

CURSUL – I

PROBABILITATI

DISTRIBUTII

VARIABLE ALEATOARE

ELEMENTE DE TEORIA PROBABILITĂȚILOR

CÂMPURI DE PROBABILITATE

Teoria matematică a probabilităților pornește de la faptul că fiecărui rezultat posibil al unui experiment aleator, rezultat pe care îl vom denumi eveniment, i se asociază o valoare numerică, numită “probabilitatea” evenimentului respectiv. Această valoare este o caracteristică obiectivă a evenimentului în condițiile experimentului dat.

Să efectuăm, de exemplu, un experiment de m ori. Dacă în cele m experiențe un eveniment A s-a produs de k ori, atunci $0 \leq k \leq m$, de unde rezultă pentru frecvența relativă:

$$0 \leq \frac{k}{m} \leq 1,$$

adică frecvența relativă a unui eveniment este întotdeauna un număr cuprins între 0 și 1. Ținând cont că frecvența relativă oscilează în jurul probabilității evenimentului considerat și că probabilitate este acea caracteristică a evenimentului care ne indică în ce proporții se produce evenimentul în cazul repetării experimentului de un număr foarte mare de ori, rezultă că și probabilitatea este tot un număr între 0 și 1. Din definiția probabilității ca generalizare a conceptului de frecvență relativă, rezultă că probabilitatea unui eveniment imposibil este 0, iar probabilitatea unui eveniment sigur este 1.

Evenimentele pot fi simple, în sensul că nu se pot descompune mai departe, sau compuse din alte evenimente ce se petrec simultan. În acest context putem considera două operații între evenimente.

Scriem $A \cap B$ și înțelegem prin aceasta un eveniment care constă în producerea evenimentelor A și B , simultan. Scriem $A \cup B$ pentru cazul când se produce cel puțin unul din cele două evenimente.

Fiind date două rezultate A și B ale unui experiment efectuat de n ori, să presupunem că A s-a obținut de k_1 ori și B de k_2 ori. Evenimentul $A \cup B$, deci obținerea unui eveniment din cele două rezultate, s-a obținut ca atare, de $\frac{k_1+k_2}{n} = \frac{k_1}{n} + \frac{k_2}{n}$ ori, ceea ce sugerează o regulă de tipul

$$\text{Probabilitate}(A \cup B) = \text{Probabilitate}(A) + \text{Probabilitate}(B)$$

În cele ce urmează vom introduce o prezentare axiomatică a conceptului de probabilitate, după Kolmogorov¹.

Corp borelian

Definiție:

Fie E o mulțime și \mathcal{K} o familie nevidă de părți ale lui E , $\mathcal{K} \subset \mathcal{P}(E)$ cu proprietățile:

1. $A \in \mathcal{K} \Rightarrow CA \in \mathcal{K}$
2. $(A_i)_{i \in N} \subset \mathcal{K} \Rightarrow \prod_1^\infty A_i \in \mathcal{K}$
3. $E \in \mathcal{K}$

Deci, este închisă la operațiile de complementare și reuniune.

Se spune, în acest caz, că familia \mathcal{K} , împreună cu operațiile menționate, formează un corp borelian. Denumirea de borelian vine de la matematicianul Emil Borel, unul dintre fondatorii teoriei probabilităților.

Consecință:

Un corp borelian este o familie închisă față de operațiunea de intersecție, indiferent de numărul elementelor sale pe care le intersectăm:

$$(A_i)_{i \in N} \subset \mathcal{K} \Rightarrow \prod A_i \in \mathcal{K}$$

Demonstrația se face imediat folosind faptul că $\prod_i A_i = C\left(\bigcup_i A_i^c\right)$ și proprietățile 1 și 2.

Propoziție:

Fiind dată o familie de corpuri boreliene $(\mathcal{K}_i)_{i \in I}$, intersecția lor este tot un corp borelian.

Demonstratia se face imediat, folosind proprietățile corpului borelian și ale operațiilor de intersecție, reuniune și complementare.

Definiție:

Fie H o familie oarecare de părți ale unei mulțimi E . H poate fi completată la un corp borelian, numit *corpul generat de H* , dacă i se adaugă E și toate mulțimile ce se formează prin reuniune, intersecție și complementare pornind de la elementele $H \in H$.

Dacă luăm pe dreapta, mulțimea intervalelor deschise de forma $(-\infty, a)$, $a \in \mathbb{R}$, corpul borelian generat se numește simplu "borelianul pe dreapta" și constituie baza teoriei probabilităților, așa cum va fi ea abordată în prezenta lucrare. Deoarece orice interval închis se poate obține prin operațiile meționate din intervale deschise și invers,

orice interval deschis poate fi generat pornind de la intervale închise, borelianul pe dreapta este în același timp generat de mulțimea intervalelor închise.

Într-adevăr, se poate scrie:

$$[a, b] = \bigcap_{n=1}^{\infty} \left(a - \frac{1}{n}, b + \frac{1}{n} \right) \text{ și } (a, b) = \bigcap_{n=1}^{\infty} \left[a + \frac{1}{n}, b - \frac{1}{n} \right]$$

Definiție:

O familie $(A_i)_{i \in I}$ se numește *desfacere* a lui E dacă:

1. I este cel mult numărabilă;
2. $\forall i, \forall j \Rightarrow A_i \cap A_j = \emptyset$
3. $\cup A_i = E$

Spații măsurabile

Definiție

O mulțime E împreună cu un corp borelian \mathcal{K} formează un *spațiu măsurabil* (E, \mathcal{K}) . Elementele lui \mathcal{K} se numesc *mulțimi măsurabile*.

Definiție

Fiind date (E, \mathcal{K}) și (F, \mathcal{L}) spații măsurabile, o funcție $f: (E, \mathcal{K}) \rightarrow (F, \mathcal{L})$ se numește *funcție măsurabilă* dacă îndeplinește condiția:

$$\forall A, A \in \mathcal{L} \Rightarrow f^{-1}(A) \in \mathcal{K} \text{ sau, altfel spus: } f^{-1}(\mathcal{L}) \subset \mathcal{K}$$

Proprietăți

- a) Dacă f și g sunt măsurabile, atunci f o g, f + g și f * g sunt măsurabile.
- b) Dacă f este continuă, atunci f este borelian măsurabilă.

Observație

Se poate face un paralelism între spațiile topologice și spațiile măsurabile, între funcțiile continue și funcțiile măsurabile. Astfel, o funcție este continuă dacă preimaginea oricărei mulțimi deschise este o mulțime deschisă iar măsurabilă este atunci când preimaginea oricărei mulțimi măsurabile este măsurabilă. Deasemenea, dacă f și g sunt două funcții continue, atunci f + g și f * g sunt continue.

Definiție

Se numește *măsură* orice funcție pozitivă definită pe corpul mulțimilor măsurabile, $\mu: K \rightarrow \mathbb{R}_+$, “aditivă” pe orice familie $(A_i)_{i \in I}$ numărabilă de mulțimi măsurabile

disjuncte: $\forall n, \forall m, A_n \cap A_m = \Phi \Rightarrow \mu\left(\bigcup_{i=1}^{\infty} A_n\right) = \sum_{i=1}^{\infty} \mu(A_n)$

Consecințe

a) $\mu(\Phi) = 0$

Într-adevăr, dacă luăm $A_1 = A, A_2 = \Phi \Rightarrow \mu(\Phi) = \mu(\Phi \cup \Phi) = 2\mu(\Phi) \Rightarrow \mu(\Phi) = 0$

b) Fie un șir de mulțimi $A_1 \subseteq A_2 \subseteq \dots$ și fie $A = \bigcup A_n$, atunci $\mu(A_n) \rightarrow \mu(A)$

Demonstrație:

Fie $B_n = A_{n+1} \setminus A_n$. Mulțimile B_n sunt disjuncte și $A_n = B_1 \cup B_2 \cup \dots \cup B_n$.

Din aditivitatea lui μ rezultă $\mu(A_n) = \mu\left(\bigcup_{i=1}^n B_i\right) = \sum_{i=1}^n \mu(B_i) = s_n$

$s_n \rightarrow s = \mu\left(\bigcup_{i=1}^{\infty} B_i\right) = \mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \mu(A)$

$A = \bigcup A_n$ și $\mu(A_i) < \infty \Rightarrow \mu(A_n) < \mu(A)$

Altfel, $A_n = \{n, n+1, \dots\}$, $\bigcup A_n = \Phi$ dar $\mu(A_n) = \infty$

Exemple

a) Fie μ definită după cum urmează:

- $\mu(A) = \infty$ dacă A este infinită și
- $\mu(A) =$ numărul elementelor din A , dacă A este finită.

Această măsură se numește în mod natural “măsura de numărare”.

b) Fie un punct exterior $x_0 \in E$ fixat. Definim:

- $\mu_{x_0}(A) = 1$ dacă $x_0 \in A$ și
- $\mu_{x_0}(A) = 0$ dacă $x_0 \notin A$

Măsura este utilizată în mecanica cuantică și se numește “măsura lui Dirac”.

Probabilitate Vom defini probabilitatea ca o măsură particulară.

Definiție:

Fiind dat un spațiu măsurabil (E, K) . O funcție $P: K \rightarrow [0,1]$ cu proprietățile:

a) P – măsură și

$$b) P(E)=1$$

se numește probabilitate.

Deci, probabilitatea ar fi o măsură “normată”.

Proprietăți:

Pe baza proprietăților măsurii și a faptului că $P(E)=1$, se pot demonstra cu ușurință următoarele proprietăți:

1. $A \supset B \Rightarrow P(A/B) = P(A) - P(B)$
2. $(\forall n), A_n \subset A_{n+1} \Rightarrow P(\bigcup A_n) = \lim_{n \rightarrow \infty} P(A_n)$
3. $(\forall n), A_n \supset A_{n+1} \Rightarrow P(\bigcap A_n) = \lim_{n \rightarrow \infty} P(A_n)$
4. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
5. $P(\bigcup A_n) \leq \sum P(A_n)$, numită subaditivitate numărabilă
6. $P(\Phi) = 0$
7. $P(\bar{A}) = 1 - P(A)$

În contextul teoriei probabilităților, mulțimile măsurabile devin *evenimente*, “spațiul măsurabil” devine *câmp de evenimente*, iar E devine *evenimentul total*.

Definiție:

Un câmp de evenimente (E, K) înzestrat cu probabilitatea P, se numește *câmp de probabilitate*.

Definiție:

Un eveniment care nu mai poate fi inclus în alt eveniment

$$A \in K, \forall B \in K, A \subset B \text{ sau } A \cap B = \Phi$$

se numește *eveniment elementar* sau atom.

Observații

Prezentarea axiomelor teoriei probabilităților în contextul mai larg al teoriei măsurii, dincolo de formalismul simplu și riguros, oferă și avantajul unor interpretări “fenomenologice” și “picturale” pentru unele formule. Astfel, dacă probabilitatea este o măsură, la fel ca aria pentru figurile plane, formula:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

se poate citi ca:

$$\text{aria}(A \cup B) = \text{aria}(A) + \text{aria}(B) - \text{aria}(A \cap B)$$

ceea ce pare ca evident.

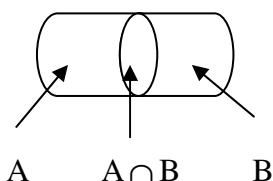


Fig. 1.

Definiția clasică elementară a probabilității derivă în mod natural din noțiunea de frecvență, despre care am vorbit mai sus.

Dacă un eveniment A se poate realiza în m feluri diferite dintr-un număr total n de evoluții posibile $(e_j)_{j=1, \dots, n}$, egal probabile, atunci :

a) $P(e_j) = \frac{1}{n}$ și

b) $P(A) = \frac{m}{n}$

Exemplu

Exemplul clasic de câmp de probabilitate finit îl constituie evenimentele ce pot apărea atunci când, dintr-o urnă în care se află bile albe și negre se extrag n bile. Dacă proporția bilor albe în urnă este p, și deci a celor negre este $q = 1 - p$, probabilitatea evenimentului A, ca din n bile extrase, k să fie albe, conform definiției clasice definite mai sus, se calculează imediat și este:

$$P(A) = C_n^k p^k q^{n-k}$$

De exemplu, evenimentul ca din trei bile extrase, două să fie albe - a - și una să fie neagră - n- se poate descompune în felul următor :

$$A = (a a n) \cup (a n a) \cup (n a a)$$

și

$$P(A) = P(a a n) + P(a n a) + P(n a a) = p^2q + p^2q + p^2q = 3 p^2q = C_3^2 p^2q^{3-2}$$

Probabilitate condiționată

Fie B un eveniment a cărei probabilitate este diferită de 0. Probabilitatea unui eveniment A, reprezintă proporția în care ne așteptăm să se realizeze A în cadrul tuturor evenimentelor câmpului de probabilitate la care aparține A

Probabilitatea lui A se mai poate analiza însă și în contextul în care știm că s-a produs anterior evenimentul B. Probabilitatea evenimentului A condiționată de B se notează, în acest caz, cu: $P(A/B)$ sau $P_B(A)$.

Dacă s-a constatat experimental o frecvență de apariție k_A și, respectiv k_B , pentru A și B, *frecvența relativă de apariție a lui A*, când deja a apărut B, va fi:

$$\frac{k_{AB}}{k_B} = \frac{\frac{k_{AB}}{n}}{\frac{k_B}{n}} \cong \frac{P(A \mid B)}{P(B)}$$

În acest context apare naturală definiția probabilității evenimentului A, condiționată de B, prin formula:

$$P_B(A) = \frac{P(A \mid B)}{P(B)}$$

Un caz special îl constituie acela în care probabilitatea de apariție a evenimentului A este aceeași, indiferent dacă s-a produs sau nu evenimentul B:

$$P(A) = P_B(A)$$

Spunem, în acest caz, că evenimentele A și B sunt *evenimente independente*.

Observăm că, rescriind formula anterioară

$$P_B(A) = \frac{P(A \mid B)}{P(B)} \Rightarrow P(A \mid B) = P_B(A) * P(B) = P(A) * P(B)$$

se poate lua ca definiție că *două evenimente sunt independente* atunci când:

$$P(A \mid B) = P(A) * P(B)$$

Formula probabilității cauzelor (Bayes)

Fie A_1, A_2, \dots, A_n o desfacere a lui E pe care, în contextul teoriei probabilităților, o numim *sistem complet de evenimente*. Ea reprezintă în același timp o desfacere pentru E cât și pentru orice eveniment $X \subset E$.

$$E = \bigcup A_j$$

$$X = \bigcup (A_i \mid X)$$

Dat fiind că evenimentele $A_i \mid X$ sunt disjuncte, avem $P(X) = \sum P(A_i \mid X)$.

Să presupunem că $\forall i, P(A_i) \neq 0$. În aceste condiții avem următoarea teoremă:

Teorema probabilității cauzelor

Probabilitatea producerii oricărui eveniment X , este egală cu suma probabilităților de producere a lui X , condiționate de evenimentele complete ale sistemului $(A_i)_{i=1, n}$ și

$$P_X(A_j) = \frac{P(A_j)P_{A_j}(X)}{\sum P(A_i)P_{A_i}(X)}$$

Demonstrație:

$$\text{Din definiție avem } P_X(A_j) = \frac{P(X \mid A_j)}{P(X)}$$

$$\text{deci, } P_X(A_j) = \frac{P(X \mid A_j)}{\sum_i P(A_i \mid X)} = \frac{P(X \mid A_j) \frac{P(A_j)}{P(A_j)}}{\sum_i P(A_i \mid X) \frac{P(A_i)}{P(A_i)}} = \frac{P(A_j)P_{A_j}(X)}{\sum P(A_i)P_{A_i}(X)}$$

$P_X(A_j)$ poate fi interpretat ca fiind probabilitatea ca X să aibă cauza A_j . În acest caz, formula calculează probabilitatea lui X în funcție de probabilitățile cauzelor care ar fi putut determina evenimentul X . Probabilitățile $P(A_k)$ se numesc *apriorice*, pentru că ele se cunosc înainte de eveniment. Probabilitățile $P_X(A_j)$ sunt probabilitățile aceluiași cauze, dar după ce s-a întâmplat evenimentul X , și se numesc din acest motiv, probabilități *aposteriorice*.

Exemplu, când un pacient intoxicat este adus la urgență el prezintă anumite simptome și medicul, folosind experiența sa, rezultatele determinărilor în sânge și un sistem computerizat elaborează o listă cu probabilitățile ca intoxicația să se fi făcut cu o anumită substanță.

În fizica statistică parametrii termodinamici sau cuantici ai unui sistem rezultă din însumarea unui număr foarte mare de evenimente. Probabilitatea de trecere de la o stare inițială la o stare finală este dată de suma probabilităților de trecere pe anumite căi A_i ponderate fiecare cu probabilitatea, sau altfel spus ponderea lor, $p(A_i)$. Deoarece numărul căilor poate fi de puterea continuului, în locul sumelor apar integrale.

Sau, dacă s-ar produce o crimă, aposteriori, ne punem problema ierarhizării suspiciunilor privind potențialii criminali.

Problema nu este de loc “teoretică” dacă suntem de exemplu o societate de asigurări sau dacă testul este un test de malignitate.

Bayer a fost un episcop care s-a preocupat de cauzele evenimentelor din lumea aceasta și legătura lor cu cauza finală – Dumnezeu.

Formula probabilității cauzelor ne arată cum se transformă probabilitățile apriorice în probabilități aposteriorice, după apariția evenimentului X.

De exemplu, știind că un medicament se absoarbe în, și se elimină din sânge pe mai mult căi, cu diferite probabilități date de considerente fizico-chimice și fiziologice, în funcție de rezultatul unor determinări a concentrației ale acestora în sângele unui pacient, ne putem pune problema stabilirii ponderilor efective ale acestor căi, în scopul “individualizării” tratamentului.

Observație:

Putem deasemenea să considerăm cazul particular al desfacerii evenimentului total în două evenimente A și complementul său CA.

Formula lui Bayes devine în acest caz:

$$P_X(A) = \frac{P_A(X)P(A)}{P_A(X)P(A) + P_{CA}(X)P(CA)}$$

Aplicație:

Dacă, de exemplu, P(B) este proporția (probabilitatea) unei boli în populație și cunoscând proporția în care un test diagnostic este pozitiv la bolnavi – $P_B(T)$ – și la sănătoși – $P_{NB}(T)$ – putem calcula probabilitatea ca un pacient la care rezultatul testului este pozitiv să fie bolnav:

$$P_+(B) = \frac{P_B(T)P(B)}{P_B(T)P(B) + P_{NB}(T)P(NB)}$$

unde:

$P_B(T)$ este probabilitatea ca un bolnav să fie catalogat pozitiv de către test și se numește “sensibilitatea” testului.

$P_{NB}(T)$ este probabilitatea ca un sănătos să fie catalogat negativ de către test și se numește “specificitatea” testului.

Problema devine teribil de importantă dacă, de exemplu, este vorba de un test de depistare a cancerului.

VARIABLE ALEATOARE

Definiții:

a) Se numește *variabilă aleatoare* (întâmplătoare sau statistică) o funcție reală f definită pe mulțimea \mathcal{K} a evenimentelor, cu proprietatea că, oricare ar fi numărul real a , mulțimea $x \in \mathcal{K}$ pentru care $f(x) \leq a$ este un eveniment din \mathcal{K} .

În termeni de teoria măsurii, o variabilă aleatoare este o funcție $f : (E, \mathcal{K}, P) \rightarrow (\mathbb{R}, \mathcal{B})$, măsurabilă.

Practic vorbind avem definită probabilitatea ca variabila să aibă valori mai mici decât orice număr dat a .

b) O variabilă aleatoare se numește *variabilă aleatoare simplă* dacă ia un număr finit de valori: $f : E \rightarrow \mathbb{R}$, $f(E)$ finită și $P(f(x) = x_i) = P(f^{-1}(x_i)) = p_i$

c) Vom lucra, în cele ce urmează, ca regulă, cu *variabile aleatoare independente*, adică variabile ce iau valori independente una de cealaltă:

$$P((f(x) = x_i) \cap (g(y) = y_j)) = P(f(x) = x_i) * P(g(y) = y_j), \forall x_i, y_j$$

Observație:

Se poate verifica ușor că variabilele aleatoare formează o algebră, adică suma, și produsul a două variabile aleatoare este tot o variabilă aleatoare; mai mult compunerea a două variabile aleatoare este tot o variabilă aleatoare.

Trebuie în acest context să fim atenți la independența sau nonindependența variabilelor aleatoare implicate în operație.

De exemplu putem citi $X+X$ unde X este o variabilă aleatoare în două feluri. Putem, de exemplu, să considerăm un experiment repetat de două ori rezultatele fiind independente

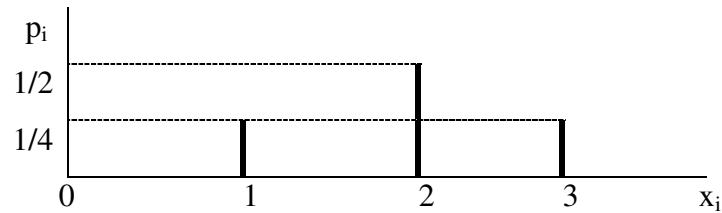
$$\begin{pmatrix} 1 & 2 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 2 & 3 & 4 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix},$$

în timp ce, dacă considerăm că X și X nu iau valori independent, atunci

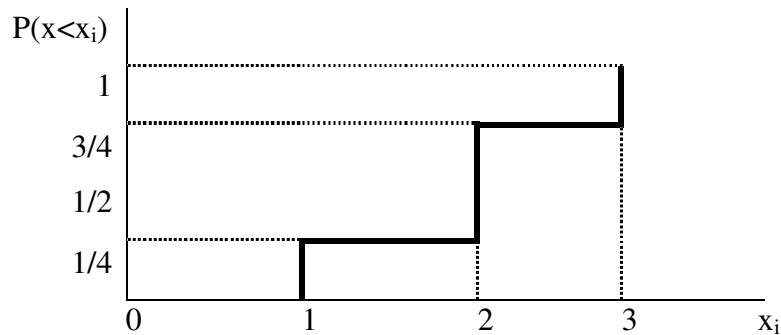
$$X+X = 2X = \begin{pmatrix} 2 & 4 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

Putem reprezenta grafic aceste probabilități.

De exemplu, $X = \begin{pmatrix} 1 & 2 & 3 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}$ apare sub forma



Dar putem reprezenta curba cumulativă a distribuției



Definiție

Funcția de repartiție asociată lui f este funcția $F(x)$, $F: \mathbb{R} \rightarrow [0,1]$ definită de formula:

$$F(x) = P(f < x) = P(f^{-1}(-\infty, x))$$

Importanța acestei funcții constă în faptul că, dacă $F(x)$ este dată se poate determina probabilitatea ca f să ia valori într-un interval $I \subset \mathbb{R}$, oricare ar fi acel interval.

În cazul în care f ia un număr finit de valori, de exemplu $\{1,2,3\}$, când cunoaștem $P(f < k) \quad \forall k = 1,2,3$, cunoaștem practic și $P(f = k) \quad \forall k = 1,2,3$.

$$\text{Într-adevăr, } P(f = 1) = P(f < 2)$$

$$P(f = 2) = P((f < 3) \cap (f > 2)) = P(f < 3) * P(f > 2) = P(f < 3) * (1 - P(f < 2))$$

$$P(f = 3) = 1 - P(f = 1) - P(f = 2)$$

$$\text{Ca regulă generală: } P(f = k) = 1 - P(f < k+1) - P(f < k)$$

Deci am determinat o distribuție de probabilitate care poate fi reprezentată sub forma unei matrici:

$$P(f = k) = \begin{pmatrix} 1 & 2 & 3 \\ p_1 & p_2 & p_3 \end{pmatrix}$$

Proprietăți

Funcția de repartiție are următoarele proprietăți:

a) $a \leq b \Rightarrow F(a) \leq F(b)$

b) $\lim_{a \rightarrow -\infty} F(a) = 0$

c) $\lim_{a \rightarrow +\infty} F(a) = 1$

d) F este continuă la stânga.

Dacă F este continuă spunem că f este *variabilă aleatoare continuă*. În acest caz, probabilitatea ca f să ia orice valoare particulară este 0.

$$\forall \xi, P(f(x) = \xi) = 0$$

Exemplu:

Dacă ne punem problema probabilității ca temperatura în cameră să fie $t = 20,347562$ aceasta este evident zero și de fapt problema nici nu are sens – în măsura în care temperatura este o valoare medie în jurul căreia avem fluctuații continue. Dacă ne punem problema ca temperatura să fie într-un anumit interval noțiunea de funcție de repartiție capătă un conținut concret.

Definiție

Fie F(x) funcția de repartiție a unei variabile aleatoare ξ . Dacă există o funcție $\rho(x)$, integrabilă pe intervalul $(-\infty, +\infty)$, cu proprietatea că pentru orice $x \in \mathbb{R}$ este verificată egalitatea:

$$\rho(x) = \frac{\partial F}{\partial x}$$

atunci, $\rho(x)$ se numește *densitatea de repartiție* sau *densitatea de probabilitate* a variabilei aleatoare ξ ,

În acest caz, probabilitatea ca variabila aleatoare să ia valori într-un interval $(-\infty, a)$ este dată de formula:

$$P(\xi(x) < a) = F(a) = \int_{-\infty}^a \rho(t) dt$$

și respectiv:

$$P(b \leq \xi(x) < a) = F(a) - F(b) = \int_{-\infty}^a \rho(t) dt - \int_{-\infty}^b \rho(t) dt = \int_b^a \rho(t) dt$$

Definiție

Se numește *valoare medie* (sau *speranță matematică*) a unei valori aleatoare f , numărul

$$M(f) = \sum x_i p_i, \text{ atunci când } \xi \text{ este o variabilă aleatoare simplă și, respectiv}$$

$M(f) = \int_{-\infty}^{+\infty} x \rho(x) dx$, atunci când ξ este o variabilă aleatoare continuă, cu densitatea de probabilitate ρ .

În literatură, operatorul de medie se mai notează și cu E , de la “expectation” – speranță în engleză.

În cazul variabilelor simple se observă că valoarea medie a variabilei f este media ponderată a valorilor sale x_i , cu ponderile p_i , care reprezintă “frecvențele” de apariție ale valorilor respective.

Proprietăți ale mediei:

Dacă f și g sunt independente, atunci avem:

- a) $M(af) = aM(f)$
- b) $M(f+g) = M(f) + M(g)$
- c) $M(fg) = M(f)M(g)$

Vom schița o demonstrație a proprietății b):

$$M(f+g) = \sum_{k,l} P(F_k \cap G_l)(x_k + x_l) = \sum_k \left(\sum_l P(F_k \cap G_l) \right) x_k + \sum_l \left(\sum_k P(F_k \cap G_l) \right) x_l$$

Dar, pe de altă parte, folosind proprietățile intersecțiilor și reuniunilor de mulțimi, respectiv distributivitatea intersecției față de reuniune și a intersecției față de reuniune, și faptul că $\bigcup_l G_l = E$ avem $\sum_l P(F_k \cap G_l) = P(F_k \cap (\bigcup_l G_l)) = P(F_k)$ și similar, $\sum_k P(F_k \cap G_l) = P(G_l)$.

Deci,

$$M(f+g) = \sum_k P(F_k) x_k + \sum_l P(G_l) x_l = M(f) + M(g)$$

Noțiunea de medie se generalizează, definindu-se *momentul de ordin k* al unei variabile aleatoare:

$$M_k(f) = \sum x_i^k p_i, \text{ atunci când } \xi \text{ este o variabilă aleatoare simplă}$$

și respectiv,

$M_k(f) = \int_{-\infty}^{+\infty} x^k \rho(x) dx$, atunci când ξ este o variabilă aleatoare continuă.

Se numește *moment centrat de ordin k* al variabilei aleatoare f momentul de ordinul k al abaterii sale față de medie.

$$M_k^c(f) = \sum (x_i - \mu_f)^k p_i$$

și respectiv, $\mu_k^c = \int_{-\infty}^{+\infty} [x - M(f)]^k \rho(x) dx$, în cazul unei variabile aleatoare continue.

Dispersia de selecție, sau varianta unui șir de rezultate numerice ale unui experiment este media aritmetică a pătratelor abaterilor acestor valori față de media lor aritmetică \bar{X} .

Dacă x_1, x_2, \dots, x_n sunt cele n valori ale seriei, dispersia de selecție a acestora, s_x^2 este:

$$s_x^2 = \frac{\sum (x_i - \bar{X})^2}{n}$$

După cum vom vedea mai departe la statistică, o formulă mai utilă pentru dispersia de selecție este: $s_x^2 = \frac{\sum (x_i - \bar{X})^2}{n-1}$

Dispersia de selecție este indicatorul principal al împrăștierii datelor unui experiment.

Dispersia unei variabile aleatoare este conceptul ce generalizează dispersia de selecție.

Definiție

Dispersia variabilei aleatoare X de notează $D(X)$ sau σ^2 și este, în particular, momentul centrat de ordinul doi.

$$D(X) = \sigma^2 = M[(X - M(X))^2] = \int_{-\infty}^{+\infty} (x - M(X))^2 \rho(x) dx$$

și respectiv

$$\sigma^2 = M[(X - M(X))^2] = \sum (x_i - \mu_x)^2 p_i, \text{ atunci când variabila aleatoare este discretă.}$$

Rădăcina pătrată a dispersiei, σ , se numește **abaterea medie pătratică** a variabilei X , iar s_x abaterea standard.

Proprietăți

a) Pentru orice variabilă aleatoare X și orice constante a și b

$$D(aX+b) = a^2D(X)$$

b) Dacă X, Y sunt două variabile aleatoare independente

$$D(X+Y) = D(X) + D(Y)$$

Demonstrație:

Pentru orice două variabile aleatoare X și Y, cu mediile μ_X și respectiv μ_Y , avem

$$D(X+Y) = M(X+Y - \mu_X - \mu_Y)^2 = M(X - \mu_X)^2 + M(Y - \mu_Y)^2 + 2 M[(X - \mu_X)(Y - \mu_Y)] = D(X) + D(Y) + 2 M[(X - \mu_X)(Y - \mu_Y)]$$

Dar, atunci când X și Y sunt independente $\Rightarrow M(XY) = \mu_X \mu_Y$,

$$M[(X - \mu_X)(Y - \mu_Y)] = M(XY - X\mu_Y - Y\mu_X + \mu_X \mu_Y) = \mu_X \mu_Y - \mu_X \mu_Y - \mu_X \mu_Y + \mu_X \mu_Y = 0$$

$$\Rightarrow M[(X - \mu_X)(Y - \mu_Y)] = 0$$

și deci $D(X+Y) = D(X) + D(Y)$

c) Între dispersie, valoarea medie și momentul de ordinul doi există relația:

$$D(f) = M(f^2) - (M(f))^2$$

Demonstrație:

$$D(X) = \sum (x_i - \mu_X)^2 p_i = \sum x_i^2 p_i - 2 \sum x_i \mu_X p_i + \sum \mu_X^2 p_i = M(f^2) - 2 \mu_X^2 + \mu_X^2 = M(f^2) - (M(f))^2$$

Observație

Dacă numim $M(f^2)$ – media pătratului și $(M(f))^2$ – pătratul mediei formula capătă o formulare ușor de reținut: **”Dispersia este egală cu media pătratului, minus pătratul mediei”**.

Relația se mai poate scrie sub forma $M(X^2) = \mu_X^2 + \sigma_X^2$ și am putea s-o numim „teorema lui Pitagora în probabilitate”.

Exemplu

În modelul clasic al urnei cu bile pe care l-am prezentat mai sus, probabilitatea evenimentului “din n bile extrase, k sunt albe” era $p_k = C_n^k p^k q^{n-k}$.

Media variabilei aleatoare X care da numărul de bile albe din n bile extrase va fi, prin definiție,

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

Pentru a calcula această sumă considerăm următoarea identitate

$$(pt + q)^n = \sum C_n^k p^k t^k q^{n-k}, \text{ pe care o derivăm în raport cu } t$$

$$((pt + q)^n)' = \left(\sum C_n^k p^k t^k q^{n-k} \right)'$$

$$np(pt + q)^{n-1} = \sum C_n^k p^k kt^{k-1} q^{n-k} \text{ și apoi facem } t = 1 \Rightarrow np = \sum C_n^k p^k kq^{n-k}$$

Am obținut, deci, $M(X) = np$

Folosind aceiași identitate, dar derivând de două ori se arată că: $D(X) = npq$

Cunoașterea mediei și dispersiei unei variabile aleatoare dă o indicație asupra intervalului în care se află valorile variabilei, cu cea mai mare probabilitate. Mai exact, după cum arată teorema următoare, cu cât ne îndepărtăm mai mult de valoarea medie, cu atât valorile respective sunt mai puțin probabile ca valori ale variabilei date.

Inegalitatea lui Cebâșev

Dacă σ^2 este dispersia variabilei aleatoare X , probabilitatea ca modulul abaterii sale de la valoarea medie să ia valori mai mari decât un număr $\varepsilon > 0$ este mai mică decât

$$\frac{\sigma^2}{\varepsilon^2}.$$

$$P(|x_i - m| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

Demonstrație:

Pornim de la definiția dispersiei $\sigma^2 = M[(x_i - m)^2] = \sum (x_i - m)^2 p_i$ și împărțim suma în doi termeni: unul corespunzător valorilor x_i pentru care $|x_i - m| \geq \varepsilon$ și unul corespunzător valorilor lui x_i pentru care $|x_i - m| < \varepsilon$.

$$\sigma^2 = \sum (x_i - m)^2 p_i = \sum_{|x_i - m| < \varepsilon} (x_i - m)^2 p_i + \sum_{|x_i - m| \geq \varepsilon} (x_i - m)^2 p_i$$

Dacă neglijăm primul termen al sumei și minorăm $|x_i - m|$ înlocuindu-l cu ε în al doilea termen, se obține

$$\sigma^2 \geq \sum_{|x_i - m| \geq \varepsilon} \varepsilon^2 p_i = \varepsilon^2 (p_{k_1} + p_{k_2} + \dots + p_{k_n}),$$

cu $p_{k_1} + p_{k_2} + \dots + p_{k_n}$ suma probabilităților valorilor x_{k_i} pentru care $|x_{k_i} - m| \geq \varepsilon$.

Dar $p_{k_1} + p_{k_2} + \dots + p_{k_n} = P(|x - m| \geq \varepsilon)$ și deci am obținut $\sigma^2 \geq \varepsilon^2 P(|x - m| \geq \varepsilon)$ ceea ce implică următoarea relație: $P(|x - m| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$.

Deoarece suma între probabilitatea unui eveniment A și probabilitatea evenimentului contrar CA este 1, avem $P(CA) = 1 - P(A)$ și inegalitatea se mai poate scrie sub forma

$$P(|x_i - m| < \varepsilon) \geq 1 - \frac{\sigma^2}{\varepsilon^2}$$

Exemplu:

Fie $\varepsilon = 3\sigma$, atunci inegalitatea Cebâșev dă: $P(|x_i - m| < 3\varepsilon) = 1 - \frac{1}{9} = \frac{8}{9} = 0.88$

Exprimat în cuvinte, această inegalitate aparent banală, spune din punct de vedere fenomenologic, enorm de mult:

Probabilitatea ca orice variabilă aleatoare să ia valori mai îndepărtate de valoarea sa medie decât de trei valori standard, este mai mică decât 0,12.

Vom vedea mai departe că, în cazul în care variabila aleatoare are suplimentar unele proprietăți de regularitate, această probabilitate este chiar mult mai mică.

Aceiași inegalitate ne permite înțelegerea legăturii între frecvența și probabilitate, legătura care exprimă însăși fundamentarea statisticii pe teoria probabilităților.

Să considerăm variabila aleatoare care dă numărul de bile albe într-o extracție de n bile din urnă. Pentru această variabilă avem următoarea teoremă, care se generalizează în teoria probabilităților în forme care depășesc însă cadrul acestei lucrări.

Teorema lui Bernoulli:

Dacă se notează cu p probabilitatea ca un eveniment A (de exemplu apariția bilei albe) să se realizeze într-un experiment și $f_n = \frac{k}{n}$ este frecvența cu care se realizează evenimentul A în n experimente identice consecutive, șirul (f_n) converge către p în probabilitate. Altfel spus:

Frecvența tinde în probabilitate la probabilitatea teoretică.

Demonstrație:

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{k}{n} - p\right| \geq \varepsilon\right) = \lim_{n \rightarrow \infty} P(|k - np| \geq n\varepsilon) = \lim_{n \rightarrow \infty} P(|k - M(k)| \geq n\varepsilon)$$

Dar, aplicând inegalitatea lui Cebâșev: $P(|k - M(k)| \geq n\varepsilon) \leq \frac{\sigma^2}{n^2 \varepsilon^2}$ și deci

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{k}{n} - p\right| \geq \varepsilon\right) \leq \lim_{n \rightarrow \infty} \frac{\sigma^2}{n^2 \varepsilon^2} = 0$$

Teorema lui Bernoulli afirmă numai că inegalitatea $|f_n - p| \geq \varepsilon$ nu are șansa să fie realizată sau că inegalitatea $|f_n - p| < \varepsilon$ are șanse mari să fie îndeplinită dacă n este suficient de mare.

DISTRIBUȚII DE PROBABILITATE

Distribuția normală

Spunem că o variabilă aleatoare este normal repartizată $N(m, \sigma)$, atunci când densitatea sa de probabilitate este data de formula:

$$\rho(x, m, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

O primă condiție ca $\rho(x)$ să fie distribuție de probabilitate este aceea că

$$\int_{-\infty}^{+\infty} \rho(x) dx = P(-\infty < f(t) < +\infty) = 1$$

Pentru a verifica această condiție, plecăm de la un rezultat care s-a obținut la cursul de matematică folosind integrala dublă, și anume :

$$\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi}$$

În cazul nostru, dacă facem schimbarea de variabilă $u = \frac{x-m}{\sigma}$ avem

$$\int_{-\infty}^{+\infty} \rho(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{(x-m)^2}{2\sigma^2}} dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}} \sigma du = 1$$

Vom arăta în continuare că o variabilă aleatoare normal repartizată are media m și dispersia σ^2 .

Să calculăm mai întâi media:

$$M[X] = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} x e^{-\frac{(x-m)^2}{2\sigma^2}} dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} (x-m+m) e^{-\frac{(x-m)^2}{2\sigma^2}} dx =$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{\sigma(x-m)}{\sigma} e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2} dx + m = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} u e^{-\frac{u^2}{2}} \sigma du + m = 0 + m = m$$

Integrala este nulă deoarece funcția de integrat este impară.

Pentru calculul dispersiei ne folosim de identitatea:

$$D(X) = M[X - M(X)]^2 = M(X^2) - [M(X)]^2$$

$$M(X^2) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 e^{-\frac{(x-m)^2}{2\sigma^2}} dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} (m + \sigma u)^2 e^{-\frac{u^2}{2}} \sigma du =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \left(m^2 e^{-\frac{u^2}{2}} + 2m\sigma u e^{-\frac{u^2}{2}} + \sigma^2 u^2 e^{-\frac{u^2}{2}} \right) du =$$

$$= \frac{1}{\sqrt{2\pi}} \left(m^2 \sqrt{2\pi} + \sigma^2 \int_{-\infty}^{+\infty} u^2 e^{-\frac{u^2}{2}} du \right)$$

Calculăm separat integrala rămasă și obținem:

$$\int_{-\infty}^{+\infty} u^2 e^{-\frac{u^2}{2}} du = - \int_{-\infty}^{+\infty} u \left(-u e^{-\frac{u^2}{2}} \right) du = u e^{-\frac{u^2}{2}} \Big|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} 1 \cdot \left(-e^{-\frac{u^2}{2}} \right) du = \sqrt{2\pi}$$

unde am integrat prin părți, luând $u = \varphi$ și $-u e^{-\frac{u^2}{2}} = \psi'$

Deci am obținut $M(X^2) = \frac{1}{\sqrt{2\pi}} (m^2 \sqrt{2\pi} + \sigma^2 \sqrt{2\pi})$ și înlocuind în expresia lui

$D(X)$ obținem:

$$D(X) = \frac{1}{\sqrt{2\pi}} (m^2 \sqrt{2\pi} + \sigma^2 \sqrt{2\pi}) - m^2 = \sigma^2$$

Pornind de la proprietățile operatorilor de medie și dispersie

$$M(X - m) = M(X) - m$$

$$D(X - m) = D(X) \text{ și}$$

$$D\left(\frac{X}{a}\right) = \frac{1}{a^2} D(X)$$

se obține că, dacă o variabilă aleatoare este normal repartizată $N(m, \sigma)$, variabila aleatoare redusă $\frac{X - m}{\sigma}$ este repartizată $N(0,1)$, deci cu distribuția de probabilitate

$$\rho(x) = e^{-\frac{x^2}{2}}$$

Funcția de repartiție asociată este funcția $\Phi(t) = \int_{-\infty}^t e^{-\frac{x^2}{2}} dx$ numită *funcția lui Laplace* și ale cărei valori se găsesc în tabelele din practic toate cărțile de statistică și probabilități.

Distribuție binomială

Distribuția binomială apare, așa cum s-a arătat mai sus, la descrierea evenimentelor asociate extracțiilor dintr-o urnă cu bile albe și bile negre.

Distribuția variabilei aleatoare “numărul de bile albe din n bile extrase” se poate reprezenta și sub formă matricială:

$$X = \begin{pmatrix} 0 & 1 & k & n \\ C_n^0 p^0 q^n & C_n^1 p^1 q^{n-1} & \dots & C_n^k p^k q^{n-k} & \dots & C_n^n p^n q^0 \end{pmatrix}$$

După cum am arătat media și dispersia unei variabile aleatoare repartizate binomial sunt $M = np$ și $D = npq$

Repartiția binomială apare întotdeauna atunci când un experiment cu numai două răspunsuri posibile se repetă de n ori. Un caz particular îl prezintă experimentele care se repetă de un număr foarte mare de ori, iar evenimentul în a cărui apariție suntem interesați are o probabilitate foarte mică, categorisit uzual ca “eveniment rar”.

La limită, când $n \rightarrow \infty$, $p \rightarrow 0$, dar np rămâne constant, $np = \lambda$, se obține distribuția Poisson.

Distribuția POISSON

Considerăm deci că $np = \lambda$ și trecem la limită după n

$$\begin{aligned} \lim_{n \rightarrow \infty} C_n^k p^k q^{n-k} &= \lim_{n \rightarrow \infty} \frac{n(n-1)\dots(n-k+1)}{k!} \frac{\lambda^k}{n^k} \left(1 - \frac{\lambda}{n}\right)^{n-k} = \\ &= \frac{1}{k!} * \lim_{n \rightarrow \infty} \frac{n(n-1)\dots(n-k+1)}{n^k} \lambda^k \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^{n-k} \end{aligned}$$

$$\text{dar } \lim_{n \rightarrow \infty} \frac{n(n-1)\dots(n-k+1)}{n^k} = 1 \text{ și } \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^{n-k} = \lim_{n \rightarrow \infty} \left[\left(1 - \frac{\lambda}{n}\right)^{-\frac{n}{\lambda}} \right]^{\frac{n-k}{n}(-\lambda)} = e^{-\lambda}$$

și deci,

$$\lim_{n \rightarrow \infty} C_n^k p^k q^{n-k} = \frac{\lambda^k}{k!} e^{-\lambda}$$

Deci, distribuția Poisson este dată de matricea

$$X = \begin{pmatrix} 0 & 1 & k & n \\ e^{-\lambda} & \frac{\lambda}{1!} e^{-\lambda} & \dots \frac{\lambda^k}{k!} e^{-\lambda} & \dots \frac{\lambda^n}{n!} e^{-\lambda} \end{pmatrix}$$

Calculând, după definiție, media și dispersia unei variabile aleatoare distribuite Poisson și ținând cont că

$$\sum_{k \geq 0} \frac{\lambda^k}{k!} = e^\lambda, \quad \sum_{k \geq 0} k \frac{\lambda^k}{k!} = \lambda e^\lambda, \quad \sum_{k \geq 2} k(k-1) \frac{\lambda^k}{k!} = \lambda^2 e^\lambda, \quad \sum_{k \geq 1} k \frac{\lambda^k}{k!} = \lambda e^\lambda$$

se obține

$$\begin{aligned} M(X) &= \sum_{k \geq 0} k \frac{\lambda^k e^{-\lambda}}{k!} = e^{-\lambda} \sum_{k \geq 1} \frac{\lambda^k}{(k-1)!} = e^{-\lambda} \lambda \sum_{k \geq 1} \frac{\lambda^{k-1}}{(k-1)!} = e^{-\lambda} \lambda e^\lambda = \lambda \\ D(X) &= e^{-\lambda} \sum_{k \geq 0} \frac{(k-\lambda)^2 \lambda^k}{k!} = e^{-\lambda} \left(\sum_{k \geq 0} \frac{k^2 \lambda^k}{k!} - 2\lambda \sum_{k \geq 0} \frac{k \lambda^k}{k!} + \lambda^2 \sum_{k \geq 0} \frac{\lambda^k}{k!} \right) = \\ &= e^{-\lambda} \left(\sum_{k \geq 1} [k(k-1) + k] \frac{\lambda^k}{k!} - \lambda^2 e^\lambda \right) = e^{-\lambda} \left[\sum_{k \geq 2} k(k-1) \frac{\lambda^k}{k!} + \sum_{k \geq 1} k \frac{\lambda^k}{k!} \right] - \lambda^2 = \\ &= e^{-\lambda} (\lambda^2 e^\lambda + \lambda e^\lambda) - \lambda^2 = \lambda \end{aligned}$$

Exemplu:

Numărul evenimentelor adverse la un medicament dat este repartizat Poisson.

Cel mai mult este utilizată distribuția Poisson în fizica statistică.

Aproximarea normală a distribuției binomiale

Ca o regulă generală, dacă np și nq sunt mai mari sau egale cu 5, poate fi folosită aproximarea normală. Pentru distribuțiile binomiale în care p < 0,5 aproximarea este bună

pentru valori ale lui np și nq mai mici decât 5. În aceste condiții, $\frac{k - np}{\sqrt{npq}} = \frac{\frac{k}{n} - p}{\sqrt{\frac{pq}{n}}}$ este

aproximativ normal distribuit cu media 0 și deviația standard 1.

Această transformare înlesnește de obicei calculul probabilităților binomiale.

Repartitia χ^2 Helmer - Pearson

Se consideră n observații independente x_1, x_2, \dots, x_n (variabile aleatoare independente) normal distribuite $N(\xi, \sigma^2)$.

Variabilele standard $u_i = \frac{x_i - \xi}{\sigma}$, $i = \overline{1, n}$ sunt de asemenea independente, iar suma pătratelor lor va avea o distribuție ce poate fi determinată.

Se definește $X = \sum_1^n u_i^2$.

Distribuția variabilei X rezultate se notează $\chi^2(n)$ și este diferită pentru fiecare valoare a lui n , iar parametru n se definește ca numărul de grade de libertate.

Vom determina în continuare parametrii (media și dispersia) unei variabile distribuite χ^2 .

Pentru a afla media distribuției χ^2 este necesară aflarea lui $M[u_i^2]$.

Deoarece $M[u_i] = 0$, $M[u_i^2] = M[(u_i - M[u_i])^2] = D[u_i] = 1$

Ca urmare $M[\chi^2(n)] = M[\sum_1^n u_i^2] = \sum_1^n M[u_i^2] = n * 1 = n$

Dispersia va fi:

$$D[\chi^2(n)] = D[\sum_1^n u_i^2] = \sum_1^n D[u_i^2] = nD[u_i^2] = n[M(u_i^4) - (M(u_i^2))^2] = n[M(u_i^4) - 1]$$

Pentru a obține $M[u_i^4]$ se folosește regula integrării prin părți:

$$\int f(x)g'(x)dx = f(x)g(x) - \int f'(x)g(x)dx$$

În acest caz se va identifica:
 $f(x) = u^3 \Rightarrow f'(x) = 3u^2$
 $g(x) = e^{-\frac{u^2}{2}} \Rightarrow g'(x) = ue^{-\frac{u^2}{2}}$, deci se va obține:

$$M[u_i^4] = \int_{-\infty}^{+\infty} u^4 \rho(u) du = \int_{-\infty}^{+\infty} u^4 \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} u^3 u e^{-\frac{u^2}{2}} du = \frac{1}{\sqrt{2\pi}} u^3 \left(e^{-\frac{u^2}{2}} \right) \Big|_{-\infty}^{+\infty} - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} 3u^2 \left(-e^{-\frac{u^2}{2}} \right) du = 3 \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} u^2 e^{-\frac{u^2}{2}} du = 3M[u^2] = 3$$

Atunci,

$$D[u_i^2] = M[u_i^4] - (M[u_i])^2 = 3 - (1)^2 = 2$$

și substituind în relația de mai sus se va obține

$$D[\chi^2(n)] = nD[u_i^2] = 2n$$

Deci variabila $x^2 = x_1^2 + x_2^2 + \dots + x_n^2$ este repartizată $\chi^2(n)$, cu n grade de libertate, având media $E(\chi^2) = n$, respectiv dispersia $D(\chi^2) = 2n$.

Se poate arăta că densitatea de probabilitate este dată de funcția

$$f(\chi^2) = \frac{1}{\Gamma\left(\frac{n}{2}\right) 2^{\frac{n}{2}}} e^{-\frac{\chi^2}{2}} (\chi^2)^{\frac{n}{2}-1},$$

unde Γ este funcția Euler de speța I-a studiată la cursul de matematică șă anume : $\Gamma(\alpha) = \int_0^{+\infty} e^{-t} t^{\alpha-1} dt$.

Repartitia χ^2 se folosește foarte mult în statistica matematică în verificarea ipotezelor asupra egalității dispersiilor.

Repațiția STUDENT

Analog cu distribuția χ^2 , repațiția t a fost propusă de Student (pseudonimul lui W.S.Gosset, chimist statistician englez), pentru statistica selecțiilor mici și exprimă deviațiile mediilor de selecție \bar{x} , față de media întregii populații μ , măsurate în $\frac{s}{\sqrt{n}}$ (abaterea standard a mediilor de selecție).

Dacă sunt date două variabile aleatoare $Z \in N(0,1)$ și $V \in \chi^2(n)$ independente, se

spune că variabila $t = \frac{Z}{\sqrt{\frac{V}{n}}} \in t(n)$ este repartizată Student cu n grade de libertate.

Mărima t nu depinde decât de numărul gradelor de libertate.

Distribuția de probabilitate a unei variabile aleatoare repartizate Student tinde

pentru $n \rightarrow \infty$, la distribuția normală $\rho(t) \rightarrow \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$

Densitatea de probabilitate este dată de funcția:

$$f(x) = \frac{1}{\sqrt{n\pi}} * \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} * \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}} \text{ unde } x \in R \text{ și } n \in N.$$

Repartiția F (Behrens - Fisher - Snedecor) sau distribuția raportului a două dispersii

Se consideră frecvent în statistică raportul a două dispersii care estimează aceeași dispersie generală a unei colectivități. Dintr-o colectivitate generală se extrag două selecții $U \in \chi^2(n_1)$, $V \in \chi^2(n_2)$. Raportul lor este o variabilă aleatoare repartizată F

$$F = \frac{\frac{U}{n_1}}{\frac{V}{n_2}} \in F(n_1, n_2)$$

Examinând acest raport se observă că el nu conține dispersia colectivității generale σ^2 , de unde rezultă că distribuția acestui raport nu depinde decât de numărul gradelor de libertate n_1 și n_2 ale celor două dispersii.

Densitatea de probabilitate este dată de funcția:

$$f(x) = \frac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right) * \Gamma\left(\frac{n_2}{2}\right)} * \left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} * x^{\frac{n_1}{2}-1} * \left(1 + \frac{n_1}{n_2} * x\right)^{-\frac{n_1+n_2}{2}}, \text{ când } x > 0.$$

¹Andrei Nicolaevici Kolmogorov (1903-1987), fost profesor la Universitatea din Moscova, a avut contribuții deosebite în analiza matematică, analiza funcțională și teoria probabilităților. Cartea sa "Grundbegriffe der Wahrscheinlichkeitsrechnung", Berlin, 1933, a însemnat o revoluție în teoria probabilităților, arătând că, formal, această teorie se poate trata ca un caz particular de teorie a integralei (sau "teoria măsurii").