Variabile aleatoare

Sorana D. BOLBOACĂ

- Variabile aleatoare
- Distribuții de probabilitate:
  - Distribuția normală
  - Distribuția binomială
  - ....

Cuprins

»Fie X o variabilă cantitativă măsurată sau observată rezultată dintr-un experiment

»Valoarea pe care o ia variabila X în urma experimentului este o variabilă aleatoare

»**Exemple**:

- Numărul de globule roşii dintr-un frotiu
- Numărul de bacterii de pe mâinile studenților
- Scorul de depresie mediu obținut la aplicarea unui test pe un eşantion de pacienți cu patologie terminală

Definiție

- Media aritmetică a eşantionului
- Deviația standard
- Proporția
- Frecvența
  - Toate sunt variabile aleatoare

Variabile aleatoare

Discrete:

» Poate lua un număr finit măsurabil de valori

- Numărul de persoane cu RH- dintr-un eşantion
- Numărul de copii cu gripă dintr-o colectivitate
- Numărul de studenți anorexici din universitate
- Pulsul

Continue:

» Poate lua orice valoare din nenumăratele valori posibile într-un interval definit

» Variază în mod continuu în intervalul dat

- Temperatura corporală
- Concentrația zahărului în sânge
- Tensiunea arterială

Tipuri de variabile aleatoare

»În general, mediile sunt variabile aleatoare continue iar frecvențele sunt discrete

- Media capacității pulmonare a unei persoane care munceşte în domeniul minier.
- Numărul de pacienți cu edentație parțială sau totală din Cluj.

Tipuri de variabile aleatoare

Spațiul unui eveniment

» Fie X numărul de fețe "cap" obținute la aruncarea de 3 ori a unei monede

» X este o variabilă aleatoare care poate lua una din următoarele valori {0,1,2,3}

Spațiul unui eveniment

» Dintr-un sac care conține bile albe şi negre sunt extrase 2 bile. La extragerea unei bile albe se câştigă 1 Ron iar la extragerea unei bile negre se pierde 1 Ron.

» X este o variabilă aleatoare care poate lua una din valorile {-2,0,2}

Distribuția de probabilitate: variabila aleatoare discretă

Sorana D. BOLBOACĂ

»Probabilitatea distribuției lui X: listă de valori ale spațiului de evenimente şi probabilitățile asociate acestora

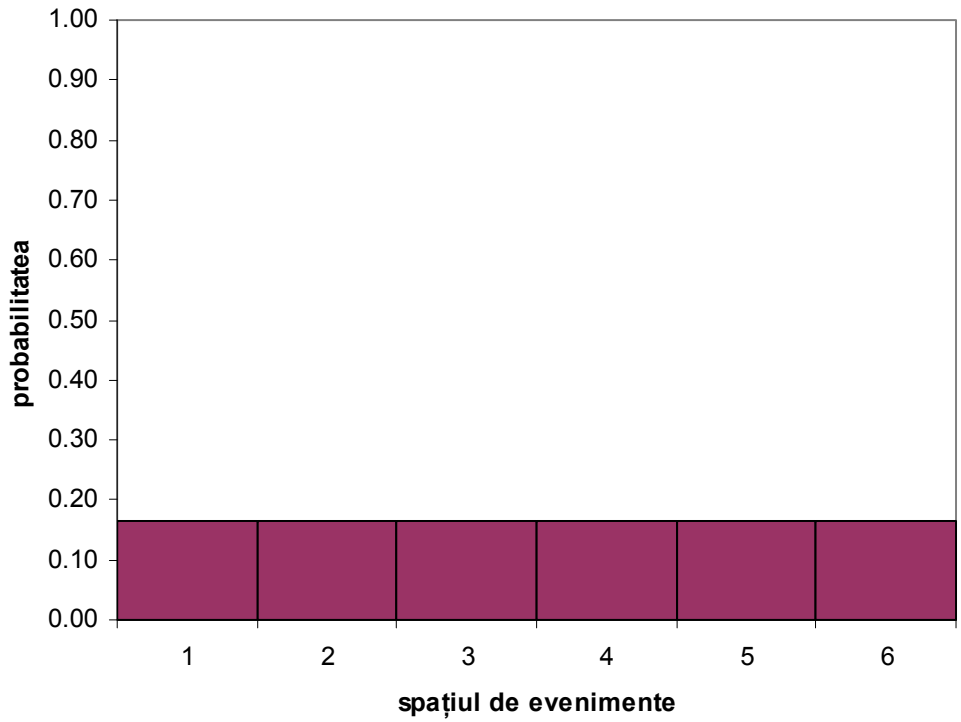
- Fie X rezultatul aruncării unui zar
- X este o variabilă aleatoare care ia una din următoarele valori 1, 2, 3, 4, 5, 6

X_i	Pr_i
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6

Distribuția de probabilitate: variabila aleatoare discretă

Sorana D. BOLBOACĂ

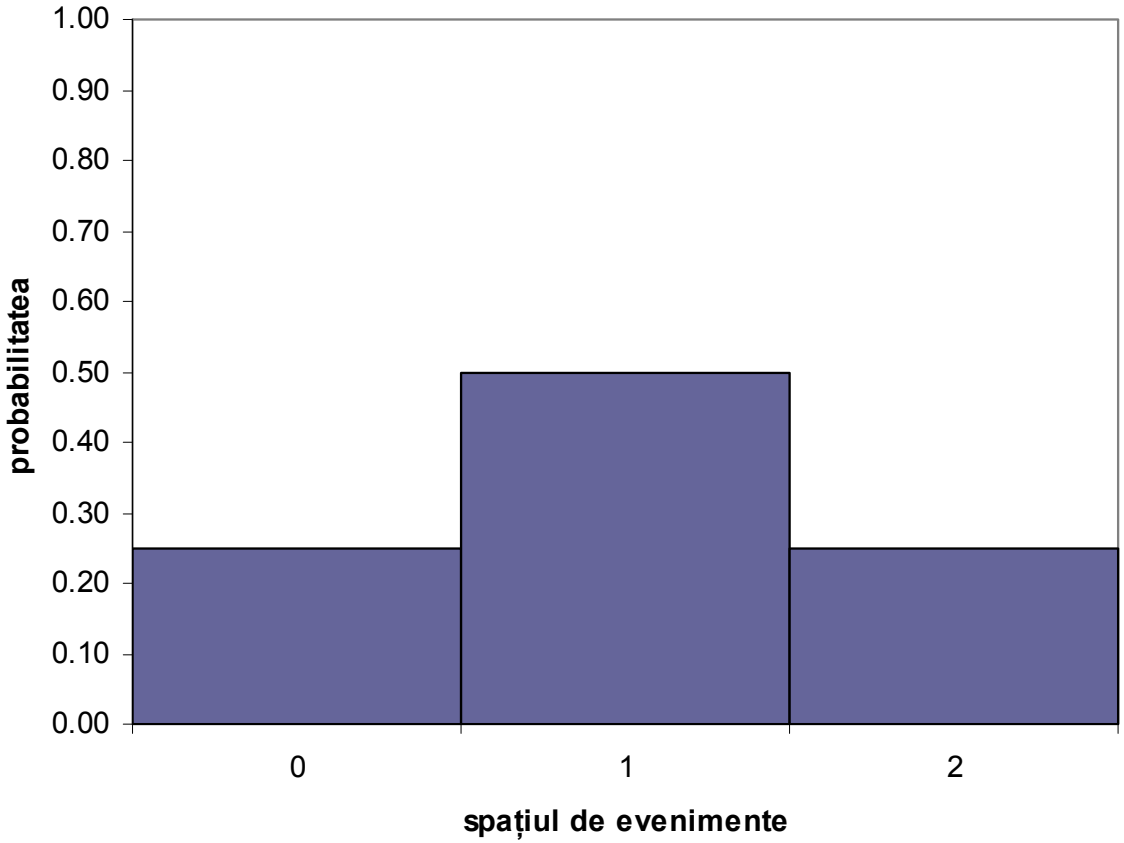
»Probabilitatea distribuției lui X listează valorile spațiului de evenimente şi probabilitățile asociate



X_i	Pr_i
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6

Distribuția de probabilitate: variabila aleatoare discretă

» Fie X numărul de fețe ‘cap’ rezultate la aruncarea a două monezi de două ori. Care este distribuția de probabilitate?



X_i	Pr_i
0	1/4
1	2/4
2	1/4

Distribuția de probabilitate: variabila aleatoare discretă

» Legea de probabilitate: simbolistică

» **Proprietate:** probabilitățile care apar în distribuția unei variabile aleatoare finite X verifică

$$X:\begin{pmatrix} X_1 & X_2 & \dots & X_n \\ \Pr(x_1) & \Pr(x_2) & \dots & \Pr(x_n) \end{pmatrix}$$

$$\sum_{i=1}^{n} \Pr(X_i) = 1$$

Distribuția de probabilitate: variabila aleatoare discretă

» **Media** distribuției de probabilitate discretă (denumită şi **valoare expectată** sau **speranța matematică**) este dată de formula

$$M(X) = \sum_{i=1}^{n} X_i \cdot Pr(X_i)$$

» Este media ponderată a valorilor posibile, fiecare valoare fiind ponderată cu probabilitatea ei de apariție

Distribuția de probabilitate: variabila aleatoare discretă

Exemplu:

» Fie X o variabilă aleatoare reprezentând numărul de episoade de otită în primii doi ani de viață într-o colectivitate. Această variabilă aleatoare are distribuția:

$$X:\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 \\ 0.129 & 0.264 & 0.271 & 0.185 & 0.095 & 0.039 & 0.017 \end{pmatrix}$$

» Care este numărul aşteptat (mediu) de episoade de otită în primii doi ani de viață?

Distribuția de probabilitate: variabila aleatoare discretă

$$X:\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 \\ 0.129 & 0.264 & 0.271 & 0.185 & 0.095 & 0.039 & 0.017 \end{pmatrix}$$

- Care este numărul aşteptat (mediu) de episoade de otită în primii doi ani de viață?
- $M(X) = 0\cdot0.129 + 1\cdot0.264 + 2\cdot0.271 + 3\cdot0.185 + 4\cdot0.095 + 5\cdot0.039 + 6\cdot0.017$
- $M(X) = 0 + 0.264 + 0.542 + 0.555 + 0.38 + 0.195 + 0.102$
- $M(X) = 2.038$

Distribuția de probabilitate: variabila aleatoare discretă

»Variația: media ponderată a pătratului deviației lui X

$$V(X) = \sum_{i=1}^{n} (X_i - M(X))^2 \cdot Pr(X_i)$$

»Abaterea standard sau ecartul tip:

$$\sigma(X) = \sqrt{V(X)} = \sqrt{\sum_{i=1}^{n} (X_i - M(X))^2 \cdot Pr(X_i)}$$

Distribuția de probabilitate: variabila aleatoare discretă

X_i	$Pr(X_i)$	$X_i*Pr(X_i)$	$X_i-M(X)$	$(X_i-M(X))^2$	$(X_i-M(X))^2*Pr(X_i)$
0	0.129	0	-2.038	4.153	0.536
1	0.264	0.264	-1.038	1.077	0.284
2	0.271	0.542	-0.038	0.001	0.000
3	0.185	0.555	0.962	0.925	0.171
4	0.095	0.38	1.962	3.849	0.366
5	0.039	0.195	2.962	8.773	0.342
6	0.017	0.102	3.962	15.697	0.267
		M(X)=**2.038**			V(X)=**1.967**
					σ(X)=**1.402**

Variabila aleatoare discretă: V(X),σ(X)

- **Bernoulli** (cap versus pajură): două rezultate posibile
- **Binomială** (numărul de ‘cap’ în *n* încercări): variabile aleatoare finite
- **Poisson** (numărul de pacienți care sunt consultați în serviciul de urgență într-o zi): variabile aleatoare discrete infinite

Principalele distribuții de probabilitate: variabile aleatoare discrete

- »Un experiment e alcătuit din repetarea unei încercări elementare de *n* ori (*n* = un număr natural dat)
- »Rezultatele posibile ale fiecărei încercări elementare sunt două evenimente numite *succes* şi *eşec*
- »Probabilitatea de succes este notată cu *p* iar probabilitatea de eşec este notată cu *q* (*q = 1-p*)
- »Cele *n* încercări repetate sunt independente

Distribuția Binomială

»Numărul X de succese obținute în cele n încercări este o variabilă aleatoare de tip binomial care depinde de parametrii n şi p şi se notează cu Bi(n,p)

»Variabila aleatoare X poate să ia valorile 0,1,2,...n

»Probabilitatea ca X să fie egal cu o valoare k este dată de formula:

$$\Pr(X = k) = C_n^k p^k q^{n-k}$$

$$C_n^k = \frac{n!}{k! \cdot (n-k)!}$$

Distribuția Binomială

»Media sau speranța matematică a distribuției binomiale: $M(X) = n \cdot p$

»Variația: $V(X) = n \cdot p \cdot q$

»Abaterea standard: $\sigma(X) = \sqrt{(n \cdot p \cdot q)}$

»Binomial? 4 condiții de verificat:

- Experimentele sunt independente.
- Numărul de experimente (n) este fix.
- Rezultatul unui experiment poate fi clasificat ca succes sau eșec.
- Probabilitatea de succes (p) este identică pentru fiecare experiment.

Distribuția Binomială

»Care este probabilitatea de ca din 5 copii ai unei familii 2 să fie băieți dacă probabilitatea de a naşte un băiat este de 0,47 pentru fiecare naştere şi sexul copiilor născuți succesiv în familie este considerat o variabilă aleatoare independentă?

»p=0,47

»q=1-0,47=0,53

»n=5

»k=2

»**Pr(X=2)=10·0,47²·0,53³**

»**Pr(X=2) = 0,33**

$$\Pr(X = k) = C_n^k p^k q^{n-k}$$

$$C_5^2 = \frac{5!}{2!\cdot(5-2)!} = \frac{5\cdot 4\cdot 3\cdot 2\cdot 1}{2\cdot 1\cdot(3\cdot 2\cdot 1)} = \frac{120}{12} = 10$$

Distribuția Binomială

» Variabila aleatoare POISSON ia o infinitate numărabilă de valori: 0,1,2,...,k,... , care reprezintă numărul de realizări într-un interval dat de timp sau spațiu ale unui eveniment:

- numărul de intrări pe an într-un spital
- numărul de globule albe de pe frotiu
- numărul de dezintegrări ale unei substanțe radioactive într-un interval de timp T dat

» Poisson?

- O variabilă aleatoare urmează o distribuție Poisson dacă
  - + evenimentul de interes este rar
  - + populația este mare
  - + și evenimentele sunt independente

Distribuția POISSON

»Variabila aleatoare POISSON

»Este caracterizată de parametrul teoretic θ (numărul mediu aşteptat de realizări ale evenimentului în intervalul considerat)

»Simbol: Po(θ)

»Legea de distribuție:

$$X : \begin{pmatrix} k \\ e^{-\theta} \cdot \dfrac{\theta^k}{k!} \end{pmatrix} \qquad \Pr(X = k) = \frac{e^{-\theta} \cdot \theta^k}{k!}$$

Distribuția POISSON

»Speranța matematică: $M(X) = \theta$

»Variația: $V(X) = \theta$

»Rata de mortalitate pentru a anumită patologie virală este de 7 la 1000 de cazuri. Care este probabilitatea ca într-un grup de 400 entități această patologie să determine 5 decese?

»n=400

»p=7/1000=0,007

»$\theta = n \cdot p = 400 \cdot 0{,}007 = 2{,}8$

»e=2,718281828=2,72

Pr(X=5)

$=(2{,}72^{-2{,}8} \cdot 2{,}8^{5}) / (5 \cdot 4 \cdot 3 \cdot 2 \cdot 1)$

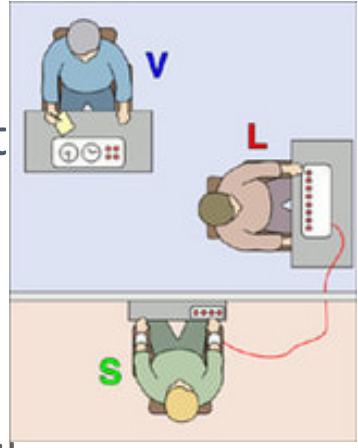
=10,45/120

=0,09

Distribuția POISSON

» Otto Adolf Eichmann (holocaust)

"Could it be that Eichmann and his million accomplices in the Holocaust were just following orders? Could we call them all accomplices?"



Experimentul Milgram:

- » Eşantion de 45 bărbați – participarea la experiment s-a plătit cu 4,50 dolari
- » Generator de şocuri electrice: **30 volți** + 15 volți → max 450 volți (swich-urile au fost etichetate ca “şoc uşor”, “şoc mediu” şi “şoc sever”)
- » S-a evaluat obediența/rezistența la autoritate a individului atunci când ordinal contravine principiilor morale
- » 1 “elev” + 1 “profesor” + cercetătorul
- » Profesorul avea o listă cu perechi de cuvinte – citea primul cuvânt din pereche urmat de o listă de 4 opțiuni posibile – la fiecare greşeală i se aplica “elevului” de către “profesor” un şoc electric a cărui intensitatea creştea odată cu numărul de răspunsuri greşite
- » Dacă “profesorul” dorea să abandoneze, cercetătorul spunea: “Vă rog, continuați” – “E important pentru reușita experinentului să continuați” – “E absolut esențial să continuați” – “Nu aveți altă opțiune; trebuie să continuați”
- » P(electroşoc) = 0.61-0.66

Experimentul Milgram

Sorana D. BOLBOACĂ

- Fiecare persoană din experimentul Milgram poate fi văzut ca un experiment
- Un participant e văzut ca succes dacă refuză să administreze electroșocul sever și eșec dacă admonistrează electroșocul sever
- Probabilitatea de succes P = 0,35
- Din moment ce experimentul are doar două posibilități de răspuns se numește variabilă aleatoare Bernoulli.
- Dorim să includem într-un experiment Milgram prin randomizate 4 persoane. Care este probabilitatea ca exact una din cele 4 persoane să refuze administrarea șocului sever?

Experimentul Milgram

» Dorim să includem într-un experiment Milgram prin randomizate 4 persoane. Care este probabilitatea ca exact una din cele 4 persoane să refuze administrarea șocului sever?

» 4 persoane: Ștefan (A), Liana (B), Mihaela (C), Mihai (D)

» 4 scenarii posibile:

- Scenariul 1: $\frac{0{,}35}{\text{A refuză}} \times \frac{0{,}65}{\text{B aplică}} \times \frac{0{,}65}{\text{C aplică}} \times \frac{0{,}65}{\text{D aplică}} = 0{,}0961$
- Scenariul 2: $\frac{0{,}65}{\text{A aplică}} \times \frac{0{,}35}{\text{B refuză}} \times \frac{0{,}65}{\text{C aplică}} \times \frac{0{,}65}{\text{D aplică}} = 0{,}0961$
- Scenariul 3: $\frac{0{,}65}{\text{A aplică}} \times \frac{0{,}65}{\text{B aplică}} \times \frac{0{,}35}{\text{C refuză}} \times \frac{0{,}65}{\text{D aplică}} = 0{,}0961$
- Scenariul 4: $\frac{0{,}65}{\text{A aplică}} \times \frac{0{,}65}{\text{B aplică}} \times \frac{0{,}65}{\text{C aplică}} \times \frac{0{,}35}{\text{D refuză}} = 0{,}0961$

= 4×0,0961 = 0,3844

Experimentul Milgram

Sorana D. BOLBOACĂ

»Se știe că prevalența consumului de droguri în rândul populației de tineri cu vârsta de 16 ani in anul 2012 a fost de 10%.

- Care este probabilitatea ca exact 6 tineri dintr-un eșantion aleator de 10 persoane să fie consumatori de droguri?
- Care este probabilitatea ca exact 4 tineri dintr-un eșantion aleator de 10 persoane să nu fie consumatori de droguri?

Probleme

»Se știe că prevalența consumului de droguri în rândul populației de tineri cu vârsta de 16 ani in anul 2012 a fost de 10%.

- Care este probabilitatea ca exact 6 tineri dintr-un eșantion aleator de 10 persoane să fie consumatori de droguri?

»p=0,10

»P(K=6) = combin(10,6) × p^6 × $(1-p)^{10-6}$ = 0,000138

- Care este probabilitatea ca exact 4 tineri dintr-un eșantion aleator de 10 persoane să nu fie consumatori de droguri?

»p=1-0,10=0,9

»P(K=4) = combin(10,4) × $0,9^4$ × $0,1^6$ = 0,000138

Probleme