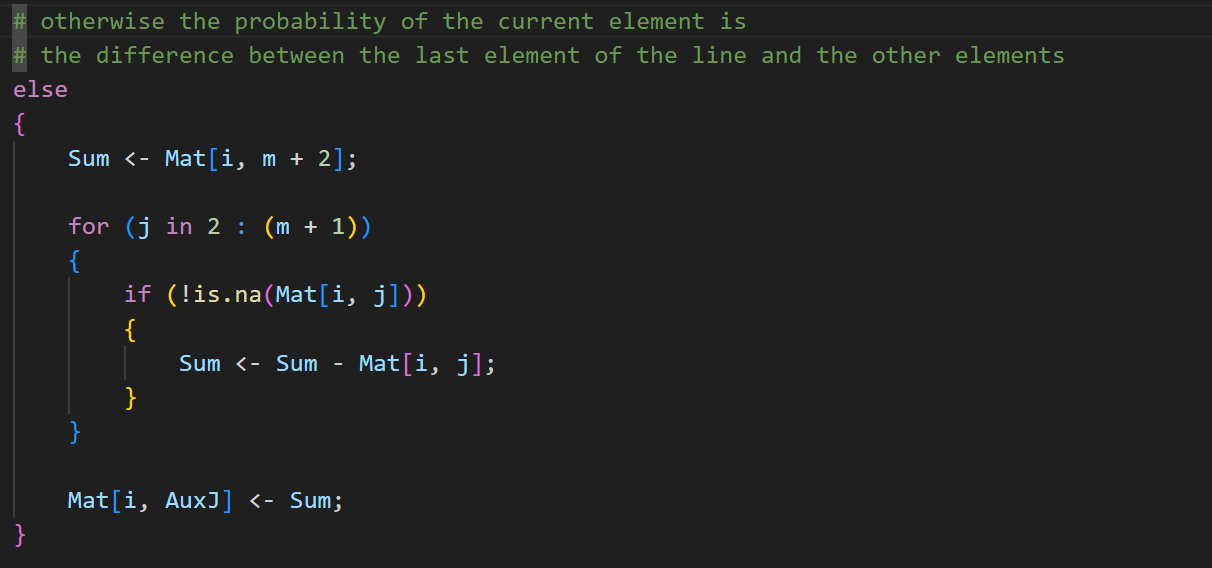
Construiţi o funcţie fcomplrepcom care completează repartiţia comună generată la punctul anterior(pentru cazul particular sau pentru cazul general).

Pentru a rezolva acest subpunct am aplicat un algoritm similar cu cel de la punctul a) doar ca de data aceasta cat timp exista valori goale in repartitia comuna cautam o linie sau coloana care are doar un spatiu gol, si il umplem folosindu-ne de suma suma de pe linia sau coloana respectiva.

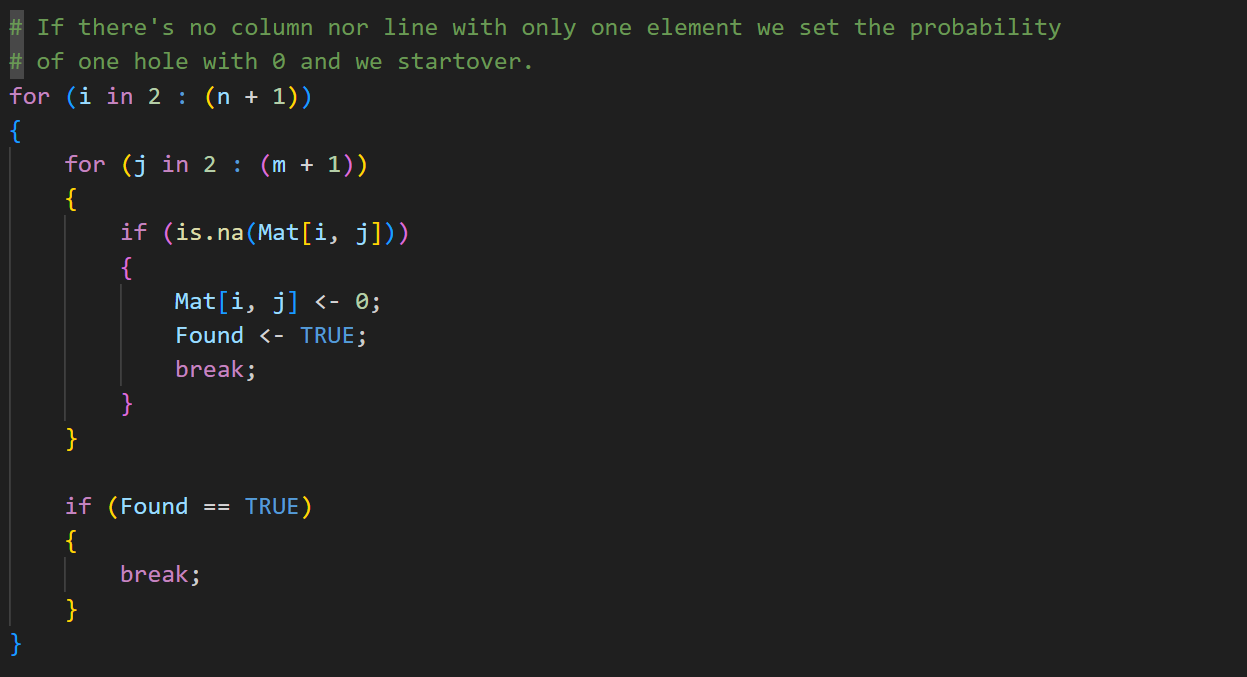
Cand golul este ultimul element de pe linie / coloana atunci probabilitatea sa este suma celorlalte elemente de pe acea linie sau coloana.



Cand golul este inauntrul unei linii/coloane, probabilitatea sa este diferenta dintre ultimul element de pe linia / coloana respectiva si suma celorlalte elemente de pe linia/coloana respectiva.



Daca nu exista nicio linie/coloana unde lipseste un singur element seteaza probabilitatea unui gol aleator cu 0 si se reia procesul.

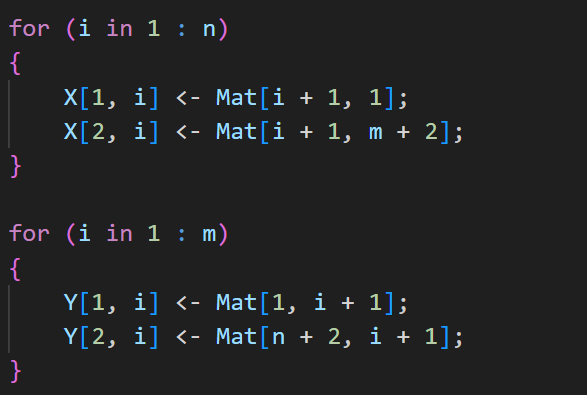


## Subpunctul c)

Construiți o funcție frepmarginal care construiește repartițiile marginale pentru X și Y pornind de la repartiția lor comună.

Repartitia marginala a lui X este formata din prima (valorile lui X) si ultima coloana (probabilitatile lui X) din repartitia comuna.

Repartitia marginala a lui Y este formata din prima (valorile lui Y) si ultima linie (probabilitatile lui Y) din repartitia comuna.

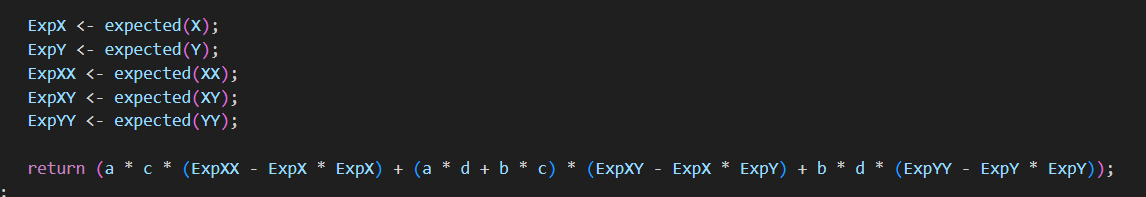


## Subpunctul d)

Construiți o funcție fpropcov care aplică proprietățile covarianței pentru calculul acesteia pentru v.a. Z = aX + bY și respectiv T = cX + dY considerȃnd că toate informațiile necesare despre X și Y sunt date de intrare.

Pentru a rezolva aceasta problema am rescris cov(Z, T), folosind proprietatile covariantei ca: ac \* Var(x) + (ad + bc) \* cov(X, Y) + bd \* Var(Y).

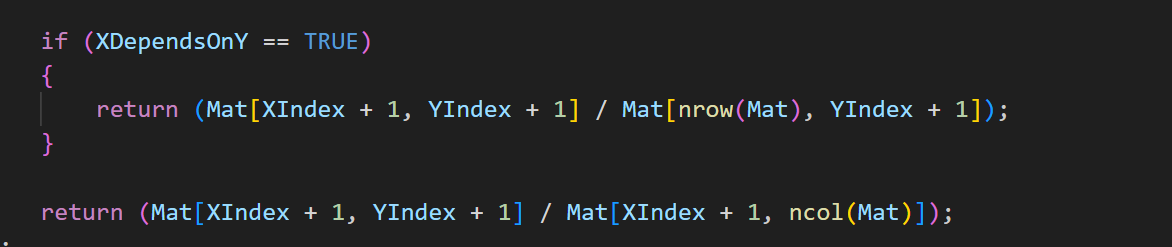
Pentru a calcula Var(X), cov(X, Y) si Var(Y) am creat variabilele aleatoare X^2, XY si Y^2 in memorie.



## Subpunctul e)

Construiți o funcție fPcond care calculează probabilitatea condiționată pentru v.a. X și Y pornind de la repartiția comună.

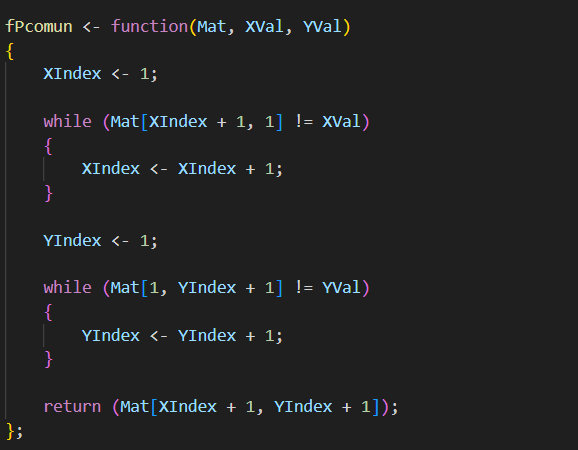
Pentru a rezolva acest subpunct am implementat functia fPcond care primeste repartitia comuna a lui X si Y , un x si un y si tipul de conditionare (x | y sau y | x). Programul cauta linia corespunzatoare lui x si coloana corespunzatoare a lui y si returneaza valoarea de pe indX indY impartita la px sau la qy in functie de tipul de conditionare.



## Subpunctul f)

Construiți o funcție fPcomun care calculează o probabilitate legată de perechea (X,Y) pornind de la repartiția comună.

Pentru a rezolva problema aceasta cautam linia in care X = x si coloana in care Y = y si returnam valoarea de pe pozitia (indX, indY)



## Subpunctul g)

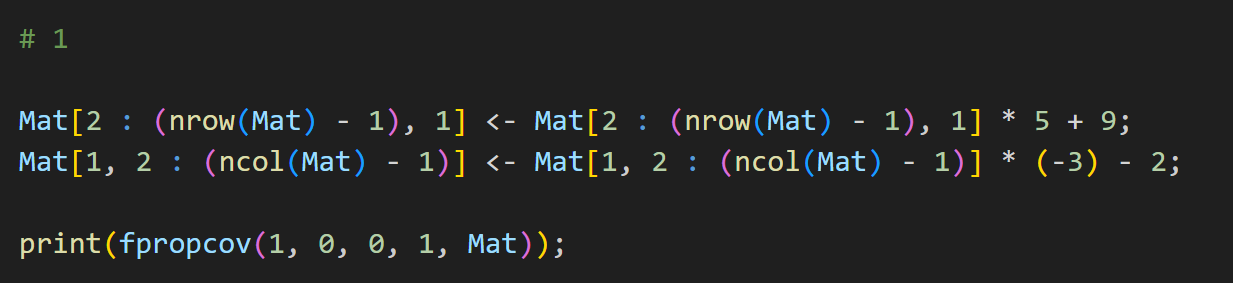
Având la dispoziţie repartiţia comună a v.a. X şi Y de la punctul b) calculaţi:

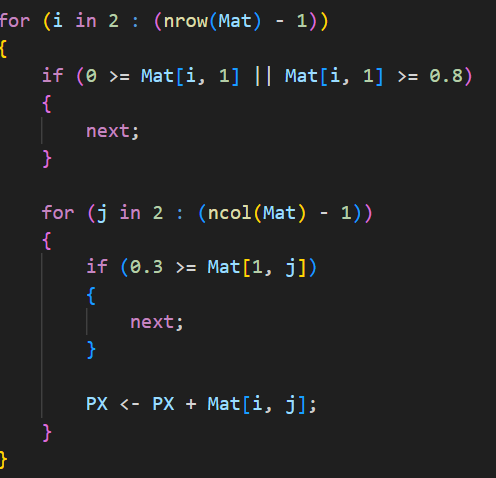
1) Cov(5X+9,-3Y-2)

2) P(0<X<0.8|Y>0.3)

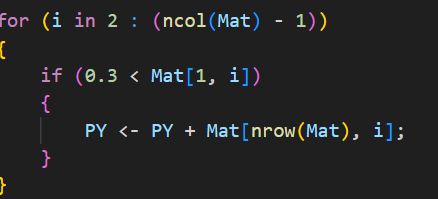
3) P(X>0.2,Y<1.7)

Pentru 1) am aplicat g(x) = 5x + 9 si h(y) = -3y – 2 peste valorile lui X respectiv Y din repartitia comuna, iar apoi ne-am folosit de functia fpropcov() de la subpunctul anterior cu a = 1, b = 0, c = 0, d = 1.

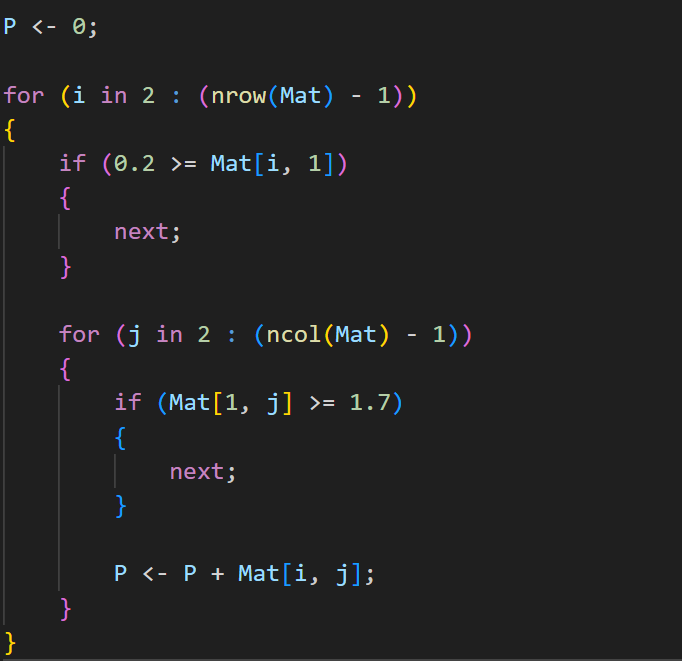


Pentru 2) folosim formula conditionarii P(0<X<0.8|Y>0.3) = P(0<X<0.8, Y>0.3) / P(Y>0.3). Pentru a calcula P(0<X<0.8, Y>0.3) am insumat probabilitatile cand valoarea lui X este mai mare decat 0, mai mica decat 0.8 si valoarea lui Y este mai mare decat 0.3 

Pentru a calcula P(Y > 0.3) am insumat probabilitatile cand Y este mai mare decat 0.3.



Pentru 3) am insumat probabilitatile cand valoarea lui x este mai mare decat 0.2 si cand valoarea lui y este mai mica decat 1.7.



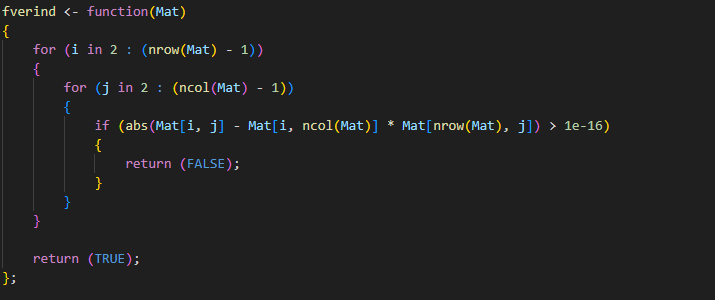
## Subpunctul h)

Pentru exemplul obţinut la punctul b) construiţi două funcţii fverind şi respectiv fvernecor cu ajutorul cărora să verificaţi dacă variabilele X şi Y sunt:

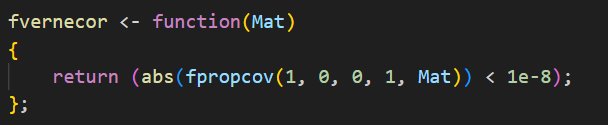
1) independente

2) necorelate

Pentru a verifica daca 2 evenimente sunt independente ne folosim de repartitia comuna a variabilelor x si y. Daca P(X = i, Y = j) = P(X = i) \* P(Y = j) pentru oricare i, j atunci X si Y sunt independente.



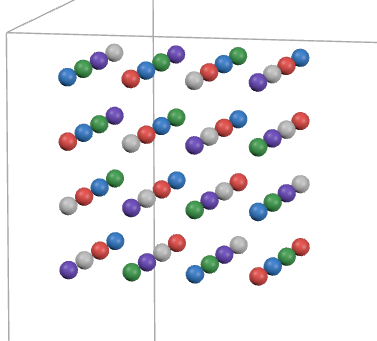
Pentru a verifica daca 2 evenimente sunt necorelate calculam covarianta lui X si Y, iar daca aceasta este 0, inseamna ca si coeficientul de corelatie este 0 si evenimentele sunt necorelate.



## Subpunctul i)

Adăugȃnd ȋncă o v.a. propuneți o manieră vizuală de reprezentare a repartiției comune pentru v.a. X, Y și Z. Care ar fi interpretarea repartițiilor marginale ȋn cazul acestei v.a. tridimensionale și cum ar putea fi obținute?

Pentru a rezolva aceata problema am extins repartitia comuna de la o matrice (2D) la o latice (3D) cu lungimea = N + 2, latimea = M + 2, inaltimea = K + 2 pe care o vizualizam in nivele.



Nivel K+2

Nivel 2

...

Nivel 3

Nivel 1

Formatul laticei este definit in felul urmator:

Colturile toate sunt not assigned, in afara de cel de la M[N + 2, M + 2, K + 2], care este 1.

Partea interioara a laticei este repartitia comuna XYZ, adica M[2 : (N + 1), 2 : (M + 1), 2 : (K + 1)].

Fata de jos si cea de sus sunt repartitia comuna XY. Fata de la stanga si cea de la dreapta sunt repartitia comuna YZ. Fata din fata si cea din spate sunt repartitia comuna XZ. Prin “fata” intelegem doar elementele de pe fata care nu sunt muchii.

Pe muchia M[2 : (N + 1), 1, 1] sunt valorile lui X.

Pe muchia M[1, 2 : (M + 1), 1] sunt valorile lui Y.

Pe muchia M[1, 1, 2 : (K + 1)] sunt valorile lui Z.

Pe muchiile M[2 : (N + 1), M + 2, 1], M[2 : (N + 1), 1, K + 2] si M[2 : (N + 1), M + 2, K + 2] sunt P(X = x).

Pe muchiile M[1, 2 : (M + 1), K + 2], M[N + 2, 2 : (M + 1), 1] si M[N + 2, 2 : (M + 1), K + 2] sunt P(Y = y).

Pe muchiile M[1, M + 2, 2 : (K + 1)], M[N + 2, 1, 2 : (K + 1)] si M[N + 2, M + 2, 2 : (K + 1)] sunt P(Z = z).

Prin “muchie” intelegem doar elementele de pe muchie care nu sunt colturi.

A screen shot of a computer program

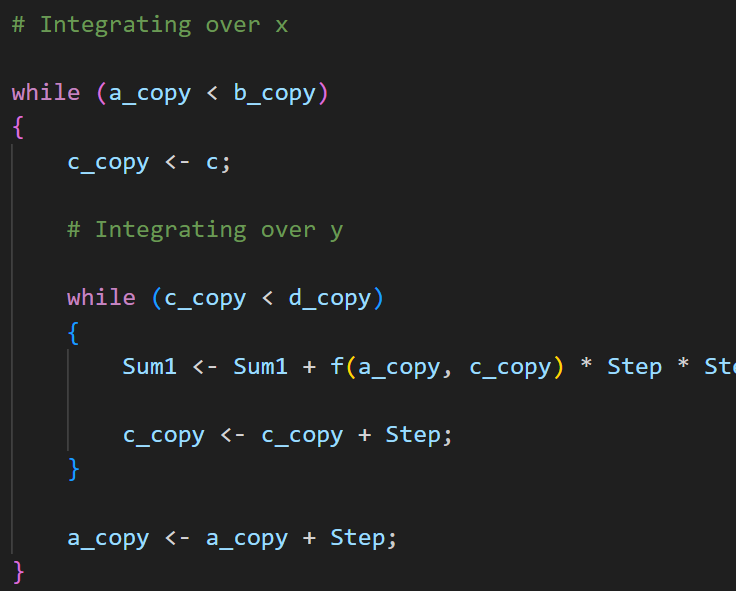
Description automatically generated

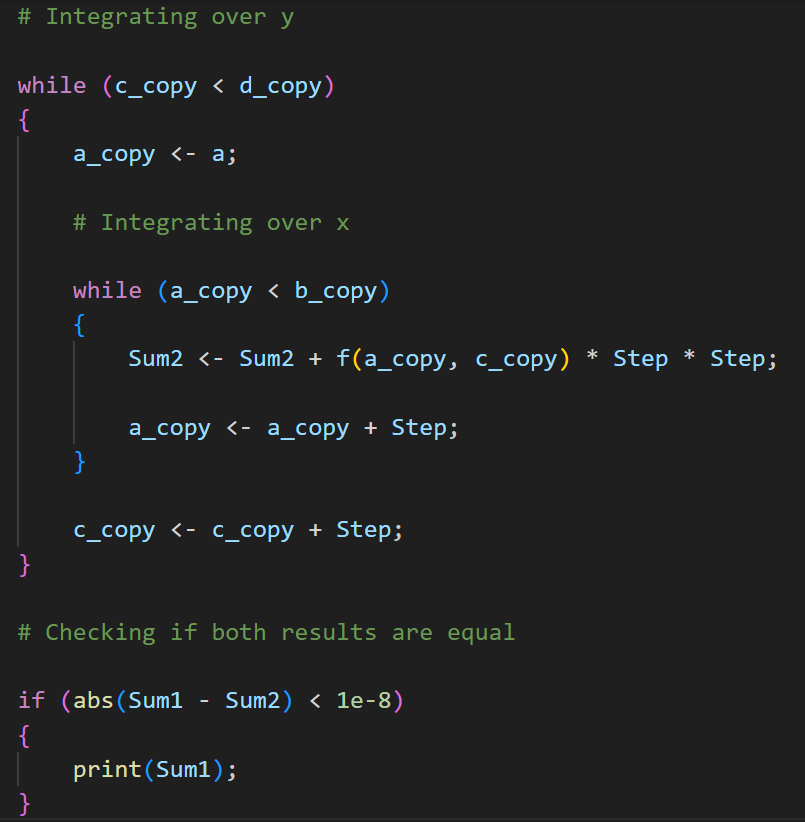
# Exercitiul 2

## Subpunctul a)

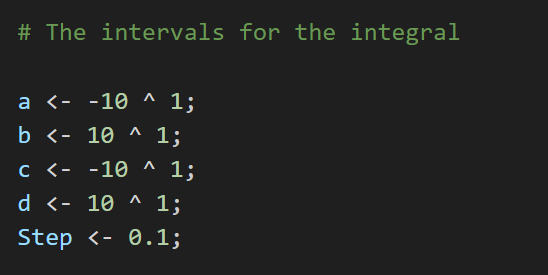
Verificarea posibilitații de aplicare a teoremei lui Fubini pentru calculul integralei duble dintr-o funcție f , introdusă de utilizator și afișarea unui mesaj corespunzător către utilizator. Calculul propriu-zis al integralei ȋn această manieră, atunci cȃnd este posibil.

Pentru a rezolva acest subpunct am creat o functie care primeste o referinta catre functia scrisa de utilizator si verifica aplicabilitatea teoremei lui Fubini integrand functia in ordinea dy, dx si apoi in ordinea dx, dy. Daca rezultatele ambelor integrale sunt egale, atunci putem aplica teorema lui Fubini.





Pentru calculul de integrala simulam pe rezultatul pe x [-10, 10], y [-10, 10] cu un pas de 0.1.

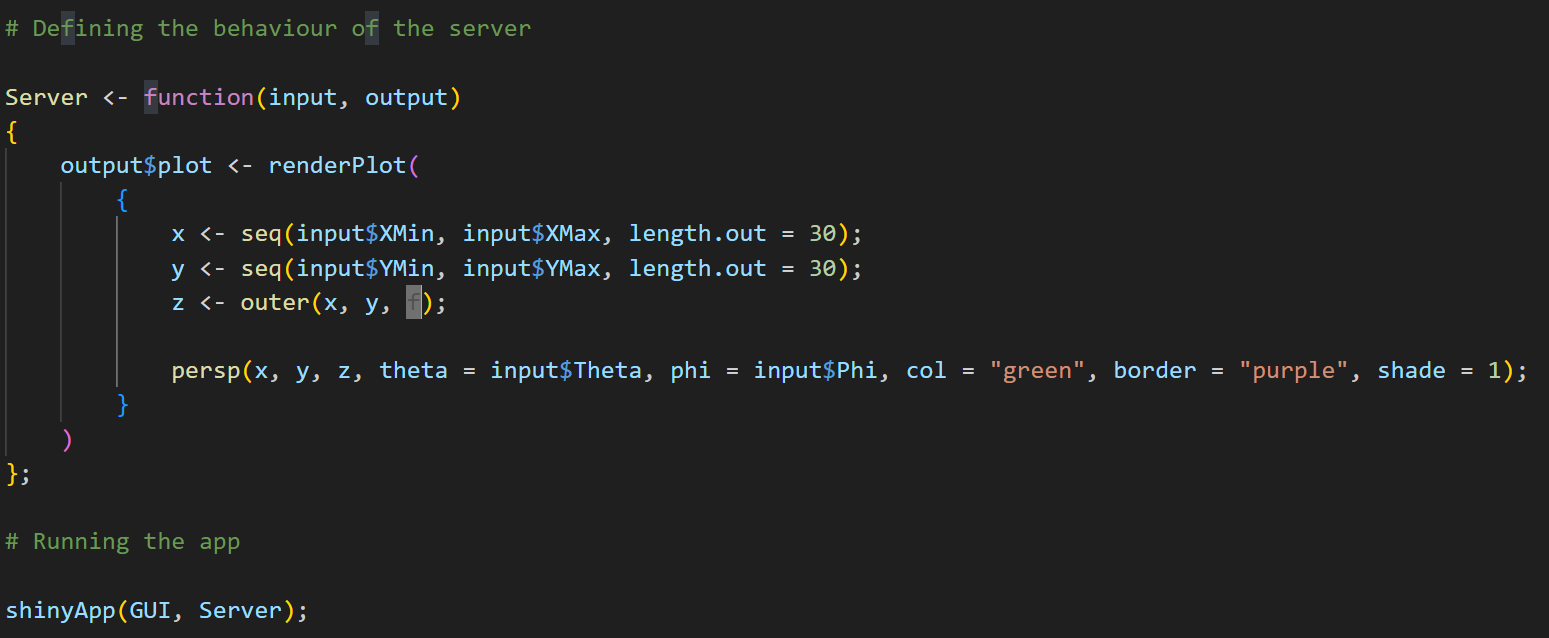


## Subpunctul b)

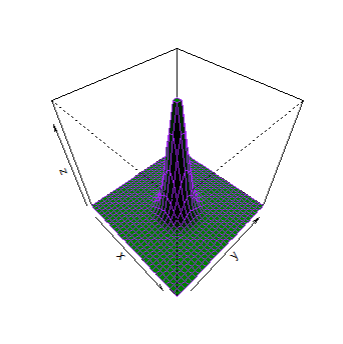
Interpretarea geometrică a integralei duble.

Pentru a rezolva exercitiul acesta am folosit pachetul shiny. UI-ul este de tip fluidPage si prezinta un sidebarPanel de unde utilizatorul poate schimba intervalele pentru x si y, dar si unghiurile de vizualizare ale interpretarii geometrice.





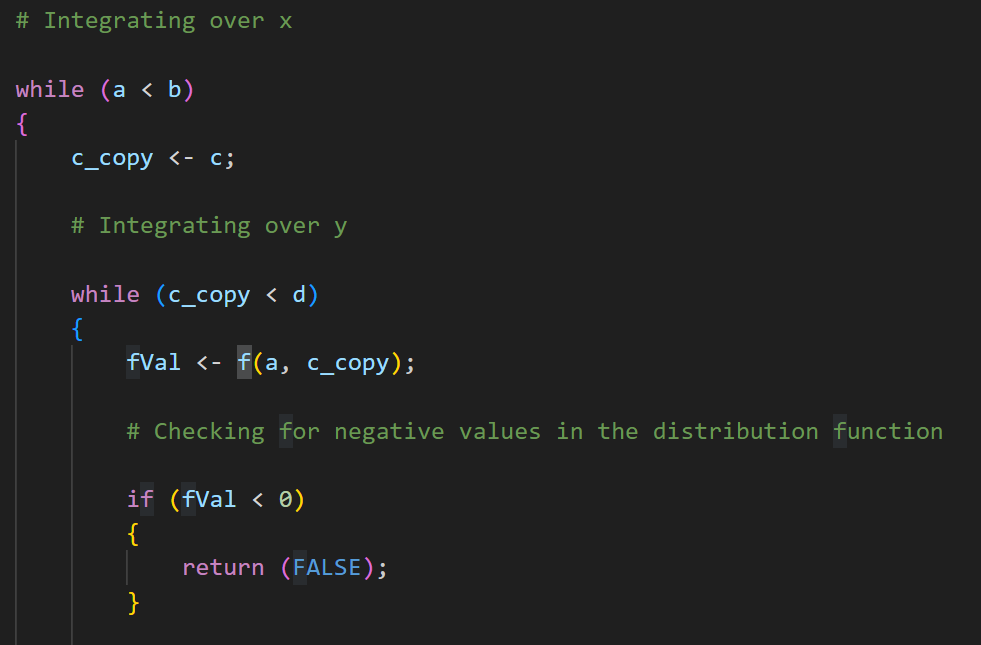
Rezultat obtinut pentru f(x, y) = exp( – ) / :



## Subpunctul c)

Verificarea dacă o funcție cu două variabile f(x,y), introdusă de utilizator este densitate de probabilitate.

Pentru a rezolva acest subpunct simulam calculul integralei in ordinea dy, dx ca la subpunctul anterior, daca in simulare f(i, j) are vreo valoare negativa atunci f nu este densitate de probabilitate.

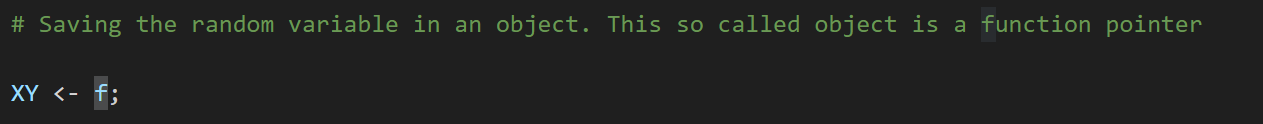


Daca rezultatul integralei este 1 atunci f este densitate de probabilitate.

## Subpunctul d)

Crearea unui obiect de tip variabilă aleatoare continuă pornind de la o densitate de probabilitate introdusă de utilizator. Funcția trebuie să aibă opțiunea pentru variabile aleatoare unidimensionale si respectiv bidimensionale.

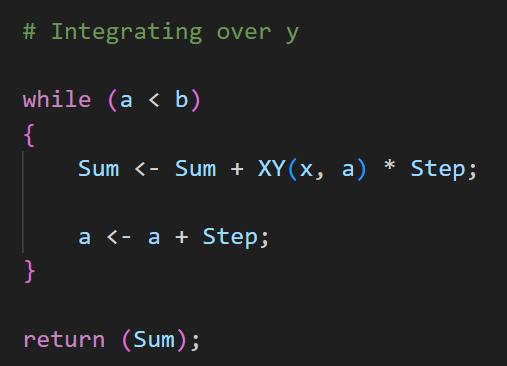
Pentru a rezolva acest subpunct am stocat un pointer la functia f definita de catre utilizator intr-un obiect R.



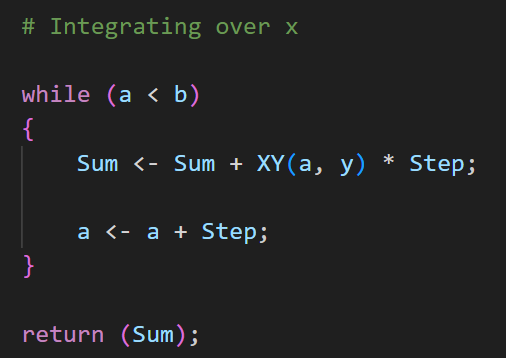
## Subpunctul e)

Construirea densităților marginale și a celor condiționate pornind de la densitatea comună f(x,y) a două v.a. unidimensionale X și Y.

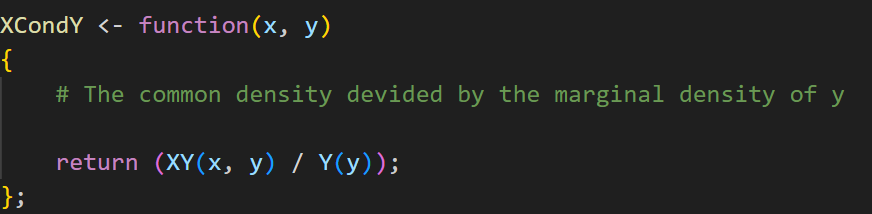
Pentru a construi densitatea marginala a lui X calculam integrala densitatii comune in functie de y.



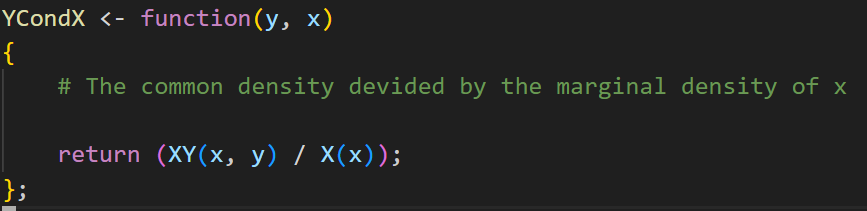
Pentru a construi densitatea marginala a lui Y calculam integrala densitatii comune in functie de x.



Functia de densitate a variabilei X conditionata de Y este raportul dintre densitatea comuna si densitatea marginala a lui Y.



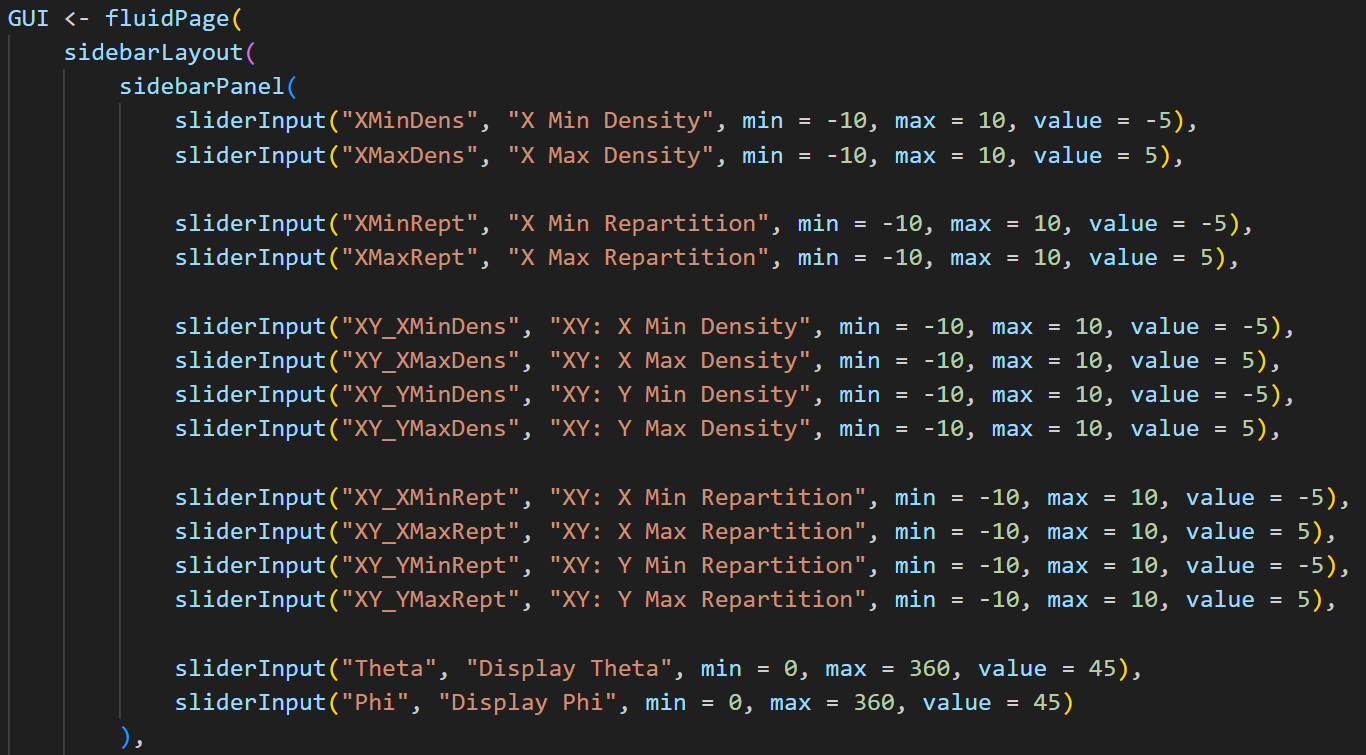
Functia de densitate a variabilei Y conditionata deX este raportul dintre densitatea comuna si densitatea marginala a lui X.

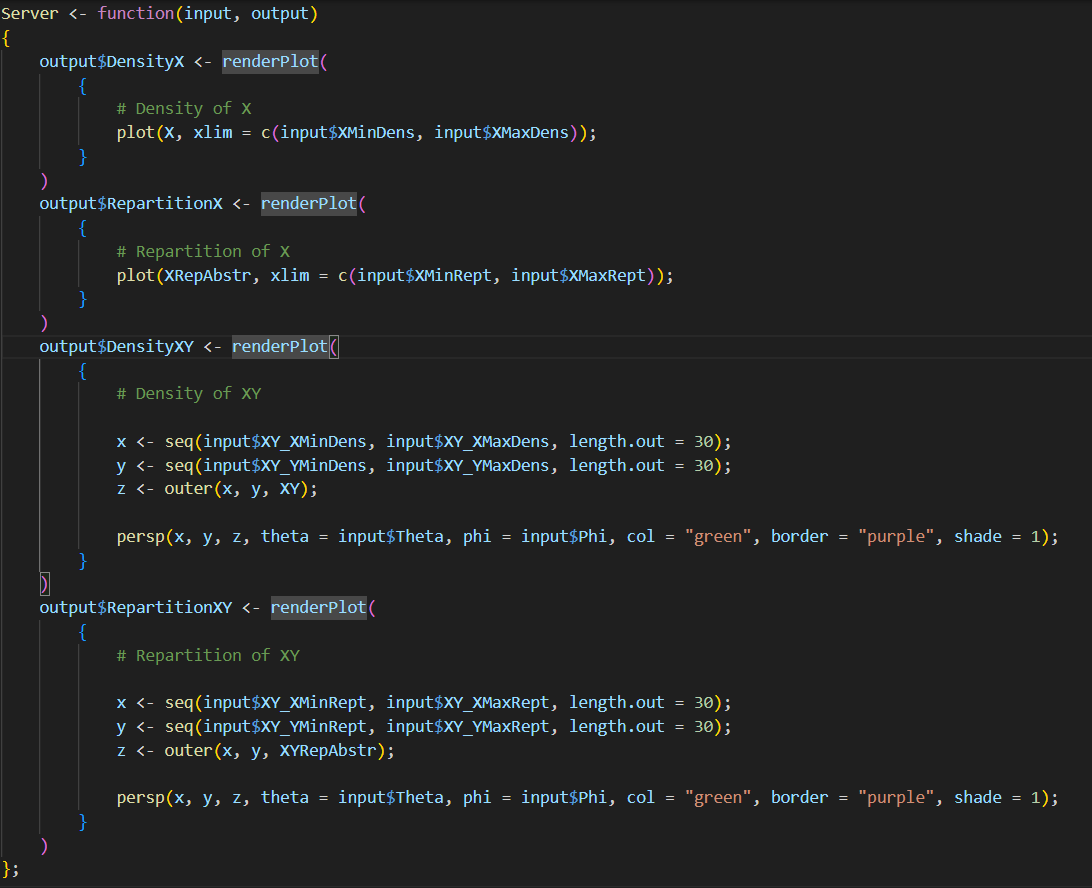


## Subpunctul f)

Reprezentarea grafică a densității și a funcției de repartiție a unei v.a. unidimensionale/bidimensionale pentru diferite valori ale parametrilor repartiției. Ȋn cazul ȋn care funcția de repartiție nu este dată ȋntr-o formă explicită(ex. repartiția normală) se acceptă reprezentarea grafică a unei aproximări a acesteia. Se obține punctaj suplimentar dacă se realizează o animație care să pună ȋn valoare modificarea funcției reprezentate la schimbarea parametrilor repartiției.

Pentru a rezolva exercitiul acesta am folosit pachetul shiny. UI-ul este de tip fluidPage si prezinta un sidebarPanel de unde utilizatorul poate schimba intervalele pentru densitatea minima/maxima si repartitia minima/maxima a lui X, pentru variabilele aleatoare unidimensionale, iar pentru variabilele aleatoare bidimensionale repartitia slidere pentru densitatea minima/maxima a lui X/Y si repartitia minima/maxima a lui X/Y.

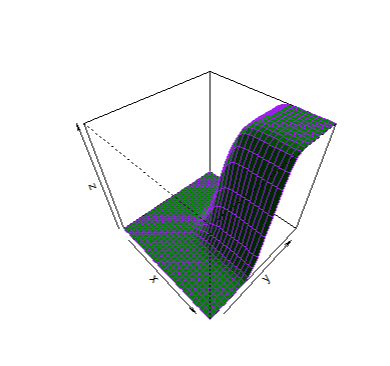




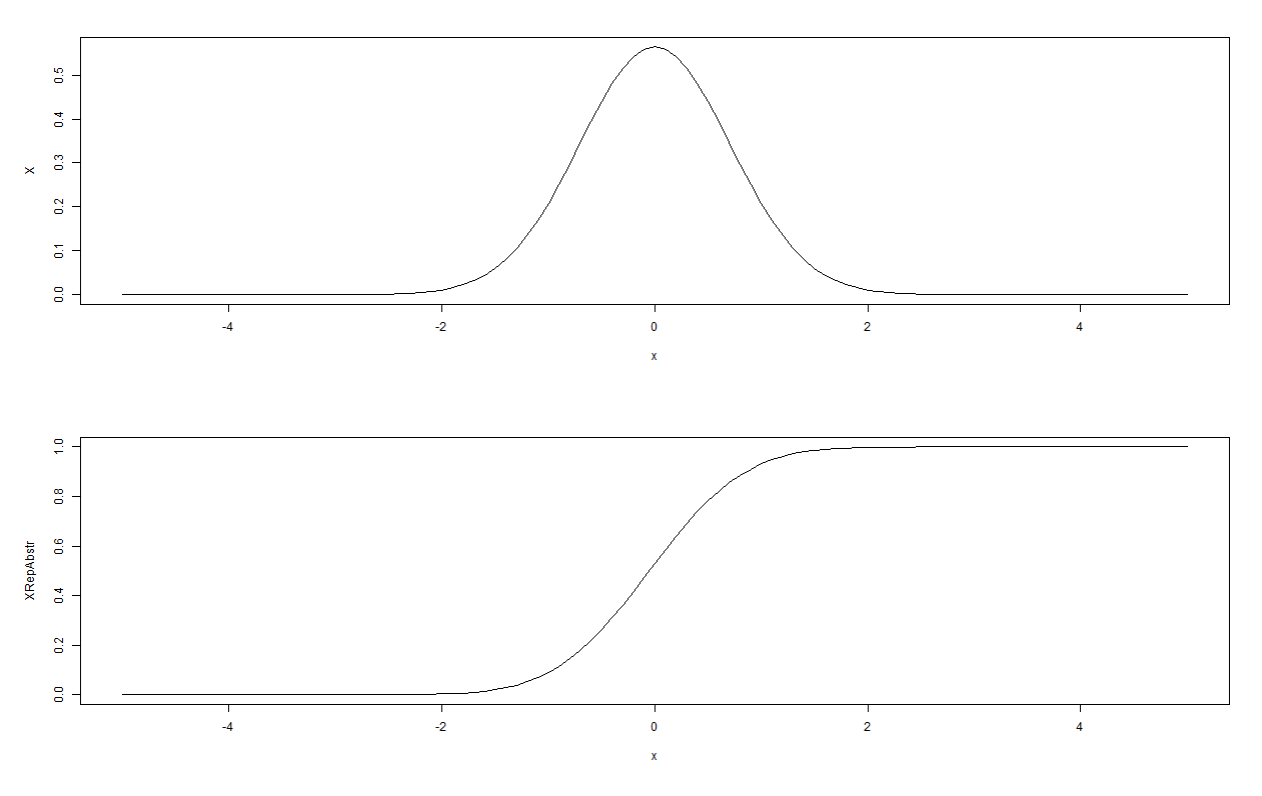
A drawing of a tower

Description automatically generated

Distributie XY = exp(-x ^ 2 - y ^ 2) / 3.1415926535897



Repartitie XY



Distributie si repartitie X

## Subpunctul g)

Calculul mediei, dispersiei și a momentelor inițiale și centrate pȃnă la ordinul 4(dacă există) atȃt pentru v.a. bidimensională cȃt și pentru v.a. unidimensionale ce o compun . Atunci cȃnd unul dintre momente nu există, se va afișa un mesaj corespunzător către utilizator

Am facut functiile Medie, Medie2, MomCentr si MomCentr2 care calculeaza media unei variabile aleatoare continue, media unei variabile aleatoare continue bidimensionale, momentul centrat al unei variabile aleatoare continue si respectiv momentul centrat al unei variabile aleatoare continue bidimensionale. Dupa care le am folosit astfel incat sa obtinem valorile necesare si le-am afisat. In cazul in care momentul centrat avea valoarea 0, afisam pe ecran ”Nu exista”.

A screen shot of a computer program

Description automatically generated

## Subpunctul h)

Calculul mediei și dispersiei unei variabile aleatoare g(X), unde X are o repartiție continuă unidimensională cunoscută iar g este o funcție continuă precizată de utilizator. Se obține punctaj bonus pentru realizarea aceleiași cerințe și pentru cazul bidimensional.

Am modificat functiile de la g) astfel incat sa primeasca un functor care sa fie aplicat pe valoarea variabilei aleatoare. Acest functor in cazul nostru este functia g. Pentru calcului fara functor este definita o functie identitate care este setata ca parametru default. Dupa aceea sunt folosite functiile exact ca la exercitiul anterior.



## Subpunctul i)

Crearea unei funcții P care permite calculul diferitelor tipuri de probabilități asociate unei variabile aleatoare continue unidimensionale/bidimensionale.

Functia P primeste repartitia si capetele intervalelor pe care se doreste calculul de probabilitate. Pentru a obtine rezultatul functia calculeaza integrala pe acest interval/aceste intervale folosind sume Riemann. La final returneaza valoarea obtinuta.

A black background with white text

Description automatically generated

## Subpunctul j)

Calculul covarianței și coeficientului de corelație pentru două variabile aleatoare continue(Atenție:Trebuie să folosiți densitatea comună a celor două variabile aleatoare!)

Pentru a rezolva subpunctul acesta, am folosit formulele covariantei si a coeficientului de corelatie. Ne-am folosit de media lui XY, X si Y de la subpunctul g).

