

Proiect la Probabilități și Statistică

Seria 24

1. Fie două variabile aleatoare discrete X și Y cu repartițiile:

$$X : \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix} \quad Y : \begin{pmatrix} y_1 & y_2 & \dots & y_m \\ q_1 & q_2 & \dots & q_m \end{pmatrix}$$

a) Construiți o funcție **frepcomgen** care primește ca parametri m și n și care generează un tabel cu repartiția comună a v.a. X și Y *incompletă*, dar într-o formă în care poate fi completată ulterior.

Observație: Se cere la a) să generați valorile lui X , valorile lui Y și suficient de multe valori pentru p_i , q_j și respectiv π_{ij} astfel încât să poată fi determinată repartiția comună a celor două v.a.

Nota: În construirea algoritmului puteți începe de la cazul particular $m=2$ și $n=3$. Dacă reușiți să oferiți soluția doar pentru acest caz particular, dar nu și pentru cazul general veți primi punctaj parțial.

b) Construiți o funcție **fcomplrepcom** care completează repartiția comună generată la punctul anterior (pentru cazul particular sau pentru cazul general).

Nota: În cazul în care nu știți să rezolvați punctul a) puteți construi o funcție care să determine repartiția comună pornind de la un exemplu discutat la seminar.

c) Construiți o funcție **frep marginal** care construiește repartițiile marginale pentru X și Y pornind de la repartiția lor comună.

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

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

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

g) Având la dispoziție repartiția comună a v.a. X și Y de la punctul b) calculați:

1) $\text{Cov}(5X+9, -3Y-2)$

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

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

h) Pentru exemplul obținut la punctul b) construiți două funcții **fverind** și respectiv **fvernecor** cu ajutorul cărora să verificați dacă variabilele X și Y sunt:

1) independente

2) necorelate

- i) Adăugând încă o v.a. $Z : \begin{pmatrix} z_1 & z_2 & \dots & z_k \\ r_1 & r_2 & \dots & r_k \end{pmatrix}$ propuneți o manieră vizuală de reprezentare a repartiției comune pentru v.a. X, Y și Z. Care ar fi interpretarea repartițiilor marginale în cazul acestei v.a. tridimensionale și cum ar putea fi obținute?

2. Folosind pachetele R **shiny**(<https://shiny.rstudio.com/>), **animate**(<https://cran.r-project.org/web/packages/animate/vignettes/introduction.html>) și orice alte surse de documentare considerați potriviți construiți un proiect R care să permită lucru cu *variabile aleatoare continue bidimensionale*. Opțiunile din proiect trebuie să implementeze următoarele funcționalități:

- Verificarea posibilității 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 în această manieră, atunci când este posibil.
- Interpretarea geometrică a integralei duble.
- Verificarea dacă o funcție cu două variabile $f(x,y)$, introdusă de utilizator este densitate de probabilitate.
- 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 și respectiv bidimensionale.
- Construirea densităților marginale și a celor condiționate pornind de la densitatea comună $f(x,y)$ a două v.a. unidimensionale X și Y.
- Reprezentarea grafică a densității și a funcției de repartiție a unei v.a. unidimensionale/bidimensionale pentru diferite valori ale parametrilor repartiției. În cazul în care funcția de repartiție nu este dată într-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 în valoare modificarea funcției reprezentate la schimbarea parametrilor repartiției.
- Calculul mediei, dispersiei și a momentelor inițiale și centrate până la ordinul 4 (dacă există) atât pentru v.a. bidimensională cât și pentru v.a. unidimensionale ce o compun. Atunci când unul dintre momente nu există, se va afișa un mesaj corespunzător către utilizator.
- 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.
- Crearea unei funcții **P** care permite calculul diferitelor tipuri de probabilități asociate unei variabile aleatoare continue unidimensionale/bidimensionale.
- Calculul covarianței și coeficientului de corelație pentru două variabile aleatoare continue (**Atenție:** Trebuie să folosiți *densitatea comună* a celor două variabile aleatoare!)

Precizări importante

1. Proiectele se realizează în echipă de 2-4 persoane. Fiecare echipa va desemna un lider care va fi precizat în documentație.
 2. Liderul echipei va trimite pe adresa simona.cojocsa@fmi.unibuc.ro până la data de **4 februarie 2024 ora 22:00** o singură arhivă care va conține fișierele sursă ale proiectului împreună cu documentația.
 3. Documentația este **obligatorie** și lipsa ei atrage necorectarea proiectului.
 4. Documentația trebuie să conțină :
 - numele membrilor echipei
 - descrierea problemei
 - aspecte teoretice folosite în rezolvarea problemei care depășesc nivelul cursului
 - reprezentări grafice și orice altă formă multimedia oportună proiectului
 - precizări privind pachete software folosite și surse de inspirație
 - codul și comentarea acestuia, precum și a soluției prezentate
 - identificarea unor eventuale dificultăți în realizarea cerințelor
 - probleme care au rămas deschise în urma implementării actuale
 - concluzii
- OBS:** Documentația se dorește o prezentare *completă* a proiectului astfel încât evaluarea acestuia să poată fi făcută facil și nu o documentație uzuală pentru un produs software.
5. Dacă la oricare din cele 3 proiecte se realizează cerințe suplimentare față de cele date, cerințe care să fie relevante, se poate obține un bonus de **5p**, fără însă ca nota finală asociată laboratorului să poată depăși **50 p**.