

PRELEGAREA VI STATISTICĂ MATEMATICĂ

I. Variabile aleatoare

6.1 Repartiția și densitatea de probabilitate a unei variabile aleatoare

Caracteristica, variabila studiată din științele experimentale se modelează în teoria probabilităților cu ajutorul variabilelor aleatoare. De exemplu, într-un studiu antropometric, a spune că studentul \mathbf{X} are înălțimea de **1.70 m**, înseamnă că mai întâi s-a selectat studentul \mathbf{X} (**s-a realizat evenimentul**) și apoi s-a făcut măsurarea (**o valoare s-a atribuit evenimentului**). Această mărime, **înălțimea studentului**, este o variabilă aleatoare.

Prin urmare, o **variabilă** se numește **aleatoare** dacă, în cazul mai multor experimente efectuate în aceleași condiții, aceasta primește valori diferite.

Teoretic, o variabilă aleatoare se caracterizează prin valorile primite și probabilitățile de apariție a acestor valori. Mulțimea formată din perechile ordonate **valoare – probabilitatea** corespunzătoare valorii respective definește **repartiția (distribuția) variabilei aleatoare**.

Vom nota cu majuscule bold variabilele aleatoare, \mathbf{X} , \mathbf{Y} , \mathbf{Z} etc., eventual, cu indici, acolo unde este cazul și cu litere mici corespunzătoare valorile primite.

Vom distinge două cazuri:

a) Cazul finit.

Fie $(\Omega, \mathbf{P}(\Omega), \mathbf{P})$ un spațiu de probabilitate finit, unde $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$, $\omega_1, \omega_2, \dots, \omega_n$ fiind evenimente elementare.

Definiția 5.1.1.

Orice funcție definită pe Ω cu valori reale, $\mathbf{X} : \Omega \rightarrow \mathfrak{R}$, se numește **variabilă aleatoare discretă** (cu prescurtarea **v.a.d.**). V.a.d. nu poate lua decât un număr finit de valori întregi și nenegative.

De exemplu, în experiența aruncării cu un zar ideal, rezultatul pe care îl obținem este o v.a.d, $\mathbf{X} : \Omega \rightarrow \mathfrak{R}$. \mathbf{X} poate lua numai valori discrete, $\mathbf{x}_i = 1, 2, 3, 4, 5, 6$, unde $i = 1, 2, 3, 4, 5, 6$, iar $\mathbf{P}(\mathbf{X} = \mathbf{x}_i) = \frac{1}{6}$. Suma tuturor probabilităților $\mathbf{P}(\mathbf{X} = \mathbf{x}_i)$ este egală cu 1.

Deci, fiecărui eveniment elementar ω_i i se asociază un număr real $\mathbf{X}(\omega_i)$. Cum aceste numere nu sînt în mod necesar distincte între ele, vom nota cu $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$, $k \leq n$, valorile distincte luate de v.a. \mathbf{X} . Se notează cu:

$\mathbf{A}_i = \{\omega \mid \mathbf{X}(\omega) = \mathbf{x}_i\}$, $i = 1, 2, \dots, k$ acele evenimente, nu neapărat elementare, pentru care v.a.d. \mathbf{X} primește valori diferite. Altfel spus, \mathbf{A}_i conține toate evenimentele a căror realizare conduce la luarea valorii \mathbf{x}_i de către v.a.d. \mathbf{X} .

Familia $\{\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_k\}$ formează o partiție a mulțimii de bază Ω , numită **partiția indusă** de v.a.d. \mathbf{X} . Orice eveniment elementar aparține unui eveniment \mathbf{A}_i și numai unuia singur.

Plecînd de la un eveniment \mathbf{A} , după cum se observă în figura de mai jos, se pot evidenția cele două noțiuni studiate, legate de acesta, probabilitatea evenimentului \mathbf{A} , $\mathbf{P}(\mathbf{A})$ și valoarea $\mathbf{X}(\mathbf{A})$, asociată prin intermediul v.a.d. \mathbf{X} .

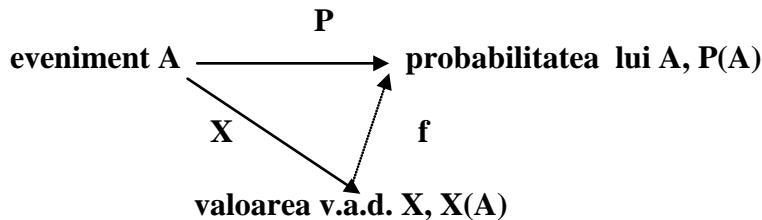


Figura 6.1.1. Tipuri de corespondență

Legătura marcată prin linie punctată precizează că **v.a.d. X ia valori cu anumite probabilități**, și anume cu probabilitățile cu care se realizează evenimentele asociate valorilor lui X. Astfel, se poate considera **funcția de probabilitate** (sau **legea de probabilitate, repartiția de probabilitate**) a v.a.d. X, care se definește în modul următor:

$f : \{x_1, x_2, \dots, x_k\} \rightarrow [0, 1]$, unde $f(x_i) = P(A_i)$, $i = 1, 2, \dots, k$.
 Funcția de probabilitate a v.a.d. X se notează fie

$$X: \begin{pmatrix} x_1 & x_2 & \dots & x_k \\ p_1 & p_2 & & p_k \end{pmatrix}, \text{ unde } p(x_i) = f(x_i) \text{ reprezintă probabilitatea cu}$$

care X ia valoarea x_i , fie

$$p_i = P(X = x_i), i = 1, 2, \dots, k.$$

Reprezentarea grafică a repartiției de probabilitate se face în felul următor: pe abscisă se reprezintă valorile posibile ale v.a.d., iar pe ordonată probabilitățile corespunzătoare.

De exemplu, funcția de probabilitate a variabilei X definită, în experiența aruncării cu un zar ideal, se prezintă în tabelul 6.1.1, iar reprezentarea ei grafică în figura 6.1.2.

x_i	1	2	3	4	5	6
$P(X = x_i)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Tabelul 6.1.1. Distribuția variabilei discrete X

Dacă se consideră lățimea dreptunghiurilor formate egală cu unitatea, atunci suma ariilor dreptunghiurilor este egală cu 1, după cum se observă în figura de mai jos.

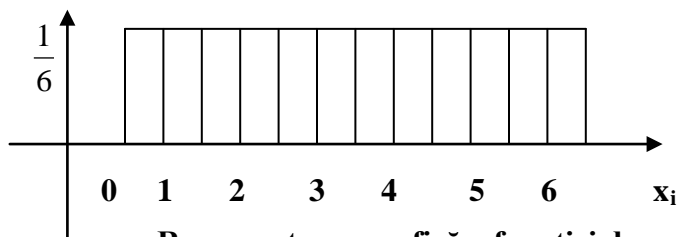


Figura 6.1.2. Reprezentarea grafică a funcției de probabilitate a v.a.d. X

Următoarele proprietăți ale funcției de probabilitate rezultă imediat din definiția și proprietățile probabilității:

1) $f(x_i) \geq 0$, $i = 1, 2, \dots, k$;

$$2) \sum_{i=1}^k f(x_i) = \sum_{i=1}^k P(X = x_i) = P(\Omega) = 1.$$

Interpretarea celor două relații este asemănătoare cu cea de la frecvențe relative, adică de a fi nenegative și de a avea suma egală cu 1.

Noțiunii de frecvență relativă cumulată îi corespunde **funcția de repartiție** a v.a.d. X , definită astfel: $F : \mathfrak{R} \rightarrow \mathfrak{R}$, $F(x) = P(X \leq x)$, unde prin $X \leq x$ s-a notat evenimentul:

$\{\omega \mid X(\omega) \leq x\}$, adică reuniunea acelor evenimente elementare pentru care v.a.d. ia valori mai mici sau egale cu x .

Funcția de repartiție a v.a. X are ca valoare, într-un punct oarecare x , probabilitatea ca v.a.d. să ia valori mai mici sau egale cu x .

Din relația $P(X \leq x) = \sum_{x_i \leq x} P(A_i)$, rezultă că suma tuturor probabilităților evenimentelor A_i determină valorile v.a. X mai mici sau egale cu x . Altfel spus, intervin probabilitățile cumulate.

Grafic, funcția de repartiție a variabilei X definită mai sus, în experiența aruncării cu un zar ideal, se prezintă ca o funcție în trepte care conține puncte de salt:

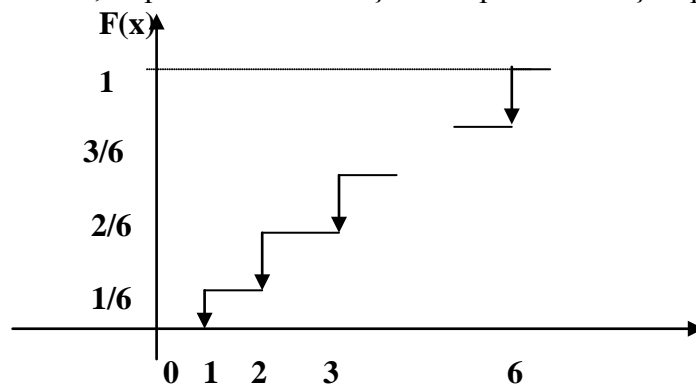


Figura 6.1.3. Reprezentarea grafică a funcției de repartiție a v.a.d. X

Probabilitatea ca numărul pe care îl obținem la aruncarea unui zar să fie mai mică decât 1 este zero, $F(1) = P(X < 1) = 0$. Continuând, vom avea:

$$F(2) = P(X < 2) = P(X = 1) = f(1) = \frac{1}{6};$$

$$F(3) = P(X < 3) = P(X = 1) + P(X = 2) = \frac{1}{6} + \frac{1}{6} = \frac{2}{6};$$

$$F(4) = P(X < 4) = \sum_{i=1}^3 P(X = i) = \frac{3}{6};$$

$$F(5) = P(X < 5) = \sum_{i=1}^4 P(X = i) = \frac{4}{6};$$

$$F(6) = P(X < 6) = \sum_{i=1}^5 P(X = i) = \frac{5}{6};$$

$$F(x > 6) = P(X < x) = \sum_{i=1}^6 P(X = i) = 1;$$

$$F(-\infty) = 0 \text{ și } F(+\infty) = 1.$$

Deci, suma tuturor salturilor este egală cu 1.

Definiția 6.1.2.

Două variabile aleatoare discrete X și Y definite pe o aceeași mulțime de bază Ω sînt independente dacă evenimentele partițiilor induse de X și Y sînt independente două cîte două, adică pentru orice i și j are loc:

$P(X = x_i \cap Y = y_j) = P(X = x_i) \cdot P(Y = y_j)$. Se observă că, v.a.d. X și Y iau valori independente una de alta. Definiția se poate generaliza la un număr nedeterminat de v.a.d.

b) Cazul infinit numărabil.

Fie (Ω, K, P) un spațiu de probabilitate.

Definiția 6.1.3.

Orice funcție definită pe Ω cu valori reale, $X : \Omega \rightarrow \mathfrak{R}$, care satisface condiția $\{\omega \mid X(\omega) \leq x\} \in K$, oricare ar fi $x \in \mathfrak{R}$, se numește **variabilă aleatoare continuă**, sau, pe scurt, **variabilă aleatoare** sau **v.a.** Condiția pusă în definiție exprimă faptul că “mulțimea evenimentelor elementare pentru care v.a. X are valori mai mici decît x , oricare ar fi x , trebuie să fie tot un eveniment”. În cazul infinit, are sens să cerem că v.a. X să aparțină unui interval, și nu probabilitatea să ia o anumită valoare. Această interpretare este apropiată de practica experimentală, deoarece, de exemplu, șansa de a găsi un individ cu înălțimea de exact **1.75 m** este aproape nulă dacă măsurarea se face cu eroare zero. De obicei, printr-o asemenea valoare se înțelege un întreg interval de înălțimi, toate acelea care prin rotunjire la două zecimale devin **1.75m**.

Această variabilă fiind susceptibilă de a lua o infinitate numărabilă de valori, nu se poate reprezenta printr-o diagramă cu bare, ci printr-o curbă de distribuție.

Funcția de repartiție a v.a. X se definește la fel ca în cazul finit:

$F(x) = P(X \leq x)$, unde prin $X \leq x$ s-a notat evenimentul $\{\omega \mid X(\omega) \leq x\}$ (aici, evenimentele elementare, fiind o infinitate, ele nu se pot decît cel mult enumera).

Probabilitatea ca v.a. X să fie cuprinsă într-un interval foarte mic “ dx ” centrat în x_0 este :

$$P((x_0 - \frac{dx}{2}) < X < (x_0 + \frac{dx}{2})) = f(x_0)dx = dF(x_0),$$

unde $f(x_0)$ este ordonata curbei în punctul de abscisă x_0 . Se observă că, $f(x)$ este derivata funcției de repartiție. Invers, $F(x)$ este integrală din $f(x)$.

Dacă există, funcția f se numește **densitatea de probabilitate** a v.a. X .

Prin convenție, $P(x_{\min} \leq X < x_{\max}) = 1$, unde x_{\min} și x_{\max} sînt valorile extreme ale lui X . De cele mai multe ori, acestea sînt $-\infty$ și $+\infty$.

Definiția 6.1.4.

Funcția de repartiție este **absolut continuă**, dacă există o funcție reală $f : \mathfrak{R} \rightarrow \mathfrak{R}$, astfel

$$\text{încît } F(x) = \int_{-\infty}^x f(u)du .$$

Grafic, funcția de repartiție se interpretează ca fiind: **mărimea ariei de sub graficul funcției f** din figura de mai jos:

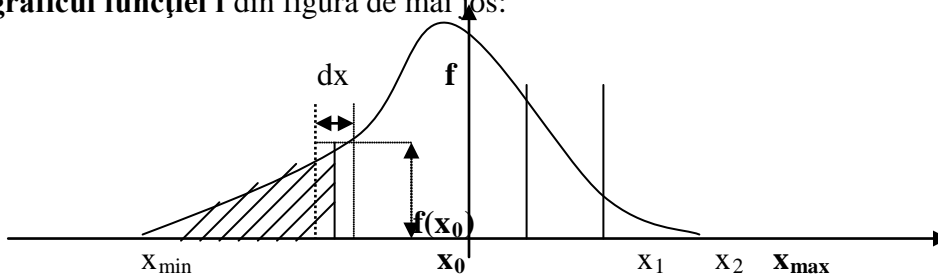


Figura 6.1.4. Reprezentarea grafică a funcției de repartiție a unei v.a.continue

Deci, $F(x_0) = P(X \leq x_0)$ este porțiunea hașurată de arie de sub graficul funcției $f(x)$ și în stînga dreptei $x = x_0$.

Ca și în cazul finit, densitatea de probabilitate a v.a. X are următoarele două proprietăți:

$$1) f(x) \geq 0, \text{ oricare ar fi } x; \quad 2) \int_{-\infty}^{+\infty} f(u)du = 1.$$

Principalele proprietăți ale funcției de repartiție sînt:

- 1) $0 \leq F(x) \leq 1$;
- 2) F este nedescrescătoare (dacă $x_1 < x_2$, atunci $F(x_1) \leq F(x_2)$);
- 3) $\lim_{x \rightarrow -\infty} F(x) = 0$, cînd $x \rightarrow -\infty$ și $\lim_{x \rightarrow +\infty} F(x) = 1$, cînd $x \rightarrow +\infty$;
- 4) $P(x_1 \leq X < x_2) = P(X < x_2) - P(X \leq x_1) =$

$$F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(u)du, \text{ dac\u0103 } F \text{ este func\u021bia de reparti\u021bie absolut}$$

continu\u0103 a v.a. X .

Prima proprietate rezult\u0103 din defini\u021bia unei probabilit\u0103\u021bi, a doua arat\u0103 c\u0103 probabilitatea unui eveniment nu poate s\u0103 scad\u0103 dac\u0103 se reune\u0219te cu un alt eveniment, a treia precizeaz\u0103 probabilitatea evenimentului sigur \u0219i respectiv a celui imposibil, iar ultima egalitate se utilizeaz\u0103 la calculul probabilit\u0103\u021bilor, atunci c\u00e2nd se cunoa\u0219te func\u021bia de reparti\u021bie.

\u00c2n situa\u021bia variabilelor aleatoare continue, reparti\u021bia cu largi aplica\u021bii \u00een domeniul practicii este **reparti\u021bia normal\u0103**, care va fi studiat\u0103 mai t\u00e2rziu, \u00eenpreun\u0103 cu reparti\u021biile **normal normat\u0103**, **log-normal\u0103**, χ^2 , **Fisher-Snedecor**, **Student**, **binomial\u0103**, etc.

II. Reparti\u021bii teoretice

Reparti\u021bia (distribu\u021bia) unei v.a. constituie caracteristica de baz\u0103 a v.a. respective, \u00een\u021beleg\u00e2nd prin aceasta fie reparti\u021bia ei de probabilitate, fie densitatea ei de reparti\u021bie. De aceea, unele reparti\u021bii teoretice au o mare importan\u021b\u0103 \u00een practic\u0103, deoarece pot constitui modelul comport\u0103rii unor date observabile din care, utiliz\u00e2nd calculul de probabilit\u0103\u021bi, pot fi trase concluzii foarte probabile \u0219i luate decizii corecte \u00een domeniul explorat. Cele mai importante r\u0103m\u00e2n reparti\u021bia Bernoulli \u0219i extensiile ei pentru cazul discret \u0219i reparti\u021biile Gauss-Laplace \u0219i normal normat\u0103 pentru cazul continuu.

6.2.1. Repartiția binomială

Repartiția binomială (repartiția Bernoulli sau repartiția newtoniană sau legea alternativei simple) se definește pentru o variabilă aleatoare discretă și are ca punct de plecare schema bilei revenite:

Într-o urnă se află în total N bile albe și negre. Probabilitatea de a extrage o bilă neagră este p (evenimentul E) și rămâne aceeași pe toată perioada experimentului. O bilă extrasă se repune în urnă (adică, o extragere neexhaustivă sau non exhaustivă). Se cere probabilitatea ca în urma a n extrageri să se obțină k bile negre (desigur, $n-k$ vor fi albe), adică evenimentul E să apară (să se realizeze) de k ori și de $n-k$ să nu apară (v.a X). Valoarea maximă a lui k fiind n .

Operația de obținere a unui rezultat într-o astfel de schemă, și nu numai, poartă denumirea de încercare (o probă), iar o suită de încercări formează un experiment.

În modelarea unei v.a. discrete după o repartiție binomială trebuie să se respecte condițiile:

- variabila trebuie să fie dihotomică, adică să primească doar două valori (alb/negru, bun/rău, bărbat-femeie, rece/cald, cifra 1/tot ceea ce nu este cifra/ 1, etc.), notate cu succes/insucces;

- procedura de obținere a încercărilor (probelor, extragerilor) trebuie să fie aceeași pe toată durata experimentului;

- încercările trebuie să fie independente între ele (probele nu se pot influența reciproc);

- probabilitățile asociate valorilor succes/insucces trebuie să fie aceleași de-a lungul experimentului. Se notează cu $p = P(\text{succes})$ (adică, probabilitatea obținerii unei valori de succes), iar cu $q = P(\text{insucces}) = 1 - p$ (deci, $p + q = 1$).

Fie X o variabilă aleatoare repartizată binomial. Valorile ei posibile reflectă numărul de succese apărute într-un experiment cu n încercări succesive. Acestea pot fi: $0, 1, \dots, n$.

În condițiile date, se demonstrează că probabilitățile ca X să primească aceste valori sînt date de expresia:

$$P(X = k) = C_n^k p^k q^{n-k}, k = 0, 1, 2, \dots, n.$$

De exemplu, presupunînd că în n încercări de extragere a bilei din urnă s-au obținut în ordinea: **NANNAAAN...** Fiecare din cele n extrageri succesive fiind independente, probabilitatea extragerii complete în această ordine este egală cu produsul probabilităților elementare (principiul probabilității compuse), adică: $p \cdot q \cdot p \cdot p \cdot q \cdot q \cdot p \dots = p^k \cdot q^{n-k}$. Ținînd cont că numărul de permutări cu repetiții în aranjarea a două tipuri de obiecte, k bile negre și $n - k$ bile albe, este numărul de combinații C_n^k . Probabilitatea căutată este:

$$P(X = k) = C_n^k p^k (1-p)^{n-k}, k = 0, 1, 2, \dots, n.$$

Se notează $P(X = k)$ cu $P(n; k)$ sau $P_n(k)$, iar repartiția binomială cu parametrii n și p , $Bi(n; p)$, unde n , p și q au semnificațiile de mai sus.

Observații. 1. Cînd $p = 0.5$, formula se simplifică și devine:

$$P(X = k) = C_n^k 0.5^n. \text{ Este cazul jocului cap sau pajură;}$$

2. Dacă N este suficient de mare, ($N > 10 \cdot n$), atunci ne putem dispensa de condiția de a repune în urnă bila extrasă, extragerea fiind considerată exhaustivă;

3. Dacă extragerea este una exhaustivă, atunci se aplică legea hipergeometrică.

4. Valorile repartiției binomiale sînt tabelate pentru diferite valori ale lui **n** și **p**.

Proprietăți.

1. $M(X) = \mu = np$. Expresia mediei se obține plecînd de la dezvoltarea binomului lui Newton și considerînd identitatea:

$$(px + q)^n \equiv \sum_{k=0}^n C_n^k p^k x^k q^{n-k} = \sum_{k=0}^n x^k P_n(k).$$

Derivînd identitatea în raport cu **x** se obține:

$$\sum_{k=0}^n kx^{k-1} P_n(k) \equiv n \cdot p \cdot (px + q)^{n-1}. \text{ Apoi, egalînd pe } x \text{ cu } 1, \text{ avem:}$$

$$\sum_{k=0}^n k P_n(k) = n \cdot p.$$

Funcția $(px + q)^n$ se numește **funcția generatoare** a legii binomiale.

2. $D^2(X) = \sigma^2 = \mu_2 = npq = np(1 - p)$;

3. $\sigma = \sqrt{npq}$;

4. Pentru calculul mai ușor al probabilității $P_n(k)$ se folosește formula de recurență: $P_n(k+1) = \frac{n-k}{k+1} \cdot \frac{p}{1-p} P_n(k)$;

5. Moda (valoarea dominantă sau valoarea cea mai probabilă) este partea întregă a expresiei $p \cdot (n+1)$:

$p \cdot (n+1) - 1 < Mo < p \cdot (n+1)$. Această valoare se obține ținînd cont de definiția modei: $Mo > P(Mo - 1)$ și $Mo > P(Mo + 1)$;

6. Funcția de probabilitate este:

$$X: \left(\begin{array}{cccccc} 0 & 1 & 2 & \dots & n \\ C_n^0 p^0 q^n & C_n^1 p^1 q^{n-1} & C_n^2 p^2 q^{n-2} & \dots & C_n^n p^n q^0 \end{array} \right).$$

Se observă că în linia a doua apar expresiile termenilor din dezvoltarea binomului $(p + q)^n$;

7. Funcția de repartiție (sau probabilitățile cumulate) este:

$$F(x) = \sum_{k=0}^x P(n; k) = \sum_{k=0}^x C_n^k p^k q^{n-k}, \quad x \in \mathfrak{R}.$$

Graficul funcției de repartiție este specific unei distribuții în trepte (o histogramă). Dacă $p = q = 0.5$, graficul este simetric. În caz contrar, este asimetric, asimetria fiind dictată de dispropoția dintre **p** și **q**.

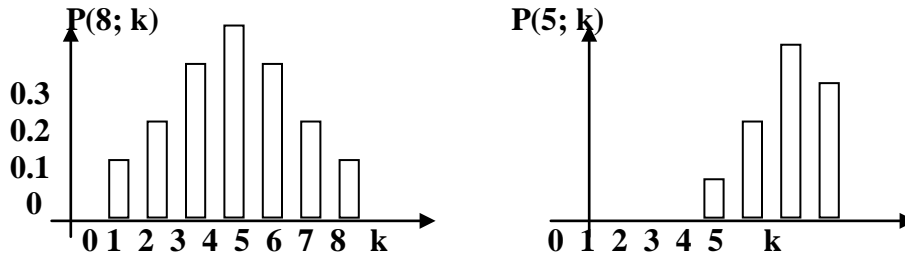


Figura 6.2.1.1. Distribuțiile binomiale $Bi(8; 0.5)$ și $Bi(5; 0.8)$

8. Pentru aplicații, valorile celor două funcții sînt tabelate.

Aplicații. 1. Care este probabilitatea ca să se obțină **2** de **3** și o altă cifră, atunci când se aruncă **3** zaruri?

În aruncarea unui zar, există două posibilități:

a) să se obțină fața **3** cu probabilitatea de $p = \frac{1}{6}$;

b) o altă cifră diferită de **3** cu probabilitatea de $1 - p = \frac{5}{6}$.

Se aplică legea binomială pentru $n = 3$ și $k = 2$:

$$P(X = 2) = C_3^2 \cdot (1/6)^2(5/6)^1 = 0.07, \mu = n \cdot p = 3 \cdot \frac{1}{6} = 0.5,$$

$$\sigma^2 = n \cdot p \cdot q = 3 \cdot \frac{1}{6} \cdot \frac{5}{6} = 0.08.$$

2. Știind că un lot de produse conține **8%** piese defecte. Care este probabilitatea ca extrăgând **10** piese să se obțină mai mult de **2** (mai mult de **20%**) piese defecte?

De fapt, se cere să se afle $P(k > 2)$ sau $P(k \geq 3)$ sau în că $1 - P(k \leq 2)$. Extrăgând din tabela cu probabilități cumulate corespunzătoare legii binomiale pentru $n = 10$ și $p = 8\%$, se obține: $P(k \leq 2) = 0.9599$. Deci, probabilitatea ca extrăgând **10** piese să se obțină mai mult de **două** piese defecte este egală cu: $1 - 0.9599 = 0.0401$.

Observație. Pentru un număr suficient de mare de observații, repartiția binomială $Bi(n, 1/2)$ poate fi folosită, mai ales, în două tipuri de aplicații;

- verificarea apropierii măsurătorilor efectuate de o repartiție normală, în situația în care datele sînt grupate într-un număr de clase;
- gradarea unei scale continue pentru o variabilă care se presupune ca este continuă și pentru care există un număr suficient de mare de observații, intervalele claselor fiind considerate egale între ele.

Aplicație. Repartizarea în clase se face ținînd cont de numerele conținute în **triunghiul lui Pascal**: **o clasă: 1**

2 clase: 1 1

3 clase: 1 2 1

4 clase: 1 3 3 1

5 clase: 1 4 6 4 1

..... Trei reguli stau la baza construcției

triunghiului:

- prima linie are un singur **1**;
- o linie ulterioară începe și se încheie cu **1**;
- un număr intermediar se obține ca suma celor două numere situate pe linia anterioară deasupra și la stînga.

De exemplu, **250** de observații pentru o variabilă repartizată normal pot fi distribuite în **5** clase, astfel:

$$\frac{a}{1} = \frac{b}{4} = \frac{c}{6} = \frac{d}{4} = \frac{e}{1} = \frac{a+b+c+d+e}{1+4+6+4+1} = \frac{250}{16} = 15.62, \text{ notînd cu } a, b, c, d \text{ și } e$$

efectivele căutate. Se obțin:

$$a = e = 1 \times 15.62 = 15.62;$$

$$b = d = 4 \times 15.62 = 62.48;$$

$$c = 6 \times 15.62 = 93.72.$$

Prin rotunjire cu adaos, avem: $a = 16$, $b = 62$, $c = 94$, $d = 62$, $e = 16$.

Aceste rezultate pot fi utilizate în teste de comparare pentru a decide dacă repartiția variabilei studiate se ajustează printr-una normală.

6.2.2. Repartiția binomială modificată

Dacă se consideră ca v.a. numărul de defecte (k) dintr-un eșantion, unde frecvența (f) de defecte dintr-un eșantion de dimensiune n este:

$$f = \frac{k}{n}, \text{ atunci v.a } f$$

urmează o lege binomială cu media: $\mu = p$ și abaterea medie pătratică: $\sigma = \sqrt{\frac{p \cdot q}{n}}$.

6.2.3. Extensii ale legii binomiale

6.2.3.1. Repartiția multinomială

Am văzut că v.a care urmează o lege binomială trebuie să fie dihotomică. Lărgirea acestei condiții se adresează extragerilor non exhaustive efectuate asupra unor populații cu mai mult de 2 categorii de obiecte.

Fie o populație cu k categorii în proporțiile $p_1, p_2, p_3, \dots, p_k$ asupra căreia se efectuează o extragere non exhaustivă de n obiecte.. Evident:

$$p_1 + p_2 + p_3 + \dots + p_k = 1 \text{ și } n_1 + n_2 + n_3 + \dots + n_k = n.$$

Probabilitatea ca cele n obiecte să fie repartizate în:

- n_1 obiecte din categoria (1),

- n_2 obiecte din categoria (2),

.....

- n_k obiecte din categoria (k) este dată de **legea multinomială**:

$$P(n_1, n_2, \dots, n_k) = \frac{n!}{n_1! n_2! \dots n_k!} p_1^{n_1} p_2^{n_2} \dots p_k^{n_k}.$$

Aplicație. O urnă conține 30% bile negre, 20% bile albe și 50% bile roșii. Care este probabilitatea ca, efectuând o extragere non exhaustivă de 6 bile, să se obțină 3 bile negre, 2 bile albe și o bilă roșie?

Se utilizează legea multinomială:

$$P(3_N, 2_A, 1_R) = \frac{6!}{3!2!1!} 0.3^3 \cdot 0.2^2 \cdot 0.5^1 = 0.0324.$$