SEMINAR NR. 2

1. Variabile aleatorii. Notații, definiții, concepte de bază

Variabila aleatorie X este o funcție *fixă* şi *deterministă* care alocă (repartizează, distribuie) un număr real X(ξ) fiecărei realizări ξ din spațiul de eşantionare S al unui experiment aleatoriu, *E*.
Spațiul de eşantionare ale varialiei aleatoare X se notează cu S_X.

1.1 Funcția de distribuție

Funcția de distribuție $F_X(x)$ a variabilei aleatoare X este probabilitatea apariției evenimentului $\{ X \le x \}$: $F_X(x) = P[\{X \le x\}]$ pentru $-\infty < x < \infty$.
Proprietățile funcției de distribuție (rezultate din axiomele probabilităților) sunt:

i). $0 \le F_X(x) \le 1$

ii). $\lim\limits_{x \to \infty} F_X(x) = 1$

iii). $\lim\limits_{x \to -\infty} F_X(x) = 0$

iv). $F_X(a) < F_X(b), \quad \forall \ a < b$ (monoton crescătoare)

v). $F_X(b) = \lim\limits_{\substack{h \to 0 \\ h>0}} F_X(b+h) \stackrel{not}{=} F_X(b^+)$ (continuitate la dreapta)

vi). $P[a < X \le b] = F_X(b) - F_X(a)$

vii). $P[X = b] = F_X(b) - F_X(b^-)$, $P[X = b]$=0 dacă X este continuă în *b;*

viii). $P[a \le X \le b] = F_X(b) - F_X(a^-)$

ix). $P[a < X < b] = F_X(b^-) - F_X(a)$

O variabilă aleatoare X este **variabilă aleatoare continuă** dacă funcția sa de distribuție $F_X(x)$ este continuă peste tot şi dacă poate fi scrisă ca o integrală dintr-o funcție pozitivă $f(x)$: $F_X(x) = \int\limits_{-\infty}^{x} f(t)dt$

O variabilă aleatoare X este o **variabilă aleatoare discretă** dacă funcția sa de distribuție $F_X(x)$ este sub formă de scară, continuă la dreapta, având o mulțime măsurabilă (finită sau infinită) de puncte de salt $x_0, x_1, x_2, ...$ aceste puncte reprezentând spațiul de eşantionare S_X.

Funcția de distribuție a unei variabile aleatoare discrete se poate exprima ca o sumă ponderată de funcții treaptă unitate.

$$F_X(x) = \sum_k p_X(x_k) \cdot u(x - x_k)$$

unde: $p_X(x_k) = P[x = x_k]$ este **probabilitatea elementară** de apariție a evenimentului $\{X = x_k\}$.

O variabilă aleatoare X este **variabilă aleatoare mixtă** dacă funcția sa de distribuție $F_X(x)$ are, pe lângă o mulțime de salturi în punctele $x_0, x_1, x_2, ...$, cel puțin un interval pe care evoluează continuu. Astfel de variabile aleatorii se pot genera în *cadrul unor experimente secvențiale dependente,* în 2 trepte: o încercare Bernoulli este urmată de generarea variabilelor aleatorii conform distribuției alese a sub-experimentului precedent.

1.2 Funcția densitate de probabilitate

Funcția de densitate de probabilitate $f_X(x)$ a variabilei aleatoare X, dacă există, este derivata funcției $F_X(x)$: $f_X(x) = \frac{dF_X(x)}{dx}$.

Funcția $f_X(x)$ reprezintă o „densitate de probabilitate" intrucât probabilitatea ca variabila aleatoare X sa apara in vecinătatea $(x, x+dx)$ a unui punct *x,* este data de relația) vezi figura 2.1:

$$p[x < X \le x + dx] \approx f_X(x)dx$$

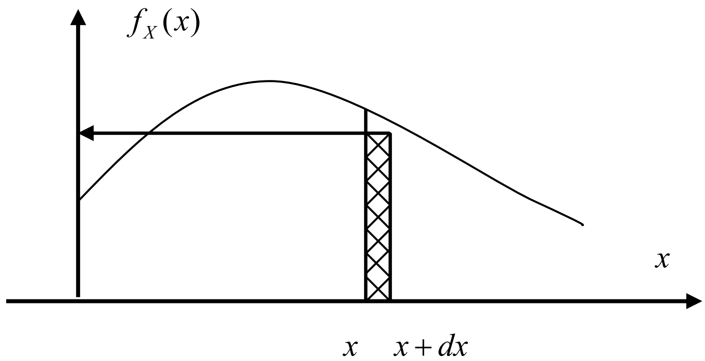

Fig. 2.1 Specificarea probabilitatii unui interval infinitezimal prin intermediul functiei densitate de probabilitate

În cazul variabilei aleatoare discrete, cu funcții de distribuții discontinue în punctele $x_0, x_1, x_2, ...$, funcția de densitate de probabilitate va conține funcții delta corespunzătoare punctelor de discontinuitate: $f_X(x) = \sum_k p_X(x_k)\delta(x - x_k)$, unde: $p_X(x_k) = P[X = x_k]$. Funcția de distribuție corespunzătoare are expresia:

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt = \sum_k p_X(x_k) \int_{-\infty}^{t} \delta(x - x_k)dt = \sum_k p_X(x_k)u(x - x_k)$$

Pentru variabile aleatorii mixte, funcția densitate de probabilitate va conține atât funcții delta, cât şi funcții continue corespunzătoare.

Principalele proprietăţi ale funcţiei densitate de probabilitate sunt:

i). $f_X(x) \geq 0$

ii). $P[a \leq X \leq b] = \int_a^b f_X(x)dx$

iii). $F_x(x) = \int_{-\infty}^{x} f_X(t)dt$

iv). $\int_{-\infty}^{\infty} f_X(x)dx = 1$ (condiţia de normare)

Aplicaţia 1

Timpul de transmisie X al unui mesaj într-un sistem de comutaţie urmează o lege de probabilitate exponenţială de parametrii λ, de forma:

$$p[X > x] = e^{-\lambda x},\ x > 0$$

Să se determine:

i). funcţia de distribuţie şi funcţia de densitate de probabilitate

ii). $p[T < X \leq 2T]$, unde $T = \frac{1}{\lambda}$

Rezolvare:

i). $F_X(x) = P[X \leq x] = 1 - P[X > x]$, deci:

$$F_X(x) = \begin{cases} 0, & pentru\ x < 0 \\ 1 - e^{-\lambda x}, & pentru\ x \geq 0 \end{cases}$$

$F_X(x)$ fiind o funcţie continuă $\Rightarrow$

$$f_X(x) = \frac{dF_X(x)}{dx} = \begin{cases} 0, & pt.\, x < 0 \\ \lambda e^{-\lambda x}, & pt.\, x \geq 0 \end{cases}$$

ii). Rezolvarea prin mai multe abordări:

M1. $P[T < X \leq 2T] = P[T < X] - P[2T < X] = e^{-\lambda\frac{1}{\lambda}} - e^{-\lambda\frac{2}{\lambda}} = e^{-1} - e^{-2} \approx 0,233$

M2. $P[T < X \leq 2T] = F_X(2T) - F_X(T) = 1 - e^{-2} - 1 + e^{-1} = e^{-1} - e^{-2} \approx 0,233$

M3. $P[T < X \leq 2T] = \int_T^{2T} f_X(x)dx = e^{-\lambda x}\,|_T^{2T} = e^{-1} - e^{-2} \approx 0,233$

Aplicaţia 2

Timpul de aşteptare W al unui client într-un sistem este zero dacă găseşte sistemul liber şi distribuit exponenţial dacă găseşte sistemul ocupat. Probabilităţile de a găsi sistemul liber sau ocupat sunt p şi respectiv $1-p$. Găsiţi:

i). Funcţia de distribuţie a variabilei aleatorie W

ii). Funcţia densitate de probabilitate

Rezolvare:

$F_W(x) = P[W \le x] = P[W \le x \mid sistemul\_liber]p + P[W \le x \mid sistemul\_este\_ocupat]\cdot(1-p)$
(s-a utilizat teorema probabilității totale)

Deoarece :

$$P[W \le x \mid sistemul\_e\_liber] = \begin{cases} 0, pentru\_x < 0 \ (nu\_poate\_exista\_timp\_de\_asteptare\_negativ) \\ 1, pentru\_x \ge 0 \ (W = 0 \le x) \end{cases}$$

şi

$$p[W \le x \mid sistemul\_e\_ocupat] = 1 - e^{-\lambda x}$$

rezultă:

$$F_W(x) = \begin{cases} 0, \ pentru\_x < 0 \\ p + (1-p)(1-e^{-\lambda x}), \ pentru\_x \ge 0 \end{cases}$$

Funcția de distribuție condiționată de un eveniment A, a variabilei aleatorie X este:

$$F_X\left(x|A\right) = \frac{P[\{X \le x\} \,\&\, A]}{P[A]}, daca\_P[A] > 0$$

Funcția de densitate de probabilitate condiționată de un eveniment A, a variabilei aleatorie X este (dacă derivatele există !) :

$$f_X\left(x|A\right) = \frac{dF_X\left(x|A\right)}{dx}$$

Aplicația 3

Timpul de servire X, al unui client urmează o funcție de distribuție continuă $F_X(x)$. Determinați funcția de distribuție condiționată şi funcția de densitate de probabilitate condiționată relativ la evenimentul $A=\{$clientul este încă în servire la momentul t$\}$. Particularizați în cazul unei distribuții exponențiale a lui X.

Rezolvare:
Evenimentul A={clientul este încă în servire la momentul t} = $\{X > t\}$

$$F_X\left(x|A\right) = P[X \le x \mid X > t] = \frac{P[\{X \le x\} \,\&\, \{X > t\}]}{P[\{X > t\}]} = \begin{cases} 0, \quad pentru\_x < t \\ \dfrac{P[t < X \le x]}{1 - P[X \le t]}, \ pentru\_x \ge t \end{cases}$$

$F_X\left(x|A\right) = 0$ pentru $x < t$ deoarece clientul nu mai poate fi în servire la momentul t dacă până la momentul x s-a încheiat servirea lui. Înlocuind probabilitățile cu expresii ce conțin funcția de distribuție, obținem:

$$F_X\left(x|A\right)=\begin{cases}0, & pentru\_x<t\\ \dfrac{F_X(x)-F_X(t)}{1-F_X(t)}, & pentru\_x\geq t\end{cases}$$

$$f_X\left(x|A\right)=\frac{dF_X\left(x|A\right)}{dx}=\begin{cases}0, & pentru\_x<t\\ \dfrac{f_X(x)}{1-F_X(t)}, & pentru\_x\geq t\end{cases}$$

Aplicația 4

Să se demonstreze că pentru n suficient de mare şi pentru p suficient de mic, Distribuția Poisson aproximează distribuția lui nomială.

Rezolvare

Notând $\alpha=n\cdot p$ (media numărului de succese). Probabilitățile generate de distribuția binominlă au expresia:

$$p_k=\binom{n}{k}p^k(1-p)^{n-k}, k=1,2,.....,n;\quad \binom{n}{k}=C_n^k \text{, cu } C_n^k=\frac{n!}{k!(n-k)!}$$

Relația de recurență care leagă aceste probabilități este:

$$\frac{p_{k+1}}{p_k}=\frac{\binom{n}{k+1}p^{k+1}(1-p)^{n-k-1}}{\binom{n}{k}p^k(1-p)^{n-k}}=....=\frac{(1-\frac{k}{n})\cdot\alpha}{(k+1)\cdot(1-\frac{\alpha}{n})}$$

Când $n\to\infty$:

$$\left.\frac{p_{k+1}}{p_k}\right|_{n\to\infty}=\frac{\alpha}{k+1}\text{, deci } p_{k+1}=\frac{\alpha}{k+1}p_k,\quad k=0,1,2......$$

Pentru a putea folosi aceasta relație de recurentă trebuie să determinăm una din probabilitățile p_k.

In cazul de față: $p_o=(1-p)^n=(1-\frac{\alpha}{n})^n\underset{n\to\infty}{=}e^{-\alpha}$

Cu ajutorul relației de secvență obținem:

$p_1=\frac{\alpha}{1}e^{-\alpha}$, $p_2=\frac{\alpha^2}{2!}e^{-\alpha}$,.., $p_k=\frac{\alpha^k}{k!}e^{-\alpha}$ …… ceea ce corespunde unei distribuții Poisson.

Interpretare: O secvență de n încercări Bernoulli se poate desfăşura în spațiu sau timp. În ultimul caz ea poate fi reprezentata conform figurii de mai jos:

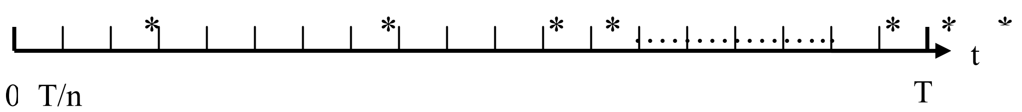

Fig.2.2 desfasurarea in timp a unei secvente de *n* incercari Bernoulli

unde

- [0,T] este intervalul in care au loc cele n încercări (durata dintre două încercări succesive se presupune constantă)

- * este un simbol ce marchează apariția succesului încercării

Distribuția binomială ce descrie acest experiment oferă media numărului de succese $n \cdot p$ în intervaul *T*. În acest caz, raportul $\frac{n \cdot p}{T}$ poate fi interpretat ca o rată medie de apariție (sosire) a proceselor:

$$\lambda = \frac{n \cdot p}{T} = \frac{\alpha}{T}$$

Variabilele aleatorii *N*, ce reprezintă numărul de succese obținute în *n* încercări Bernoulli desfăşurate în intervalul [0,*T*] urmează o distribuție binomială de parametrii *n* şi *p* , atunci când $n \to \infty$ şi $p \to 0$, N devine o variabilă aleatorie Poisson de parametru $\lambda = \frac{\alpha}{T}$.

Aplicație 5

Fie *X* numărul de surse active dintr-un număr de *N* surse independente ce generează blocuri informaționale. Sistemul de transmisii acceptă maximul *M* transferuri simultane de blocuri informaționale pe unitatea de timp. Dacă limita e depăşită, un număr $Y = X - M$ de blocuri informaționale (alese aleatoriu) sunt eliminate. Să se determine modelul probabilistic ce descrie variabila aleatorie *Y*.

Rezolvare:

$$S_y = \{0,1,2,....., N - M\}$$

$$\begin{cases} P[Y=0] = P[\{X=0\} sau \{X=1\},.....sau\{X=M\}] = \sum_{j=0}^{M} p_j \\ P[Y=k] = P[X = M+k] = p_{M+k}, \quad 0 < k \le N-M \end{cases}$$

unde: $p_j = p[X = j] = ... = \binom{N}{j} p^j (1-p)^{N-j}$

2. Distribuțiile uzuale de variabile aleatorii

2. 1 Distribuția Bernoulli

Se presupune că: variabila aleatorie *X* este **funcție indicator** (I_A) pentru un eveniment *A*:

$$X = \begin{cases} 1\_daca\_evenimentul\_A\_apare \\ 0\_in\_caz\_contrar \end{cases}$$

Parametrii: *p*-probabilitatea de apariție a evenimentului *A*, ($0 \le p \le 1$)

Probabilități elementare: $p_1 = P[X=1] = p;\ \ p_0 = p[X=0] = 1-p;$

Media: $E[X] = p$

Varianța: $VAR[X] = p \cdot (1-p)$
Aplicatie: modelarea mecanismului fundamental de generare a aleatorului

2.2 Distribuția binomială

Se presupune că variabila aleatorie *X* reprezintă numărul succeselor unui eveniment înregistrate în *n* subexperimente (repetări) independente (este o sumă de *n* variabile aleatorii independente, identic distribuite după *legea Bernoulli*). Probabilitatea succesului în cadrul unei repetări este *p*:

Spațiile realizărilor: $S_X = \{0,1,2,....n\}$
Parametrii: *n* (*n*>0), *p* ($0 \le p \le 1$).
Probabilități elementare: $p_X(k) = p[X = k] = \binom{n}{k} p^k (1-p)^{n-k}, k = \overline{0,n}$

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$ (combinări de n luate câte k)

Media : $X[X] = n \cdot p$
Varianța: $VAR[X] = n \cdot p \cdot (1-p)$
Aplicații: numărul de biți eronați într-un bloc de *n* biți transmişi printr-un canal fără zgomot, numărul de nume active dintr-un număr de *n* surse în servire, etc. .

2.3 Distribuția geometrică

Versiunea I :

Variabila aleatorie *X* reprezintă numărul eşecurilor unui eveniment în inregistrat pe parcursul unei secvențe de subexperimente independente, tip Bernoulli, până la apariția primului succes. Probabilitatea succesului în cazul unei repetări este *p*.

Spațiile realizărilor: $S_X = \{0,1,2,....\}$
Parametrii: *p* ($0 \le p \le 1$)
Probabilități elementare: $p_X(k) = P[x = k] = p(1-p)^k, k = 0,1,2,\ldots$
Media : E[x]=$E[X] = \frac{1-p}{p}$
Varianța: $VAR[X] = \frac{1-p}{p^2}$

Versiunea II:

Se presupune ca variabila aleatoare X' este numărul de repetări până ce apare *primul succes* într-o secvență de subexperimente independente tip Bernoulli.

Spațiile realizărilor: $S_{X'} = \{1,2,....\}$
Probabilități elementare: $p_{X'}(k) = p \cdot (1-p)^{k-1},\ k = 1,2,....$

Media : $E[X'] = \frac{1}{p^2}$

Varianța: $VAR[X'] = \frac{1-p}{p^2}$

Observații:

- $X' = X + 1$
- Este singura variabilă aleatoare discretă ce deține proprietatea de lipsă de memorie

Aplicații: numărul de apeluri repetate până la ocuparea unei surse; numărul de retransmisii într-un canal cu zgomot până la prima recepționare corectă ; etc.

2.4 Distribuția Poisson

Se presupune că :

i). Variabila aleatorie X este numărul de apariții ale unei anumite realizări pe parcursul unui interval ΔT de timp

ii). Momentele aparițiilor sunt distincte ;

iii). Orice interval finit conține un număr finit de apariții;

iv). Orice interval infinit conține un număr infinit de apariții;

v). Evenimentele nu apar la momente predeterminate ;

vi). Numărul aparițiilor dintr-un interval este independent de numărul aparițiilor dintr-un alt interval disjunct.

Spațiile realizărilor: $S_X = \{0, 1, 2, ...\}$

Parametrii: $\alpha > 0$ (numărul mediu de evenimente ce apar în intervalul dat)

Probabilități elementare: $p_X(k) = \frac{\alpha^k}{k!} e^{-\alpha}$

Media : $E[X] = \alpha$

Varianța: $VAR[X] = \alpha$

Observatii:

- aproximeaza distributia binomiala pentru *p* “mic” si *n* “mare”, considerand: $p \cdot n = \alpha$
- timpii între două apariții succesive sunt variabile aleatorii independente, identic distribuite după o lege exponențială de parametru: $\lambda = \frac{\alpha}{\Delta T}$.

Aplicații: - numărul mesajelor sosite într-un sistem de comunicații, numărul cererilor de conexiuni într-o rețea de comunicație, numărul defectelor dintr-un dispozitiv electronic, etc.

2.5 Distribuția uniformă discretă

Se presupune că variabila aleatorie X ia valori într-un interval finit $\{1,2,....n\}$ de numere naturale. Toate valorile sunt echiprobabile.

Spațiile realizărilor: $S_X = \{0,1,2,....n\}$

Parametrii: n

Probabilități elementare: $p_X(k) = P[X = k] = \frac{1}{n}, \quad k = 1,2,.....n$

Media : $E[X] = \frac{n+1}{2}$

Varianța: $VAR[X] = \frac{n^2 - 1}{12}$

Aplicații: Evaluarea raportului semnal-zgomot introdus de procesul de cuantizare.

2.6 Distribuția uniformă continuă

Se presupune că variabila aleatorie X ia valori reale în intervalul $[a,b]$. Probabilitatea ca variabila aleatoare să ia o valoare într-un subinterval este proporțională cu lungimea intervalului.

Spațiile realizărilor: $S_X = [a,b]$

Parametrii: $a \; si \; b$ cu $(a < b)$

Funcția densitatea de probabilitate: $f_X(n) = \begin{cases} \frac{1}{b-a}, & a \le x \le b \\ 0, & in\_rest \end{cases}$

Funcția de distribuție: $F_X(x) = \begin{cases} 0, & pentru \; x < a \\ \frac{x-a}{b-a}, & pentru \quad a \le x \le b \\ 1, & pentru \quad x > b \end{cases}$

Media : $E[X] = \frac{a+b}{2}$

Variația: $VAR[X] = \frac{(b-a)^2}{12}$

Aplicații: Valori distribuite uniform în $[0,1]$ sunt utilizate pentru a genera numere ce corespund altor distribuții.

2.7 Distribuția exponențială

Nu se fac presupuneri

Spațiul realizărilor: $[0,\infty)$

Parametrii: $\lambda > 0$

Funcția densitate de probabilitate: $f_X(x) = \lambda \cdot e^{-\lambda x}$

Funcția de distribuție: $F_X(x) = 1 - e^{-\lambda x}$

Media : $E[X] = \frac{1}{\lambda}$

Varianța: $VAR[X] = \frac{1}{\lambda^2}$

Observații:

- reprezintă forma limită a distribuției geometrice;
- este unica distribuție ce deține proprietatea de *lipsă memorie*.

Aplicații: lungimea mesajelor şi timpul între două apariții succesive de evenimente (cereri de serviciu) în sistemele de comunicații; timpul de funcționare (fiabilitatea sistemelor şi componentelor)

2.8 Distribuția Gamma

Nu se fac presupuneri

Spațiile realizărilor: $(0, \infty)$

Parametrii: $\alpha > 0 \quad \lambda > 0$

Densitatea de probabilitate: $f_X(x) = \frac{\lambda \cdot (\lambda \cdot x)^{\alpha-1} \cdot e^{-\lambda x}}{\Gamma(\alpha)}$

unde: $\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy$ cu proprietățile

i). $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

ii). $\Gamma(\alpha+1) = \alpha \cdot \Gamma(\alpha)$

iii). $\Gamma(\alpha) = (\alpha - 1)!$, pt. $\alpha \in N$

Media : $E[X] = \frac{\alpha}{\lambda}$

Varianța: $VAR[X] = \frac{\alpha}{\lambda^2}$

Aplicații: Distributia gamma, prin intermediul parametrilor α şi λ poate genera o varietate de curbe prin care se pot modela diverse experimente.

O serie de distributii sunt cazuri particulare ale distribuției gama.

Astfel:

- pentru $\alpha = 1$ se obtine distributia exponentiala;
- pentru $\lambda = 1/2$ si $\alpha = k/2$ cu $k \in N^*$ se obtine distribuția χ^2;
- pentru $\alpha = m$, $m \in N$ se obtine distribuția *m*- Erlang.

2.9 Distribuția *m*-ERLANG

Spațiile realizărilor: $S_X = (0, \infty)$

Parametrii: $m \in N^*$ şi $\lambda > 0$

Funcția densitatede probabilitate: $f_X(x) = \frac{\lambda e^{-\lambda x} (\lambda \cdot x)^{m-1}}{(m-1)!}$.

Funcția de distribuție: $F_X(x) = 1 - \sum_{k=0}^{m-1} \frac{(\lambda \cdot x)^k}{k!} e^{-\lambda x}$

Media : $E[X] = \frac{m}{\lambda}$

Varianța: $VAR[X] = \frac{m}{\lambda^2}$

Aplicații: O variabilă aleatoare *m*-ERLANG se obține adunând *m* variabile aleatorii exponențiale independente, identic distribuite, de parametru λ. De aceea, ***ea este un model general a timpilor de aşteptare în sistemele cu aşteptare***, a timpilor de viață în studiile de fiabilitate, etc..