Introducere în Calculul Probabilităţilor

I.L. Stoica

Dedic această carte memoriei
profesorului Aurel Cornea.

PREFAŢĂ. Cartea de faţă este o introducere elementară în teoria probabilităţilor şi reprezintă în principal cursul pe care îl ţin la studenţii din anul doi, profilul informatică, la Facultatea de Matematică şi Informatică. Teoria măsurii este dezvoltată la nivel discret în acelaşi timp cu construcţia modelelor probabiliste. Obiectivele principale ale cursului sunt: 1) explicarea modelării probabiliste, 2) înţelegerea noţiunii de independenţă, 3) interpretarea variabilelor aleatoare şi a repartiţiilor în modele, 4) unele idei privitoare la teoremele fundamentale, care sunt legea numerelor mari şi teorema limită centrală. Datorită timpului limitat, acestea din urmă sunt prezentate în cursul ţinut în faţa studenţilor într-o manieră foarte condensată. De altfel toate expunerile pot fi suplimentate şi cu alte complemente. Pentru aceasta, în textul scris am inclus numeroase adăugiri ce nu au loc în expunerile orale. Am însemnat cu un asterix titlurile pe care nu am ajuns să le expun în ultimii doi ani.

Prezentarea elementară pe care am adoptat -o sperăm să contribuie la popularizarea subiectului, popularizare care este tot mai cerută de aplicaţiile teoriei probabilităţilor în diverse domenii ale activităţii din societatea contemporană.

Bibliografia indică o serie de cărţi dedicate introducerii în teoria probabilităţilor, pe care le-am consultat şi care pot fi utile şi cititorului în continuarea studiului.

Cuprins

Introducere

Teoria probabilităţilor reprezintă un domeniu aparte al matematicii, ce are relaţiile sale specifice cu realitatea înconjurătoare şi care se raportează la ea prin experienţă, ca o adevărată ştiinţă a naturii. Din punct de vedere filozofic, conceptul de probabilitate, care este conceptul central al teoriei, este caracterizat în două ipostaze.

O primă ipostază este apariţia sa în exprimarea *frecvenţei* unui eveniment în timpul desfăşurării unui proces uniform. De exemplu, să presupunem că facem experienţa aruncării cu banul de un număr mare de ori. Urmărim fiecare aruncare, notând de fiecare dată rezultatul: stema sau cifra. Chiar înainte de a începe experienţa putem gândi logic că nu exisă nici un motiv să fie preferată vreuna din feţele monedei. Deci " a priori" spunem că şansele de a cădea stema sau cifra sunt egale. Faptul neaşteptat, care apare în timpul experienţei, este că după un număr nu prea mare de aruncări (de exemplu 100 de aruncări) se şi verifică aceasta. Notăm cu N_n numărul de experienţe la care rezultatul a fost cifra după ce s-au făcut n aruncări. Atunci graficul funcţiei $f(n) = \frac{N_n}{n}, n \in \mathbf{N}$, arată cam aşa cum se vede în figura 1. Graficul din figură a fost realizat pe baza datelor unui experiment ce a constat din 200 de aruncări cu o monedă. După cum se vede, prima porţiune a graficului este lipită de axa orizontală. Aceasta reflectă faptul că în experiment s-a obţinut la început de zece ori la rând faţa cu stema. Putem spune că aşa ceva este neobişnuit. A preciza

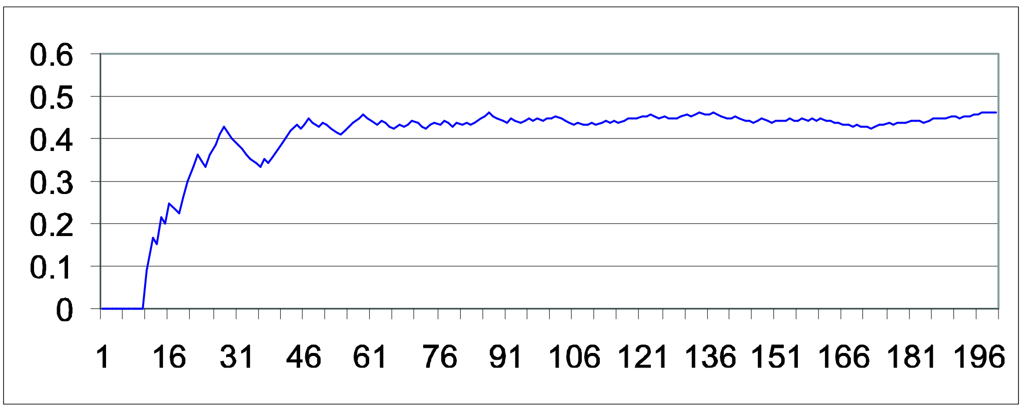


FIGURA 1. Graficul frecvenţei cu care a ieşit o faţă a monedei

însă noţiunea de „obişnuit" sau „frecvent" în legătură cu un fenomen aleator este tocmai obiectul teoriei probabilităţilor. Graficul evoluează apoi apropiindu-se de valoarea constantă $0,5$. Acest fapt este foarte obişnuit la un număr mai mare de 100 de aruncări. Oricine poate verifica acest lucru experimentând aruncarea cu banul! Deci probabilitatea egală pentru obţinerea unei feţe sau a alteia este reflectată în practică prin *frecvenţa* egală de realizare a celor două rezultate atunci când experimentul este repetat de un număr relativ mare de ori.

Cea de a doua ipostază a conceptului de probabilitate este cea în care exprimă o cuantificare a unui fenomen subiectiv, în care intervine *necunoaşterea* unor aspecte ale unui experiment. De exemplu, un medic se poate exprima cu privire la şansele de reuşită ale unui tratament sau ale unei operaţii şi atunci când pacientul are o stare diferită de cazurile ce le cunoscuse anterior. Medicul face deci aprecieri subiective ale unor aspecte necunoscute şi spune spre exemplu "şansele de însănătoşire sunt de 80%".

Acelaşi lucru se întâmplă cu un economist care trebuie să facă prognoze. Economia fiecărui an este în mare măsură unică, nu se poate vorbi de repetarea unor fenomene decât parţial, şi deci, un important aspect al prognozei este caracterul său subiectiv. Apare astfel o evaluare probabilistă a şanselor de desfăşurare a unor scenarii, necunoscându-se care vor fi condiţiile reale.

Deşi din punct de vedere filosofic bazele teoriei probabilităţilor au implicaţii foarte interesante, pentru studiul matematic, cât şi pentru aplicaţiile practice obişnuite, aspectele filosofice sunt aproape irelevante. De aceea nu ne vom ocupa deloc în cele ce urmează de fundamentele teoriei probabilităţilor, ci vom trece direct la definirea obiectelor matematice şi la explicarea procesului de modelare prin exemple. Eficienţa practică a modelării este evident suficientă pentru a înlocui orice discurs filosofic.

Din punctul de vedere al matematicianului, teoria probabilităţilor are două părţi: modelarea unor fenomene reale (în legătură cu modelarea apare şi problema verificării modelului prin metodele statisticii, care validează sau invalidează modelul) şi studiul obiectelor matematice create prin modelare. Studiul acesta conduce la noi concepte şi legături cu celelalte ramuri ale matematicii, ceea ce permite ulterior crearea altor modele, de un nivel superior.

Modelarea se face în primul rând prin construirea unui *spaţiu probabilizat*, prin intermediul căruia se asociază evenimentelor posibile numere din intervalul $[0,1]$. Unui eveniment mai probabil decât un altul trebuie să-i corespundă un număr mai mare decât celuilalt. Evenimentele cele mai probabile vor avea asociate numere apropiate de 1, de exemplu $0,95$ sau $0,99$ în timp ce evenimentele cele mai puţin probabile vor avea asociate numere mici de genul $0,05$ sau $0,01$.

CAPITOLUL 1

Modelul probabilist

În prima parte a acestui curs vom analiza mai ales teoria finită, în care există doar o familie finită de evenimente posibile legate de fenomenul studiat. De aceea vom introduce mai întâi noţiunea de *spaţiu probabilizat finit*.

Definiţia 1.1. *Fie* Ω *o mulţime şi* $\mathcal{F}$ *o familie de părţi, deci* $\mathcal{F} \subset \mathcal{P}(\Omega)$. *Spunem că* $\mathcal{F}$ *este o* algebră de părţi, *sau pe scurt o algebră, dacă sunt îndeplinite condiţiile următoare:*

(i) mulţimea vidă şi spaţiul total aparţin lui $\mathcal{F}$*, adică* $\emptyset, \Omega \in \mathcal{F}$;

(ii) familia este închisă la complementară, în sensul că $A^c \in \mathcal{F}$*, dacă* $A \in \mathcal{F}$;

(iii) familia este închisă la reuniune, adică $A \cup B \in \mathcal{F}$*, dacă* $A, B \in \mathcal{F}$.

Se mai utilizează şi termenul de *corp* pentru noţiunea de algebră de părţi. Dacă, în plus, familia $\mathcal{F}$ este finită spunem că ea este o *algebră finită*. În acest caz, perechea $(\Omega, \mathcal{F})$ este numită *spaţiu măsurabil finit*.

Exemplul cel mai simplu de algebră este algebra trivială $\{\Omega, \emptyset\}$, care conţine numai spaţiul total şi mulţimea vidă. Alt exemplu simplu este constituit de $\mathcal{P}(\Omega)$, mulţimea tuturor părţilor spaţiului Ω. În cazul în care Ω este o mulţime finită, cel mai adesea ea este considerată ca spaţiu măsurabil fiind înzestrată cu algebra $\mathcal{P}(\Omega)$.

Dacă avem dată o algebră $\mathcal{F} \subset \mathcal{P}(\Omega)$ şi $\Lambda \subset \Omega$ este o submulţime arbitrară, familia de părţi ale lui Λ, care este descrisă prin $\{A \cap \Lambda / A \in \mathcal{F}\}$, este tot o algebră. Ea este numită urma lui $\mathcal{F}$ pe Λ.

Un exemplu important de algebră este cel al părţilor măsurabile Jordan dintr-un spaţiu euclidian. O mulţime din spaţiul euclidian se numeşte *măsurabilă Jordan*, dacă frontiera sa este neglijabilă. Se verifică uşor că aceste mulţimi alcătuiesc o algebră de părţi. (Ea nu este finită.)

Se observă că o algebră $\mathcal{F}$ are proprietatea că este închisă la intersecţie şi diferenţă: dacă $A, B \in \mathcal{F}$, atunci $A \cap B, A \backslash B \in \mathcal{F}$. De asemenea, prin inducţie se constată că reuniunea şi intersecţia unui număr finit arbitrar de mulţimi din $\mathcal{F}$ sunt tot în $\mathcal{F}$.

Definiţia 1.2. *Fie* $(\Omega, \mathcal{F})$ *un spaţiu măsurabil finit. Spunem că o aplicaţie* $P : \mathcal{F} \to [0,1]$, *este o* măsură de probabilitate *dacă satisface relaţiile*

(i) $P(\Omega) = 1$,

(ii) $P(A \cup B) = P(A) + P(B)$, *pentru orice două mulţimi disjuncte,* $A, B \in \mathcal{F}$.

În acest caz, tripletul $(\Omega, \mathcal{F}, P)$ *va fi numit* spaţiu probabilizat finit.

Din definiţia dată rezultă imediat următoarele proprietăţi:

1. $P(A^c) = 1 - P(A)$, pentru orice $A \in \mathcal{F}$. În particular, se deduce de aici că are loc relaţia $P(\emptyset) = 0$. Mai observăm că relaţia $P(A) = 1$ este echivalentă cu relaţia $P(A^c) = 0$. În acest caz, se spune că mulţimea A *suportă* măsura P şi că măsura P nu *încarcă* A^c.

2. $P(A \backslash B) = P(A) - P(B)$ şi $P(B) \leq P(A)$ ori de câte ori $A, B \in \mathcal{F}$ satisfac relaţia $B \subset A$.

3. $P(A_1 \cup ... \cup A_n) = P(A_1) + ... + P(A_n)$, pentru orice număr finit de mulţimi disjuncte $A_1, ..., A_n \in \mathcal{F}, n \in \mathbf{N}$.

4. $P(A_1 \cup ... \cup A_n) \leq P(A_1) + ... + P(A_n)$, în general, cu mulţimile arbitrare $A_1, ..., A_n \in \mathcal{F}, n \in \mathbf{N}$.

5. Dacă mulţimea Ω este finită, rezultă că o măsură de probabilitate P pe $\mathcal{P}(\Omega)$ este complet determinată de valorile pe care le ia pe mulţimile cu un singur punct. Mai precis, să presupunem că $\Omega = \{x_1, ..., x_n\}$ şi să notăm $P(x_i) = P(\{x_i\}), i = 1, ..., n$. Ţinând cont de punctul 3., aceste numere trebuie să satisfacă relaţia $P(x_1) + ... + P(x_n) = 1$, pentru că $\bigcup_{i=1}^{n} \{x_i\} = \Omega$. Pentru o mulţime arbitrară $A \subset \Omega$ are loc formula

$$P(A) = \sum_{i, x_i \in A} P(x_i).$$

Reciproc, fiind date numerele $a_1, ..., a_n \in [0, 1]$, astfel încât $a_1 + ... + a_n = 1$, se poate defini o măsură de probabilitate punând $P(x_i) = a_i$, pentru orice $i = 1, ..., n$ şi definind apoi probabilitatea unei mulţimi arbitrare după formula anterioară.

De multe ori un spaţiu probabilizat finit este construit prin definirea mulţimii Ω şi a unei partiţii finite a acesteia, $(A_k)_{1 \leq k \leq n}$. Deci mulţimile A_k sunt nevide şi astfel definite încât au loc relaţiile: a) $A_l \cap A_k = \emptyset$, dacă $l \neq k$ şi b) $\Omega = \cup_{1 \leq k \leq n} A_k$. Algebra generată de această partiţie, pe care o notăm $\mathcal{F}$, constă din toate mulţimile de forma $A_{k_1} \cup ... \cup A_{k_m}$, cu indicii satisfacând relaţia $1 \leq k_1 < ... < k_m \leq n$, la care se mai adaugă mulţimea vidă. (Lăsăm cititorului verificarea faptului că aceasta este o algebră de părţi.) Probabilitatea P este apoi definită prin fixarea valorilor pe elementele partiţiei $p_k = P(A_k), k = 1, ..., n$. Desigur, numerele $p_k, k = 1, ..., n$ trebuie să fie pozitive şi să aibă suma $p_1 + ... + p_n = 1$. Pentru o mulţime arbitrară din $\mathcal{F}$ se defineşte apoi

$$P(A_{k_1} \cup ... \cup A_{k_m}) = P(A_{k_1}) + ... + P(A_{k_m}),$$

pentru $k_1 < ... < k_m$. Cititorul poate uşor verifica faptul că în acest fel este definită o măsură de probabilitate pe $\mathcal{F}$.

O metodă convenabilă de construcţie a unui spaţiu probabilizat finit $(\Omega, \mathcal{F}, P)$ constă în alegera mulţimii Ω ca o parte dintr-un spaţiu euclidian, de exemplu din plan. Mulţimile din $\mathcal{F}$ sunt alese sugestiv după aspectul grafic. De exemplu cercuri, pătrate, dreptunghiuri, ovale, etc. În plan se poate uşor construi un spaţiu probabilizat desenând o diagramă Venn şi punând ca probabilitate a unei mulţimi figurate A, valoarea $P(A) = \frac{ariaA}{aria\Omega}$. Cel mai adesea în astfel de cazuri, pentru a nu complica exprimarea analitică a diverselor mulţimi, liniile desenate, care separă regiunile, sunt excluse din Ω. De exemplu, a se vedea modelul zarului turtit din paragraful următor sau modelul celor trei urne din paragraful despre formula lui Bayes.

În primele capitole ale acestei cărţi nu vom utiliza decât spaţii probabilizate finite. Totuşi, anumite noţiuni generale le vom da în contextul unui spaţiu probabilizat general, pentru a evita repetarea definiţiilor mai târziu. De aceea introducem acum şi definiţia unui spaţiu probabilizat general.

Definiţia 1.3. *O algebră $\mathcal{F}$, de părţi ale lui Ω, este numită $\sigma-$* algebră *dacă are proprietatea de a fi închisă la reuniuni numărabile: dacă $(A_n)_{n\in\mathbf{N}}$ este un şir de elemente din $\mathcal{F}$, atunci $\cup_{n\in\mathbf{N}}A_n \in \mathcal{F}$. În acest caz perechea $(\Omega, \mathcal{F})$ se numeşte* spaţiu măsurabil.

Se mai utilizează şi denumirea de *corp borelian* pentru o σ-algebră. Bineînţeles că orice algebră finită este o $\sigma-$ algebră. Familia $\mathcal{P}(\Omega)$, a tuturor părţilor mulţimii Ω, este o $\sigma-$algebră. Se observă, pe baza formulelor lui de Morgan, că o σ -algebră este închisă şi la intersecţii numărabile: dacă $(A_n)_{n\in\mathbf{N}}$ este un şir de elemente din $\mathcal{F}$, atunci $\bigcap_{n\in\mathbf{N}} A_n \in \mathcal{F}$. Dacă avem dată o $\sigma-$algebră $\mathcal{F} \subset \mathcal{P}(\Omega)$ şi $\Lambda \subset \Omega$ este o submulţime arbitrară, *urma* lui $\mathcal{F}$ pe Λ este tot o $\sigma-$algebră.

Dar cel mai important exemplu de $\sigma-$algebră este cea constituită din mulţimile măsurabile Borel dintr-un spaţiu euclidian E (de o anumită dimensiune), care este notată cu $\mathcal{B}(E)$. Această $\sigma-$algebră este, prin definiţie, cea mai mică σ -algebră care conţine toate mulţimile deschise din spaţiul euclidian E. Se mai spune că este σ -algebra generată de familia mulţimilor deschise.

Pe de altă parte, este de reţinut că mulţimile măsurabile Jordan nu formează o $\sigma-$algebră. De exemplu mulţimea lui Cantor de pe dreaptă nu este măsurabilă Jordan, dar complementara sa este obţinută ca o reuniune numărabilă de intervale deschise. Deci complementara mulţimii lui Cantor se obţine ca o reuniune numărabilă de mulţimi măsurabile Jordan, dar ea nu este măsurabilă Jordan.

O $\sigma-$algebră este închisă la intersecţii numărabile. Într-adevăr, dacă $(A_n)_{n\in\mathbf{N}}$ este un şir de elemente din $\sigma-$algebra dată putem scrie

$$\cap_n A_n = (\cup_n A_n^c)^c,$$

relaţie ce are în membrul drept numai operaţii permise în interiorul unei $\sigma-$algebre.

Trebuie spus că noţiunea de algebră de părţi are o importanţă strict tehnică, utilă pentru a explica demonstraţiile. Adevărata noţiune importantă în teoria măsurii şi în probabilităţi este cea de σ-algebră.

Definiţia 1.4. *Fie* $(\Omega, \mathcal{F})$ *un spaţiu măsurabil. O aplicaţie* $P : \mathcal{F} \to [0,1]$ *se numeşte* măsură de probabilitate *dacă satisface condiţiile:*
(i) $P(\Omega) = 1,$
(ii) $P\left(\bigcup_{n\in\mathbf{N}} A_n\right) = \sum_n P(A_n)$, *pentru orice şir de mulţimi disjuncte,* $A_n \in \mathcal{F}, n \in \mathbf{N}$.
În acest caz tripletul $(\Omega, \mathcal{F}, P)$ *se numeşte* spaţiu probabilizat*.*

Relaţia (ii) din această definiţie este numită relaţia de $\sigma-$ aditivitate. Câteva proprietăţi legate de $\sigma-$aditivitate vor fi demonstrate mai târziu în secţiunea despre partiţii finite sau numărabile. Spaţiul probabilizat mai este cunoscut şi sub denumirea de *câmp de probabilitate.*

Studiile despre fundamentele logice ale noţiunii de probabilitate au o lungă istorie. La începutul secolului XX mai mulţi mari matematicieni depun eforturi pentru a formaliza matematic analiza logică a ideilor din teoria probabilităţilor care evoluaseră anterior, de peste două secole. Printre cei mai citaţi sunt matematicienii Serghei Bernstein, Emile Borel, Francesco Cantelli, Maurice Frechet, Paul Levy, Antoni Lomicki, Evgeny Slusky, Hugo Steinhaus, Richard von Mises (vezi „The Sources of Kolmogorov's Grundbegriffe", Statistical Science, 2006, vol. 21, No. 1). Sinteza acestor eforturi este consemnată de către matematicianul Andrei Nikolaevici *Kolmogorov*(1903 - 1987) într-o lucrare rămasă de referinţă (Grundbegriffe der Wahrscheinlichkeitsrechnung, Berlin, 1933). Rezultatul acestei formalizări revine la adoptarea spaţiului probabilizat $(\Omega, \mathcal{F}, P)$ drept cadrul standard pentru discutarea problemelor de teoria probabilităţilor. Azi acest obiect este definit şi studiat în cadrul teoriei măsurii. În teoria probabilităţilor el este preluat şi există obiceiul să se utilizeze, pe lângă terminologia de teoria măsurii, o terminologie proprie, colorată de sensuri apropriate modelelor probabiliste. Astfel o mulţime $A \in \mathcal{F}$ se numeşte eveniment iar un punct $\omega \in \Omega$ se numeşte eveniment elementar. Dacă $P(A^c) = 0$, se spune că evenimentul A are loc aproape sigur. Prescurtat scriem a.s. pentru „aproape sigur". Termenul „măsură de probabilitate" este adesea înlocuit prin prescurtarea „probabilitate", care trebuie folosit totuşi cu atenţie pentru a nu crea confuzie. În mod riguros probabilitatea reprezintă un număr ataşat unui eveniment - de exemplu $P(A)$ este probabilitatea evenimentului A.

1. Câteva exemple

Modelarea propriu-zisă este ceea ce se numeşte o artă. Nu există reguli şi cu greu se pot da reţete. Vom ilustra ideea de modelare prin exemple de-a lungul întregii cărţi. Începem cu următoarea secţiune în care prezentăm câteva din modelele tipice.

1.1. Zarul turtit. Un zar obişnuit este un cub, pe ale cărui feţe sunt marcate cu puncte numerele de la unu la şase. Să presupunem că avem un zar cu latura de 1 cm. El este modificat micşorându-i-se

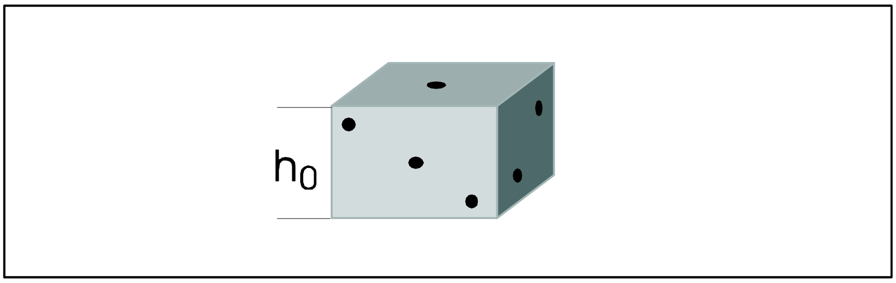


FIGURA 1. Un zar turtit

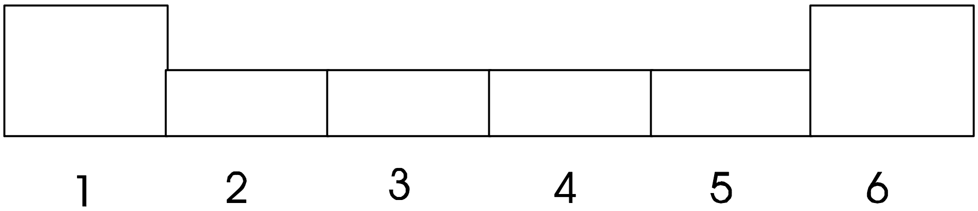


FIGURA 2. Un model probabilist pentru zarul turtit

înălţimea, astfel că devine un paralelipiped dreptunghic cu baza un pătrat cu latura de 1 cm iar înălţimea de $h \in (0,1)$.

Să presupunem că baza zarului este marcată cu şase puncte, iar faţa superioară cu unul. Când înălţimea h este foarte mică, feţele laterale devin înguste şi zarul nu poate sta decât cu greu pe ele; şi de aceea este natural să ne aşteptăm ca zarul, prin aruncare, să cadă mai ales cu feţele de sus şi jos: 1 şi 6. Atunci şansele de a cădea una din aceste feţe, se apropie pentru fiecare de $\frac{1}{2}$. Când $h = 1$ aceste feţe au şanse egale cu celelalte, deci egale cu $\frac{1}{6}$ pentru fiecare. Există o anumită valoare h_0 pentru care feţele cu un punct şi şase puncte au şansele $\frac{1}{4}$ fiecare. În acest caz, şansele feţelor cu 2,3,4 şi 5 puncte sunt egale cu $\frac{1}{8}$ fiecare. Experimentul de aruncare a zarului cu înălţimea h_0 poate fi descris de mulţimea Ω, desenată în figura 2.

Ea constă din şase dreptunghiuri (mulţimi deschise, fără laturile ce le mărginesc pentru a fi disjuncte). Primul dreptunghi şi ultimul sunt duble faţă de celelalte şi corespund la feţele cu 1 şi 6 puncte. Aria fiecărui dreptunghi este proporţională cu şansele de a cădea numărul de puncte pe care-l reprezintă. Dreptunghiurile ce formează mulţimea Ω sunt numerotate conform punctelor pe care le reprezintă, ca în figură: $D_1, D_2, D_3, D_4, D_5, D_6$. În acest caz $\mathcal{F}$ este algebra de părţi generată de aceste mulţimi. Mulţimea D_i corespunde evenimentului „în urma aruncării zarului a ieşit faţa cu i puncte". Probabilitatea este definită prin

$$P(A) = \frac{ariaA}{aria\Omega},$$

pentru orice $A \in \mathcal{F}$. De exemplu, evenimentul „la aruncarea zarului iese un număr de puncte mai mare sau egal cu 4” este reprezentat de mulţimea $D_4 \cup D_5 \cup D_6$ şi are probabilitatea

$$P\left(D_4 \cup D_5 \cup D_6\right) = \frac{ariaD_4 + ariaD_5 + ariaD_6}{aria\Omega} = \frac{1}{8} + \frac{1}{8} + \frac{2}{8} = \frac{1}{2}.$$

Pentru aceeaşi problemă se poate utiliza un model bazat pe mulţimea finită $\Omega' = \{1, 2, 3, 4, 5, 6\}$, cu $\mathcal{F} = \mathcal{P}(\Omega')$ şi punând

$$P(\{1\}) = P(\{6\}) = \tfrac{1}{4},$$
$$P(\{2\}) = P(\{3\}) = P(\{4\}) = P(\{5\}) = \tfrac{1}{8}.$$

1.2. Repartiţii uniforme. 1. Să presupunem că cineva aruncă la întâmplare o piatră într-o curte. Dacă nu am asistat la aruncare şi ne punem problema să găsim piatra, putem mai întâi să estimăm probabilitatea ca piatra să se afle într-o anumită regiune după aria regiunii. Să zicem că dreptunghiul Ω reprezintă curtea. Pentru o regiune $A \subset \Omega$, definim $P(A) = \frac{ariaA}{aria\Omega}$. În acest fel, am definit o măsură de probabilitate pe $\mathcal{B}(\Omega)$, care modelează şansele de a găsi piatra în diverse regiuni din interiorul curţii. Această măsură de probabilitate spunem că este repartizată *uniform* pe Ω, deoarece este proporţională cu aria. (Modelul se deosebeşte de cel de la zarul turtit prin aceea că, de data aceasta, măsura este definită pentru toate mulţimile boreliene. Clasa mulţimilor boreliene este infinită.)

2. Să presupunem că după o ploaie vrem să estimăm numărul de picături care au cazut în diverse regiuni din grădină. Să presupunem că s-a determinat numărul n_0 de picături ce au căzut într-un metru pătrat. Dacă grădina este un dreptunghi Ω, numărul de picături care au căzut în toată curtea este aproximativ $n_0 \cdot aria\Omega$. Pentru o regiune $A \subset \Omega$, numărul de picături va fi aproximativ $n_0 \cdot ariaA$. Dacă vrem să ştim ce proporţie din toate picăturile au căzut în regiunea A, aceasta este dată de

$$P(A) = \frac{ariaA}{aria\Omega}.$$

Din nou este o măsură de probabilitate pe $\mathcal{B}(\Omega)$, uniform distribuită pe Ω.

3. O roată de ruletă este învârtită şi apoi se aşteaptă oprirea ei. Care va fi probabilitatea ca indicatorul de la marginea roţii să se oprescă într-un anumit sector al cercului exterior? Răspunsul este dictat de o logică elementară: această probabilitate este proporţională cu lungimea curbilinie a sectorlui. Pentru a modela acest fenomen probabilist se notează cu Ω cercul exterior roţii de ruletă şi pentru o mulţime $A \in \mathcal{B}(\Omega)$ se pune

$$P(A) = \frac{\mathcal{L}(A)}{2\pi R},$$

unde $\mathcal{L}(A)$ este lungimea mulţimii A (sau măsura Lebesgue), iar R este raza cercului. Din nou P este o măsură de probabilitate pe $\mathcal{B}(\Omega)$, uniform repartizată pe cercul Ω.

Se ştie că un cerc de rază R este pus în corespondenţă bijectivă cu intervalul $[0, 2\pi R)$ prin desfăşurarea cercului. Lungimile măsurate pe cerc, după desfăşurare, corespund lungimilor măsurate pe segmentul $[0, 2\pi R)$. Atunci, un model echivalent se obţine punând $\Omega' = [0, 2\pi R)$, iar pentru o mulţime $A \in \mathcal{B}(\Omega')$ se pune

$$P'(A) = \frac{\mathcal{L}(A)}{2\pi R},$$

unde de data aceasta $\mathcal{L}(A)$ este măsura Lebesgue de pe segmentul $[0, 2\pi R)$. Măsura de probabilitate P' este uniform răspândită pe intervalul $[0, 2\pi R)$. Deoarece punctul din capătul intervalului are măsura nulă, din punct de vedere practic, acest model este echivalent cu cel obţinut pe intervalul $[0, 2\pi R]$ cu aceeaşi măsură P'.

4. La o fabrică de confecţii se utilizează şireturi. Şiretul este adus la fabrică în ghemuri şi resturile mai mici de 40 cm sunt inutilizabile. Prin urmare rămân deşeuri de şiret de toate mărimile mai mici de 40 de cm. Pentru a modela distribuirea resturilor de şiret este natural să considerăm $\Omega = [0, 40]$, $\mathcal{F} = \mathcal{B}(\Omega)$ şi pentru $A \in \mathcal{B}(\Omega)$

$$P(A) = \frac{\mathcal{L}(A)}{40}.$$

Dacă $A = (a, b)$, numărul $P((a, b))$ reprezintă proporţia acelor resturi de şiret care au lungimea mai mare decât a cm şi mai mică decât b cm, printre toate resturile de şiret. Suntem conduşi astfel la probabilitatea distribuită uniform pe intervalul $[0, 40]$.

1.3. Modele cu şanse egale. Cele mai vechi modele probabiliste consideră o mulţime $\Omega = \{\omega_1, ..., \omega_n\}$, finită, ca mulţime a posibilităţilor şi evaluează cu şanse egale fiecare posibilitate, ceea ce conduce la $P(\{\omega_i\}) = \frac{1}{n}$, pentru orice i. Exemplele la care se potrivesc modele de acest fel sunt cele mai uzuale: aruncarea monedei, unde şansele de a cădea în sus o faţă sau alta sunt evident egale (modelul are doar două posibilităţi şi deci $n = 2$), aruncarea cu zarul, unde fiecare din cele şase feţe are tot şanse egale de a ieşi (model cu $n = 6$), extragerea unei bile dintr-o urnă (n este numărul de bile din urnă) şi multe altele legate de diverse jocuri în care intervine norocul. Dacă $A \in \mathcal{P}(\Omega)$ este un eveniment arbitrar, probabilitatea sa este

$$P(A) = \frac{card A}{card \Omega}.$$

Această formulă spune că „probabilitatea evenimentului A este egală cu numărul cazurilor favorabile supra numărul cazurilor posibile”, care este definiţia clasică a probabilităţii. Pentru a decide că un fenomen probabilist este de tipul şanselor egale pentru fiecare posibilă evoluţie este însă necesară întotdeauna o analiză.

2. Aruncarea cu banul

Este cunoscut modelul dat de d'Alembert pentru aruncarea cu două monezi. El spunea că rezultatele posibile pentru un astfel de experiment sunt trei: ambele monezi cad cu cifra în sus, ambele monezi cad cu stema în sus, şi cea de a treia posibilitate, o monedă cade cu cifra în sus iar cealaltă cade cu stema în sus.

A făcut apoi greşeala să presupună că fiecare din aceste trei posibilităţi este la fel de probabilă. Dacă ar fi făcut experienţa, şi-ar fi dat seama foarte repede că nu este adevărată egalitatea şanselor pentru modelul ales. Dar nu numai experienţa, ci şi o analiză logică mai atentă relevă modelul corect. Pentru aceasta se consideră că cele două monezi sunt însemnate A şi B iar posibilitatea a treia din enumerarea dată mai sus este înlocuită cu două alte posibilităţi: A cade cu cifra şi B cade cu stema, respectiv A cade cu stema şi B cade cu cifra. Se poate uşor raţiona că dacă A cade cu cifra sunt şanse egale ca B să cadă cu cifra sau cu stema şi la fel, prin comparare, se ajunge imediat la concluzia că fiecare din cele patru posibilităţi descrise acum are şanse egale. Se vede atunci că, pentru a reflecta experimentul aruncării cu două monezi, modelul cu trei posibilităţi propus de d'Alembert nu poate da şanse egale fiecăreia. Pentru a corespunde realităţii, trebuiesc acordate şanse egale numai primelor posibilităţi şi anume $\frac{1}{4}$, în timp ce a treia posibilitate enumerată de d'Alembert ar trebui să aibă şanse duble faţă de fiecare din primele, deci $\frac{1}{2}$. Desigur că marele enciclopedist şi matematician francez Jean le Rond *d'Alembert* (1717-1783) a rămas în istorie pentru alte descoperiri importante, cazul descris mai sus nefiind decât un episod anecdotic.

În lumina discuţiei anterioare este clar că, în cazul în care considerăm o serie de n aruncări cu banul, modelul ce se impune este următorul: dacă 1 corespunde cifrei şi 0 corespunde stemei, mulţimea tuturor posibilităţilor este

$$\Omega = \{0, 1\}^n = \{(x_1, ..., x_n) \mid x_i \in \{0, 1\}, i = 1, ..., n\},$$

constituită din toate sistemele $(x_1, ..., x_n)$ de n cifre de 0 şi 1. Fiecare posibilitate are aceleaşi şanse de a ieşi. Numărul tuturor posibilităţilor este 2^n, adică $card\Omega = 2^n$. Probabilitatea corespunzătoare se defineşte prin $P(\{(x_1, ..., x_n)\}) = \frac{1}{2^n}$, pentru fiecare sistem $(x_1, ..., x_n) \in \Omega$.

Exemplu.

Doi jucători joacă aruncarea cu banul după următoarea regulă. La fiecare aruncare a monedei este pus în joc un punct. Dacă moneda cade cu stema în sus, câştigă jucătorul J_1, iar când cade faţa cu cifra, câştigă jucătorul J_2. Se aruncă moneda de mai multe ori la rând până când unul din jucători totalizează 10 puncte. În această situaţie respectivul jucător este considerat învingător şi câştigă o sumă. În cazul în care jocul este întrerupt înainte ca vreunul din jucători să iasă câştigător, se convine ca suma în joc să fie împărţită între cei doi, proporţional cu şansele pe care le-ar fi avut dacă jocul ar fi continuat. Ne propunem să calculăm cum va fi împărţită miza jocului dacă jocul este întrerupt în situaţia în care jucătorul J_1 are 7 puncte iar jucătorul J_2 are 8 puncte? (Soluţia acestei probleme o dă Blaise *Pascal* (1623-1662) într-o scrisoare, datată 24 august 1654, către Pierre de *Fermat* (1601-1665).)

Soluţie. Jocul se termină în cel mult 4 aruncări pentru că după 4 aruncări neapărat se întâmplă unul din următoarele evenimente: sau J_2 a câştigat cel puţin două puncte, sau J_1 a câştigat cel puţin trei. Cu convenţia că 0 reprezintă stema şi 1 cifra, putem reprezenta toate rezultatele posibile la o serie de 4 aruncări în felul următor:

$$\begin{array}{ccccc}
(0,0,0,0) & (0,0,0,1) & (0,0,1,1) & (0,1,1,1) & (1,1,1,1) \\
 & (0,0,1,0) & (0,1,0,1) & (1,0,1,1) & \\
 & (0,1,0,0) & (0,1,1,0) & (1,1,0,1) & \\
 & (1,0,0,0) & (1,0,0,1) & (1,1,1,0) & \\
 & & (1,0,1,0) & & \\
 & & (1,1,0,0) & &
\end{array}$$

Am listat aici 16 posibilităţi şi fiecare are şanse egale. (Vezi schema bilei întoarse, unde se arată egalitatea şanselor într-un caz mai general.) Primele 5 posibilităţi îl dau câştigător pe J_1 iar celelalte 11 pe J_2. Rezultă că şansele de a câştiga J_1 sunt $\frac{5}{16}$ şi şansele de a câştiga J_2 sunt $\frac{11}{16}$. Suma pusă în joc se va împărţi între cei doi jucători proporţional cu aceste valori.

Un învăţat al timpului i-a reproşat lui Pascal că de fapt jocul nu se termină neapărat cu încă 4 aruncări, ci el poate să se încheie după 2 sau 3 aruncări şi atunci posibilităţile ce se iau în considerare sunt de fapt următoarele:

$$\begin{array}{ccccc}
(0,0,0) & (0,0,1,0) & (0,0,1,1) & (0,1,1) & (1,1) \\
 & (0,1,0,0) & (0,1,0,1) & (1,0,1) & \\
 & (1,0,0,0) & (1,0,0,1) & &
\end{array}$$

Observaţia este corectă, însă în cazul acesta tebuie să acordăm altă pondere fiecărei posibilităţi:

$$P(\{(1,1)\}) = \frac{1}{4},\ P(\{(0,0,0)\}) = P(\{(0,1,1)\}) = P(\{(1,0,1)\}) = \frac{1}{8},$$
$$P(\{(0,0,1,0)\}) = P(\{(0,1,0,0)\}) = P(\{(1,0,0,0)\}) = P(\{(0,0,1,1)\})$$
$$= P(\{(0,1,0,1)\}) = P(\{(1,0,0,1)\}) = \frac{1}{16}.$$

Justificarea acestei ponderări se face pornind de la evaluarea şanselor pentru rezultatele a două aruncări:

$$P(\{(1,1)\}) = P(\{(1,0)\}) = P(\{(0,1)\}) = P(\{(0,0)\}) = \frac{1}{4}.$$

Rezultatul $(1,1)$ ar pune capăt jocului dar celelalte posibilităţi duc la continuarea jocului. Şansele pe care le are rezultatul $(1,0)$ se împart în mod egal pentru posibilităţile următoare: $(1,0,1), (1,0,0)$. Deci

$$P(\{(1,0,1)\}) = P(\{(1,0,0)\}) = \frac{1}{8}.$$

Rezultatul $(1,0,0)$ conduce la continuarea jocului cu încă o aruncare şi obţinerea posibilităţilor $(1,0,0,1)$ şi $(1,0,0,0)$, ce vor fi evaluate fiecare cu probabilitatea $\frac{1}{16}$. etc. □

3. Extrageri repetate din urnă

Din punct de vedere al teoriei probabilităţilor urna permite realizarea modelului tipic de experiment în care avem un număr finit de evenimente elementare cu şanse egale. Prin urnă înţelegem o cutie în care se află nişte obiecte asemănătoare, cum ar fi bile, sau bilete, etc. Când se fac extrageri din urnă întotdeauna se presupune că cel ce extrage nu poate alege şi că fiecare obiect din urnă are aceleaşi şanse de a fi extras.

3.1. Schema bilei întoarse. Într-o urnă se află m bile numerotate de la 1 la m. Se extrage din urnă o bilă, se notează numărul ei, după care se pune la loc bila în urnă. (O extragere de acest fel se numeşte extragere cu întoarcere sau cu revenire.) Repetând experienţa de n ori se obţine o secvenţă $(x_1, ..., x_n)$ formată din n numere. Dorim să modelăm probabilist experimentul global al extragerilor succesive de acest tip. Notăm $M = \{1, ..., m\}$ mulţimea ce reprezintă bilele. Mulţimea posibilităţilor este clar descrisă de produsul

$$\Omega = M^n = \{(x_1, ..., x_n) \mid x_i \in M, \forall i = 1, ..., n\}.$$

Să considerăm două secvenţe

$$\begin{gathered}(x_1, ..., x_{i-1}, x_i, x_{i+1}, ..., x_n), \\ (x_1, ..., x_{i-1}, x_i', x_{i+1}, ..., x_n)\end{gathered}$$

care au aceleaşi elemente, cu excepţia celor de pe poziţia i. Deci presupunem $x_i \neq x_i'$. Cele două secvenţe diferă doar prin rezultatul de la extracţia i. Dar la fiecare extracţie şansele de a extrage una sau alta din bile sunt egale. Concluzia este că cele două secvenţe au şanse egale de a se produce. Din aproape în aproape se deduce că toate secvenţele au aceeaşi şansă de a se produce. Numărul elementelor din Ω este m^n şi se impune $P(\{(x_1, ..., x_n)\}) = \frac{1}{m^n}$, pentru orice secvenţă.

3.2. Schema bilei neîntoarse. Aceeaşi urnă ca mai înainte este supusă unor n extrageri succesive, dar de data asta după fiecare extragere bila extrasă rămâne în afara urnei. (O astfel de extragere se numeşte fără întoarcere sau fără revenire.) Implicit trebuie să avem $n \leq m$. Mai remarcăm că secvenţa rezultatelor are toate componentele diferite. Mulţimea care modelează fenomenul este

$$\Omega = \{(x_1, ..., x_n) / x_i \in M,\ x_i \neq x_j, \forall i \neq j\}.$$

Un sistem $(x_1, ..., x_n)$ în care fiecare componentă este distinctă poate fi identificat cu mulţimea $\{x_1, ..., x_n\}$, care este ordonată prin indici. Se observă atunci că Ω poate fi identificată cu mulţimea submulţimilor de cardinal n din M ordonate în toate felurile posibile. Deci $card\Omega = A_m^n$. Vom arăta că, şi de această dată, trebuie să acceptăm ideea că fiecare posibilitate de desfăşurare a n extrageri succesive are aceleaşi şanse de apariţie.

Pentru aceasta vom face o inducţie după n. Notăm Ω_n şi respectiv Ω_{n+1} spaţiile tuturor posibilităţilor pentru experimentele corespunzătoare seriilor de n, respectiv $n+1$ extrageri. Presupunem că am stabilit că şansele tuturor secvenţelor de n extrageri sunt egale şi deci probabilitatea asociată unei astfel de secvenţe este egală cu $\frac{1}{A_m^n} = \frac{1}{m(m-1)...(m-n+1)}$. Fie $(x_1, ..., x_n) \in \Omega_n$. Această secvenţă poate fi continuată cu o nouă extragere din mulţimea $M \setminus \{x_1, ..., x_n\}$. Este o mulţime cu $m-n$ bile şi fiecare din ele are aceleaşi şanse de a fi extrasă. Notăm cu

$$A(x_1, ..., x_n) = \{(x_1, ..., x_n, x) / x \in M \setminus \{x_1, ..., x_n\}\},$$

mulţimea secvenţelor ce continuă secvenţa dată. Această mulţime are $m-n$ secvenţe şi trebuie să admitem că au şanse de a se produce egale între ele. Pe de altă parte, când secvenţa $(x_1, ..., x_n)$ parcurge Ω_n, mulţimile $A(x_1, ..., x_n)$ acoperă Ω_{n+1}, adică are loc egalitatea

$$\Omega_{n+1} = \bigcup_{(x_1,...,x_n)\in\Omega_n} A(x_1, ..., x_n).$$

Aceste mulţimi sunt disjuncte şi deci formează o partiţie a lui Ω_{n+1}. Este clar că obţinem o secvenţă de lungime $n+1$ care se află în mulţimea $A(x_1, ..., x_n)$ dacă şi numai dacă primele n extrageri sunt date de secvenţa $(x_1, ..., x_n)$. De aceea este natural să măsurăm producerea evenimentului $A(x_1, ..., x_n)$ în modelul cu secvenţe de lungime $n+1$, prin probabilitatea pe care o are evenimentul elementar $(x_1, ..., x_n)$ în modelul secvenţelor de lungime n :

$$P_{n+1}(A(x_1, ..., x_n)) = P_n(\{(x_1, ..., x_n)\}) = \frac{1}{A_m^n}.$$

Aceasta implică să punem aceeaşi valoare pentru probabilitatea fiecăruia din evenimentele $A(x_1, ..., x_n)$. Cum fiecare din aceste evenimente este compus din acelaşi număr de evenimente elementare ce au şanse egale între ele, rezultă că toate evenimentele elementare din Ω_{n+1} au aceleaşi şanse de a se produce:

$$P_{n+1}\left(\{(x_1,...,x_{n+1})\}\right)=\frac{1}{A_m^{n+1}}=\frac{1}{m\,(m-1)\,...\,(m-n)}.$$

3.3. Extrageri fără întoarcere fără ordine. Un caz particular de extrageri repetate fără întoarcere sunt extragerile în care ceea ce contează la sfârşit este numai mulţimea bilelor extrase. Deci nu mai reţinem ordinea în care s-au făcut extragerile şi privim drept rezultat al extragerii numai mulţimea bilelor extrase. Ţinând cont de simetria experimentului a n extrageri succesive fără întoarcere este normal să ne aşteptăm ca fiecare mulţime de n elemente să apară cu aceeaşi probabilitate. Vom demonstra acest lucru bazându-ne pe modelul deja construit.

Propoziţia 1.1. *Fie M mulţimea cu m elemente care reprezintă bilele din urnă şi $n\le m$. Notăm cu Ω mulţimea ce modelează seriile de n extrageri fără întoarcere, ca mai sus. Pentru fiecare submulţime $A\subset M$ având cardinalul n are loc formula*

$$P\left(\{(x_1,...,x_n)\in\Omega/\left\{x_1,...,x_n\right\}=A\}\right)=\frac{1}{C_m^n}.$$

Demonstraţie. Mulţimea $\{(x_1,...,x_n)\in\Omega/\left\{x_1,...,x_n\right\}=A\}$ constă din toate permutările posibile cu elementele din A. Cardinalul acestei mulţimi este $n!$ şi, cum probabilitatea unei secvenţe din Ω este $\frac{1}{A_m^n}$, rezultă că

$$P\left(\{(x_1,...,x_n)\in\Omega/\left\{x_1,...,x_n\right\}=A\}\right)=\frac{n!}{A_m^n}=\frac{1}{C_m^n}.\ \square$$

Să examinăm acum experimentul extragerii a n bile deodată din urna cu m bile. Pentru acest experiment mulţimea tuturor posibilităţilor este mulţimea ce o notăm cu Γ, şi care se compune din părţile lui M cu n elemente. Din motive de simetrie este natural să considerăm că fiecare parte a lui M cu n elemente are aceleaşi şanse de a fi extrasă. Rezultă ca naturală alegerea probabilităţii P' pe Γ care dă aceeaşi valoare fiecărui eveniment elementar. Cardinalul lui Γ fiind C_m^n, deducem că probabilitatea fiecărui eveniment elementar din Γ este $\frac{1}{C_m^n}$.

Să comparăm acum modelul nou construit (Γ,P') cu cel anterior (Ω,P). Pentru un eveniment elementar $A\in\Gamma$ considerăm mulţimea tuturor secvenţelor $(x_1,...,x_n)\in\Omega$ care sunt formate cu elementele mulţimii A (adică $\{x_1,...,x_n\}=A$). Notăm Λ_A această mulţime:

$$\Lambda_A=\{(x_1,...,x_n)\in\Omega\ /\ \ \{x_1,...,x_n\}=A\}\,.$$

Ea constituie un eveniment neelementar din Ω. Ansamblul acestor mulţimi $\{\Lambda_A\ /\ A\in\Gamma\}$ formează o partiţie a lui Ω care evident este în bijecţie cu Γ. Din propoziţia anterioară rezultă că $P\left(\Lambda_A\right)=P'\left(\{A\}\right).$

În concluzie, putem spune că experimentul extragerii din urnă a n bile deodată este echivalent cu experimentul a n extrageri succesive fără revenire în care se uită ordinea extragerilor şi se ia în considerare numai mulţimea rezultată din extragere. Modelarea se poate face fie pe (Ω, P) considerând evenimentele Λ_A fie pe (Γ, P').

3.4. Extrageri cu întoarcere fără ordine*. Să presupunem că dintr-o urnă cu m bile numerotate se fac extrageri cu revenire. Fie $M = \{1, 2, ..., m\}$ mulţimea ce reprezintă bilele din urnă. Să ne închipuim că rezultatele unei serii de n extrageri se notează în linie consemnându-se pe rând fiecare număr extras, iar la finalul seriei se completează m căsuţe în care se menţionează numărul de apariţii ale fiecărei bile. Aşadar, în căsuţa i se specifică de câte ori apare în lina respectivă numărul i, adică de câte ori a fost extrasă bila cu numărul i. Putem spune că datele conţinute în căsuţe rezumă rezultatele unei serii de extrageri în care nu mai interesează ordinea în care au apărut bilele. Problema pe care ne-o punem este de a afla care este proporţia seriilor pentru care cele m căsuţe de la sfârşit au un conţinut identic, unul anume, în ansamblul tuturor seriilor. Mai precis, putem formula problema astfel: fiind date numerele naturale $k_1, ..., k_m$ astfel încât $\sum_{i=1}^{m} k_i = n$, să se determine probabilitatea ca într-o serie de n extrageri să obţinem exact k_i apariţii pentru bila cu numărul i, pentru fiecare $i = 1, ..., m$.

Pentru aceasta facem socotelile pe modelul $\Omega = M^n$, corespunzător seriilor de extrageri cu întoarcere. Avem de determinat numărul elementelor $(x_1, ..., x_n) \in \Omega$ cu proprietatea că, pentru fiecare $i = 1, ..., m$, cardinalul mulţimii $\{1 \leq l \leq n / \ x_l = i\}$ este k_i. Vom nota mulţimea acestor elemente astfel:

$$\Lambda (k_1, ..., k_m) = \{(x_1, ..., x_n) \in \Omega / \ card \{1 \leq l \leq n / \ x_l = i\} = k_i\}.$$

Este vorba de o problemă de combinatorică cunoscută, care este formulată deobicei într-un cadru de teoria mulţimilor, făcând abstracţie de urnă şi bile.

Pentru a exprima problema în cadrul combinatoric tipic, vom nota cu $L = \{1, ..., n\}$ mulţimea locurilor din linia pe care sunt consemnate cele n extrageri dintr-o serie. Un element $(x_1, ..., x_n) \in \Omega$ poate fi identificat cu o partiţie a mulţimii L de tipul $(A_1, ..., A_m)$, unde $A_i = \{1 \leq l \leq n / \ x_l = i\}$. Cu alte cuvinte, A_i este mulţimea poziţiilor din L care sunt ocupate de numărul i (şi care corespund extragerilor la care a apărut bila cu numărul i). Atragem atenţia că pentru a avea o bună descriere a elementelor lui Ω este necesar să considerăm partiţii ordonate, descrise ca sisteme ordonate $(A_1, ..., A_m)$. Spre exemplu, având o astfel de partiţie, se pot obţine alte partiţii ordonate distincte prin permutarea mulţimilor. Astfel $(A_2, A_1, A_3, ..., A_m)$, sistemul în care am inversat locurile mulţimilor A_1 şi A_2, este distinct de cel anterior, în timp ce privite ca partiţii ale lui L ele sunt identice.

(În general, prin partiţie a unei mulţimi se înţelege o familie de părţi disjuncte a căror reuniune acoperă mulţimea dată, fără a exista vreo ordine pe familia de părţi.) La noi, ordinea în care apar mulţimile unui sistem indică numărul de bilă ce corespunde fiecărei mulţimi.

Cu această nouă formulare pentru problema noastră avem de determinat câte sisteme de mulţimi $(A_1, ..., A_m)$ verifică următoarele proprietăţi: 1^0 mulţimile $A_1, ..., A_m$ formează o partiţie a lui L şi 2^0 cardinalul mulţimii A_i este k_i, pentru orice $i = 1, ..., m$. Propoziţia următoare răspunde la această cerinţă.

Propoziţia 1.2. *Fiind dată o mulţime L cu n elemente şi fiind date numerele $k_1, ..., k_m$, astfel încât $n = k_1 + ... + k_m$, se pot forma*

$$\frac{n!}{k_1! \cdot ... \cdot k_m!}$$

partiţii ordonate $(A_1, ..., A_m)$ ale mulţimii L, astfel încât cardinalul mulţimii A_i să fie $k_i, i = 1, ..., m$.

Demonstraţie. Numărarea o vom face raţionând prin inducţie după m. Să presupunem că avem o mulţime $A \subset L$, de cardinal k_m, fixată. Pentru a obţine toate partiţiile ordonate ale lui L, de tipul căutat, care au drept ultimă mulţime A, trebuie să vedem în câte feluri putem face operaţia similară pentru $L\backslash A$ în loc de L, dar luând în considerare partiţiile ordonate cu $m-1$ părţi. Mai precis, căutăm numărul partiţilor ordonate $(B_1, ..., B_{m-1})$ ale lui $L\backslash A$ astfel încât cardinalul mulţimii B_i să fie k_i, pentru orice $i = 1, ..., m-1$. Conform inducţiei, numărul acestor partiţii este $\frac{(n-k_m)!}{k_1!...k_{m-1}!}$. Cum $(A, B_1, ..., B_{m-1})$ reprezintă o partiţie ordonată ca în enunţul propoziţiei şi cum toate partiţiile căutate se obţin în acest fel, rezultă că avem de calculat numărul de combinaţii de acest tip posibile. Pe de altă parte, numărul părţilor $A \subset L$ de cardinal k_m este egal cu $C_n^{k_m}$. Rezultă că numărul căutat este $C_n^{k_m} \frac{(n-k_m)!}{k_1!...k_{m-1}!} = \frac{n!}{k_1! \cdot ... \cdot k_m!}$. $\square$

Numărul $\frac{n!}{k_1! \cdot ... \cdot k_m!}$ este numit *coeficient multinomial.* Pentru $n = 2$ el devine coeficientul binomial $C_n^{k_1} = C_n^{k_2} = \frac{n!}{k_1!(n-k_1)!}$.

Revenind la problema iniţială, putem spune că mulţimea $\Lambda(k_1, ..., k_m)$ are cardinalul $\frac{n!}{k_1! \cdot ... \cdot k_m!}$ şi, prin urmare, avem

$$P(\Lambda(k_1, ..., k_m)) = \frac{n!}{k_1! \cdot ... \cdot k_m!} \frac{1}{m^n}.$$

Pe de altă parte, mulţimea Ω se scrie sub forma

$$\Omega = \bigcup_{(k_1, ..., k_m)} \Lambda(k_1, ..., k_m),$$

unde se face reuniunea după toate sistemele $(k_1, ..., k_m) \in \mathbf{N}^m$, astfel încât $k_1 + ... + k_m = n$. Cum mulţimile care participă la această reuniune sunt şi disjuncte, rezultă că are loc relaţia

$$\sum_{(k_1,...,k_m)} \frac{n!}{k_1! \cdot ... \cdot k_m!} = m^n.$$

O problemă de alocaţie.

Examinăm acum problema anterioară sub altă înfăţişare, sub care este adesea întâlnită.

Vom presupune că M este o mulţime formată din n bile numerotate de la 1 la n. Avem m cutii şi numerele $k_1, ..., k_m$ date astfel încât $k_1 + ... + k_m = n$. Se iau pe rând bilele mulţimii M şi se pun aleator în câte o cutie. Se pune apoi problema să se determine în câte feluri se pot aloca aleator cele n bile în cutii astfel încât întâia cutie să conţină k_1 bile, a doua să conţină k_2 bile, etc. Dacă notăm cu A_i mulţimea bilelor ce sunt puse în cutia i, atunci $A_1, ..., A_m$ formează o partiţie a lui M. Formulată altfel, problema revine la a determina numărul partiţiilor ordonate ale mulţimii $M = \{1, ..., n\}$ în forma $(A_1, ..., A_m)$ aşa încât cardinalul mulţimii A_i să fie $k_i, i = 1, ..., m$. Este, deci, vorba despre situaţia descrisă în lema anterioară şi numărul căutat este $\frac{n!}{k_1! \cdot ... \cdot k_m!}$.

Numărul rezultatelor posibile.

Ne punem acum problema de a determina numărul rezultatelor posibile distincte ce se consemnează în ansamblul căsuţelor de la sfârşitul unei linii ce corespunde unei serii de n aruncări cu întoarcere, când nu se mai ţine cont de ordine. Acest număr este determinat de următorul rezultat de combinatorică.

Propoziţia 1.3. *Fiind date $n, m \in \mathbf{N}$, numărul sistemelor de numere naturale $(k_1, ..., k_m) \in \mathbf{N}^m$ cu proprietatea că $n = k_1 + ... + k_m$, este egal cu $C_{m+n-1}^{m-1} = C_{m+n-1}^{n}$.*

Demonstraţie. Pentru a explica într-un mod mai intuitiv demonstraţia revenim la modelul seriilor de n extrageri cu întoarcere dintr-o urnă cu m bile numerotate de la 1 la m. Şi anume, examinăm rezultatele consemnate în căsuţele de la sfârşitul liniei de însemnări corespunzătoare unei serii de n extrageri schimbând modalitatea de consemnare a rezultatului. Anume, în fiecare căsuţă în locul numărului ce reprezintă apariţiile respectivei bile punem semnul X repetat de atâtea ori câte apariţii a avut bila. Astfel rezultatul poate fi exprimat şi sub forma unui şir de semne X care marchează în căsuţa respectivă fiecare apariţie, despărţite de bare de tipul $|$, ce delimitează căsuţele. Pentru exemplificare, în şirul de mai jos avem $m = 4$ şi $n = 5$ şi sunt marcate următoarele rezultate: pentru bila cu numărul 1 au fost bifate 2 apariţii, pentru bila cu numărul 2 a fost bifată o singură apariţie, pentru bila cu numărul 3 nu este nici o apariţie, iar pentru bila cu numărul 4 sunt bifate 2 apariţii:

$$XX \;|\; X \,|\; \;|\; XX$$

La fel, în şirul

$$| \quad | \quad X \,| \quad XXXX$$

se consemnează rezultatele unei serii de $n = 5$ extrageri din urna cu $m = 4$ bile, la care bilele 1 şi 2 nu au apărut, bila 3 a ieşit o dată, iar bila 4 a ieşit de patru ori.

Rezultatul unei serii de aruncări în care ne interesează doar numătul de apariţii ale fiecărei bile este deci descris complet de succesiunea a $m - 1$ bare şi n semne X. Dar, privind mulţimea locurilor pe care au fost trasate bare verticale sau semne X, ca o mulţime fixă cu $m + n - 1$ elemente, atunci observăm că o succesiune ce descrie o serie de aruncări este complet determinată de mulţimea poziţiilor barelor. Există C_{m+n-1}^{m-1} posibilităţi de a alege în mod diferit poziţiile barelor, ceea ce încheie demonsraţia. □

Altă demonstraţie se poate face grupând sistemele după valorile lui $k_1 = 0, 1, ..., n$ şi utilizând metoda inducţiei matematice pentru a număra sistemele din fiecare grup. În final se utilizează relaţia cunoscută

$$C_m^l = C_{m-1}^{l-1} + C_{m-2}^{l-1} + ... + C_{l-1}^{l-1}.$$

Putem trage concluzia, pentru problema seriilor de n extrageri cu întoarcere fără ordine, că numărul mulţimilor de tip $\Lambda(k_1, ..., k_m)$, care de fapt reprezintă evenimentele elementare ale experimentului, este de C_{m+n-1}^{m-1}. Probabilitatea fiecărui eveniment elementar este diferită, aşa cum am stabilit mai sus.

4. Exemple

4.1. Controlul calităţii. Exemplu.

Într-o ladă se află 1500 de mere, din care un anumit procent sunt putrede. Fără a şti acest procent, se ia un eşantion de 40 de mere şi se taie pentru a se constata starea lor şi a evalua care ar fi procentul de mere putrede în toată lada. Este o problemă tipică de controlul calităţii. Aici ne propunem să determinăm care ar fi probabilitatea ca eşantionul să conţină x mere putrede dacă am şti că 2% din numărul total de mere sunt putrede.

Soluţie. Presupunerea făcută implică faptul că 30 de mere sunt putrede. În ce priveşte extragerea celor 40 de mere, ea nu contează decât ca mulţime. Suntem în cazul extragerilor fără ordine. Fie M mulţimea tuturor merelor şi R mulţimea tuturor merelor putrede. Avem $card\,(M) = 1500$ şi $card\,(R) = 30$. Pentru un eşantion arbitrar de 40 mere putem avea $0, 1, ...,$ până la 30 mere putrede. Vrem să determinăm care este probabilitatea fiecărui caz. Spaţiul posibilităţilor pe care modelăm problema este

$$\Omega = \{A \in \mathcal{P}(M) / cardA = 40\},$$

iar evenimentele a căror probabilitate o căutăm vor avea expresia

$$\Lambda_l = \{A \in \Omega \ / \ card\,(A \cap R) = l\}, l = 0, 1, ..., 30.$$

Este clar că evenimentele acestea formează o partiţie lui Ω. Cardinalul mulţimii Ω este C_{1500}^{40}. Pentru a determina cardinalul mulţimii Λ_l vom raţiona astfel: orice mulţime $A \in \Lambda_l$ se împarte în mod natural în două părţi, $A_1 = A \cap R$ de cardinal $cardA_1 = l$ şi $A_2 = A \cap R^c$ cu $cardA_2 = 40 - l$. Pentru A_1 există C_{30}^l posibilităţi de alcătuire, iar pentru A_2 există C_{1470}^{40-l} posibilităţi de alcătuire. Deci, mulţimea ce ne interesează, Λ_l, are $C_{30}^l \cdot C_{1470}^{40-l}$ elemente. Obţinem formula

$$P(\Lambda_l) = \frac{C_{30}^l \cdot C_{1470}^{40-l}}{C_{1500}^{40}}.$$

Notând $p_l = P(\Lambda_l)$, trecem la calculul numeric al acestor probabilităţi utilizând doar un minicalculator. Începem cu

$$\begin{aligned} p_0 &= \frac{C_{1470}^{40}}{C_{1500}^{40}} = \frac{1470 \times ... \times 1431}{1500 \times ... \times 1461} = \frac{1460 \times ... \times 1431}{1500 \times ... \times 1471} = \\ &= \left(1 - \frac{40}{1500}\right) ... \left(1 - \frac{40}{1471}\right). \end{aligned}$$

Avem un produs de 30 de numere apropiate de unitate şi putem face o aproximare bazându-ne pe *lema*1.1, punctul (iii) de mai jos:

$$\approx \exp - \sum_{i=1}^{30} \frac{40}{1470 + i}.$$

Eroarea acestei aproximări este pozitivă şi majorată de $\frac{7}{12} \sum_{i=1}^{30} \left(\frac{40}{1470+i}\right)^2$, evident mai mică decât $0,02$. Mai departe, se observă că ultimul rezultat se poate aproxima, cu o nouă eroare mică, negativă de data aceasta, prin

$$\exp - \sum_{i=1}^{30} \frac{40}{1470 + i} \approx \exp - \frac{30 \times 40}{1470} = 0,442.$$

Deci avem o valoare aproximativă a lui $p_0 \approx 0,442$. Mai departe, calculăm o relaţie de recurenţă:

$$\begin{aligned} p_{l+1} &= \frac{C_{30}^{l+1} \cdot C_{1470}^{40-l-1}}{C_{1500}^{40}} = \frac{1}{C_{1500}^{40}} \times \frac{30!}{(l+1)!\,(30-l-1)!} \times \\ &\times \frac{1470!}{(40-l-1)!\,(1430+l+1)!} = \frac{30-l}{l+1} \times \frac{40-l}{1430+l+1} \times p_l. \end{aligned}$$

Deci $p_1 = \frac{30}{1} \times \frac{40}{1431} \times p_0 \approx 0,370$ şi apoi $p_2 = \frac{29}{2} \times \frac{39}{1432} \times p_1 \approx 0,146$ iar $p_3 = \frac{28}{3} \times \frac{38}{1433} \times p_2 \approx 0,036$ şi $p_4 = \frac{27}{4} \times \frac{37}{1434} \times p_3 \approx 0,006$.

Listăm în tabelul următor valorile numerice ce se obţin prin calcul pe un PC:

p_0	p_1	p_2	p_3	p_4
$0,4409$	$0,3697$	$0,1460$	$0,0361$	$0,0062$

După cum se vede, calculele noastre aproximative sunt destul de apropiate de valorile reale. Adunând probabilităţile din tabel avem $p_0 + p_1 +$

$p_2 + p_3 + p_4 = 0,9989$, ceea ce arată că restul probabilităţilor sunt neglijabile (bineînţeles că are loc relaţia $\sum_{l=0}^{30} p_l = 1$). Cu tabelul în faţă putem să facem următorul comentariu. Sub ipoteza că numai 2% dintre mere sunt putrede, probabilitatea covârşitoare o are evenimentul „numărul de mere putrede care ies la o extragere de 40 este cel mult doi, adică $l \leq 2$”, şi anume $p_0 + p_1 + p_2 = 0,9566$. Evenimentul „numărul de mere putrede ce ies la o extragere de 40 este 3 sau mai mare” este foarte rar, pentru că probabilitatea sa este $0,0434$. Faptul de a găsi 3 mere putrede printre 40 testate nu este normal. O persoană interesată, care în urma testului ar găsi 3 mere putrede, ar trebui să tragă concluzia că este foarte verosimil ca ipoteza cum că „procentul merelor putrede din ladă este de numai 2%” să nu fie reală. □

Produsul mai multor numere apropiate de unitate.*

În teoria probabilităţilor apare de multe ori problema evaluării unor produse cu mulţi factori cu valori apropiate de 1. De exemplu probabilitatea intersecţiei mai multor evenimente independente conduce la un astfel de produs. Pentru aceasta se utilizează aproximarea $1-x \approx e^{-x}$, valabilă pentru x pozitiv, mic. Această aproximare se dovedeşte în multe situaţii atât de bună că permite şi aproximarea repetată în acelaşi fel a factorilor unui produs. Formalizăm rezultatul în următoarea lemă. Vom avea prilejul de a utiliza această lemă în mai multe ocazii.

Lema 1.1. *(i) Au loc următoarele inegalităţile*

$$e^{-x-\frac{7}{12}x^2} \leq 1 - x \leq e^{-x-\frac{1}{2}x^2},$$

unde pentru prima este necesar ca $0 \leq x \leq \frac{1}{5}$, iar pentru a doua este suficient ca $0 \leq x \leq 1$.

(ii) Dacă $0 \leq x \leq 1$, atunci sunt verificate inegalităţile

$$1 - x \leq e^{-x} \leq 1 - x + \frac{1}{2}x^2.$$

(iii) Dacă $0 \leq \alpha_1, ..., \alpha_n \leq \frac{1}{5}$, atunci au loc estimările

$$\exp\left(-\sum_{k=1}^{n} \alpha_k\right) \exp\left(-\frac{7}{12}\sum_{k=1}^{n} \alpha_k^2\right) \leq \prod_{k=1}^{n} (1 - \alpha_k) \leq \exp\left(-\sum_{k=1}^{n} \alpha_k\right),$$

$$\exp\left(-\sum_{k=1}^{n} \alpha_k\right) - \prod_{k=1}^{n} (1 - \alpha_k) \leq \frac{7}{12}\left(\sum_{k=1}^{n} \alpha_k^2\right) \exp\left(-\sum_{k=1}^{n} \alpha_k\right).$$

Pentru partea dreaptă a primei estimări este suficient ca $0 \leq \alpha_1, ..., \alpha_n \leq 1$.

Demonstraţie. (i) Se porneşte de la scrierea $1-x = \exp(\ln(1-x))$ după care se utilizează dezvoltarea în serie de puteri a funcţiei $\ln(1-x)$ pentru $|x| < 1$:

$$\ln(1 - x) = -x - \frac{1}{2}x^2 - \frac{1}{3}x^3 - \frac{1}{4}x^4 - ...$$

În urma neglijării termenilor de la al treilea încolo avem $\ln(1-x) \leq -x - \frac{1}{2}x^2$, ceea ce conduce imediat la partea dreaptă a estimării din enunţ.

Pentru a deduce partea stângă a estimării vom majora restul seriei de la al treilea termen prin seria geometrică astfel:

$$\frac{1}{3}x^3 + \frac{1}{4}x^4 + ... < \frac{1}{3}x^3\left(1 + x + x^2 + ...\right) = \frac{1}{3}\frac{x^3}{1-x},$$

unde am considerat că $0 < x < 1$. Pentru $0 \leq x \leq \frac{1}{5}$ avem inegalitatea $\frac{1}{1-x} \leq \frac{5}{4}$, ceea ce conduce la

$$\frac{1}{3}x^3 + \frac{1}{4}x^4 + ... \leq \frac{1}{12}x^2.$$

De aici rezultă imediat că $\ln(1-x) \geq -x - \frac{7}{12}x^2$, ceea ce conduce la inegalitatea din enunţ.

(ii) Pentru estimările de la acest punct se utilizează dezvoltarea în serie a funcţiei exponenţiale,

$$e^{-x} = 1 - x + \frac{1}{2}x^2 - \frac{1}{3!}x^3 + ...,$$

care este o serie alternantă, pentru $0 < x < 1$, cu termenii descrescători în valoare absolută.

(iii) Primul rând de estimări de la acest punct se obţin prin aplicarea punctului (i) pentru fiecare factor al produsului. Pentru a deduce ultima inegalitate se utilizează primul rând de estimări, obţinând

$$\exp\left(-\sum_{k=1}^{n}\alpha_k\right) - \prod_{k=1}^{n}(1-\alpha_k) \leq \left(1 - \exp\frac{7}{12}\left(\sum_{k=1}^{n}\alpha_k^2\right)\right)\exp\left(-\sum_{k=1}^{n}\alpha_k\right),$$

după care se utilizează inegalitatea $1 - e^x \leq x$. $\square$

Remarcăm că pentru $0 \leq x \leq \frac{1}{10}$ estimările de la punctul (i) ne asigură că aproximarea $1 - x \approx e^{-x}$ este satisfăcătoare. În calculele concrete ale unor probabilităţi de multe ori valorile mai mici ca $0,01$ se neglijează. Punctul (ii) este util pentru a examina estimarea în sens invers $e^{-x} \approx 1 - x$. Estimările de la punctul (iii) sunt utile în măsura în care valorile α_k sunt atât de mici încât chiar cumulate pătratele dau suma $\sum_{k=1}^{n}\alpha_k^2$ cu o valoare mică.

Sigur că pentru multe situaţii concrete un PC permite calcule mai precise decât cele bazate pe inegalităţile anterioare. Totuşi sunt situaţii când vrem să facem estimări lucrând cu variabile, adică în loc de numere precise avem litere despre care ştim că desemnează valori ce sunt de un anumit ordin. Atunci sunt utile aceste inegalităţi.

Identitatea lui Van der Monde.

Calculul din paragraful antepreceedent a pus în evidenţă o identitate combinatorică care merită reţinută în forma generală următoare. Să presupunem că mulţimea M constă din bile albe şi roşii şi are cardinalul m. Submulţimea bilelor roşii o notăm R şi notăm cu r cardinalul său.

Numărul de submulţimi ale lui M care au cardinalul un număr dat $n \leq m$ este C_m^n. Aceste submulţimi pot fi grupate după numărul l de elemente din R pe care le conţin. Numărul submulţimilor de n bile dintre care l sunt roşii este $C_r^l C_{m-r}^{n-l}$ şi deducem identitatea

$$C_m^n = \sum_l C_r^l C_{m-r}^{n-l},$$

care este cunoscută sub numele de identitatea lui Van der Monde. Sumarea se face după indicele l care trebuie să satisfacă relaţiile $0 \leq l \leq n \wedge r$ şi $n - l \leq m - r$. (A doua relaţie spune că pentru o submulţime a lui M care are n elemente, numărul elementelor din mulţimea $M \backslash R$ nu depăşeşte cardinalul acestei mulţimi.)

Putem calcula probabilitatea ca extrăgând o submulţime de cardinal n din M să obţinem l elemente din R. Această probabilitate este $p_l = \frac{C_r^l C_{m-r}^{n-l}}{C_m^n}$. Conform cu condiţiile menţionate mai sus, l trebuie să satisfacă relaţiile $0 \vee (n - m + r) \leq l \leq n \wedge r$.

Vom face acum o verificare algebrică pentru identitatea lui Van der Monde, pe care o şi enunţăm mai precis sub forma: dacă numerele naturale m, r, n satisfac condiţiile $0 \leq r, n \leq m$, atunci are loc relaţia

$$C_m^n = \sum_{l=0 \vee (n-m+r)}^{n \wedge r} C_r^l C_{m-r}^{n-l}.$$

<u>*Soluţie.*</u> Vom presupune că $1 \leq r, n \leq m - 1$, pentru că altfel identitatea este trivială. Facem o inducţie după m şi utilizăm identitatea elementară $C_m^k = C_{m-1}^k + C_{m-1}^{k-1}$, valabilă pentru numerele naturale $1 \leq k < m$. Dacă $r < n < m - r$, putem scrie

$$\sum_{l=0 \vee (n-m+r)}^{n \wedge r} C_r^l C_{m-r}^{n-l} = \sum_{l=0}^{r} C_r^l C_{m-r}^{n-l} = \sum_{l=0}^{r} C_r^l C_{m-1-r}^{n-l} + \sum_{l=0}^{r} C_r^l C_{m-1-r}^{n-1-l}$$

şi, aplicând formula presupusă valabilă pentru $m - 1$, rezultă

$$= C_{m-1}^n + C_{m-1}^{n-1} = C_m^n.$$

Pentru cazul $n \leq r, n < m - r$, calculul este asemănător:

$$\sum_{l=0 \vee (n-m+r)}^{n \wedge r} C_r^l C_{m-r}^{n-l} = \sum_{l=0}^{n} C_r^l C_{m-r}^{n-l} = \sum_{l=0}^{n-1} C_r^l C_{m-r}^{n-l} + C_r^n =$$

$$\sum_{l=0}^{n-1} C_r^l C_{m-1-r}^{n-l} + C_r^n + \sum_{l=0}^{n-1} C_r^l C_{m-1-r}^{n-1-l} = C_{m-1}^n + C_{m-1}^{n-1} = C_m^n.$$

Cazul $m - r \leq n, r < n$ este şi el similar:

$$\sum_{l=0 \vee (n-m+r)}^{r} C_r^l C_{m-r}^{n-l} = \sum_{l=n-m+r}^{r} C_r^l C_{m-r}^{n-l} =$$

$$= \sum_{l=n-m+r+1}^{r} C_r^l C_{m-r}^{n-l} + C_r^{n-m+r} = \sum_{l=n-m+r+1}^{r} C_r^l C_{m-1-r}^{n-l} +$$

$$+ \sum_{l=n-m+r+1}^{r} C_r^l C_{m-1-r}^{n-1-l} + C_r^{n-m+r} = C_{m-1}^{n} + C_{m-1}^{n-1} = C_m^n.$$

Rămâne de verificat cazul $m-r \leq n \leq r$, pentru care avem un calcul ce combină situaţiile ultimelor două cazuri anterioare. Lăsăm cititorului grija acestei verificări.

Observăm că verificarea algebrică este mai greoaie. □

4.2. Problema potrivirilor. Exemplul 1.

O persoană pretinde a avea simţuri paranormale care îi permit să poată simţi culorile fără să vadă respectivele obiecte colorate. Pentru a se dovedi aceste calităţi, se face următorul experiment. Se iau 4 cartoane colorate, fiecare de altă culoare, şi se pun într-o pălărie. Persoana respectivă trece în camera alăturată şi altcineva extrage pe rând cartoanele din pălărie. Din camera alăturată persoana cu pretinsele calităţi paranormale trebuie să spună în ordine ce culori au fost extrase. Care este probabilitatea ca un singur răspuns să fie exact, dacă răspunsurile se dau prin hazard? Care este probabilitatea ca numai două răspunsuri să fie exacte, dacă ele se dau la întâmplare? Care este probabilitatea ca toate răspunsurile să fie exacte, când ele se dau prin hazard?

Soluţie. Experimentul extragerii celor 4 cartoane din pălărie este complet analog extragerii fără întoarcere a 4 bile dintr-o urnă ce are iniţial exact 4 bile. Iar dacă răspunsurile sunt date la întâmplare, aceasta revine de fapt la a repeta extragerea fără întoarcere cu cele 4 bile. Să notăm c_1, c_2, c_3, c_4, culorile extrase din pălărie. Fiecare serie de alte 4 extrageri corespunde unei permutări a acestor patru culori şi toate au şanse egale. În total sunt $4! = 24$ de posibilităţi şi probabilitatea ca să fie date toate răspunsurile corect este probabilitatea permutării identice (c_1, c_2, c_3, c_4), ce este egală cu $\frac{1}{24}$.

Să stabilim care este probabilitatea ca un singur răspuns să fie exact. Pentru aceasta vom analiza mai întâi cazul acelori permutări care lasă numai culoarea c_1 pe loc. Acestea sunt descrise de poziţiile celorlalte culori, care pot fi listate astfel

$$c_2, c_3, c_4 \qquad c_2, c_4, c_3 \qquad c_3, c_2, c_4 \qquad c_3, c_4, c_2 \qquad c_4, c_2, c_3 \qquad c_4, c_3, c_2$$

Se vede că numai două, şi anume a patra şi a cincea schimbă toate elementele primei ordonări a culorilor (ce corespunde extragerii din pălărie). Deci există exact două permutări ce schimbă locul fiecăreia din ultimele trei culori şi lasă pe loc prima culoare. La fel se poate repeta raţionamentul ţinând pe loc fiecare din culorile c_2, c_3 sau c_4. În felul acesta se obţine numărul total al permutărilor ce păstrează pe loc exact o culoare, care este egal cu $4 \times 2 = 8$. Probabilitatea ca prin hazard să se dea raspuns corect la o singură întrebare este deci $\frac{8}{24} = \frac{1}{3} = 0,333$.

Să vedem acum câte permutări se pot face cu cele patru culori astfel ca două să rămână pe loc iar celelalte două să schimbe locul. De exemplu, să zicem că c_1 şi c_2 rămân pe loc. Atunci neapărat c_3 şi c_4 schimbă locul una cu alta, adică singura permutare posibilă în acest fel este (c_1, c_2, c_4, c_3). Dacă numărăm posibilităţile în care putem fixa două din culori, găsim că acestea sunt în număr de $C_4^2 = 6$. Deci există 6 permutări ce lasă exact două culori pe loc. Rezultă că probabilitatea ca exact două răspunsuri date la întâmplare să fie corecte este de $\frac{6}{24} = \frac{1}{4} = 0,25$.

Dacă trei dintre culorile unei permutări a celor 4 culori sunt lăsate pe loc, atunci neapărat şi cea de a patra culoare trebuie să rămână neschimbată, deci permutarea în cauză este permutarea identică. Rezultă că permutările care schimbă locul tuturor culorilor în acelaşi timp sunt în număr de $24 - 1 - 8 - 6 = 9$. Putem concluziona că probabilitatea ca toate răspunsurile date la întâmplare să fie eronate este de numai $\frac{9}{24} = \frac{3}{8} = 0,375$.

Desigur, probabilitatea ca toate răspunsurile să se nimerească exacte este $\frac{1}{24} = 0,04166$, destul de mică. Putem concluziona că, dacă dorim să tragem concluzii realiste bazându-ne pe experimentul descris în această problemă, trebuie să considerăm ca întâmplătoare ghicirea a una sau două culori. Numai menţionarea completă a tuturor culorilor, în ordine, poate fi considerată ca un indiciu al unei capacităţi reale deosebite de simpla nimereală. $\square$

Exemplul 2.*

Două rânduri de cărţi de joc sunt amestecate fiecare separat, după care sunt puse pe masă în perechi, o carte dintr-un pachet alături de o carte din celălalt pachet. Se formează în acest fel 52 de perechi. Care este probabilitatea ca cel puţin una din perechi să conţină două cărţi identice?

Soluţie. Este o problemă de acelaşi tip cu cea anterioară. Spaţiul tuturor posibilităţilor este modelat de mulţimea permutărilor ce se pot face cu 52 de obiecte. Problema revine la determinarea numărului de permutări care nu lasă nici un element pe locul său. Numărul 52 fiind mare în raport cu genul de calcule implicate în soluţia anterioară, nu putem repeta raţionamentul de acolo. Vom da o soluţie bazată pe egalitatea lui Poincaré demonstrată mai jos.

Vom înlocui cifra 52 cu litera n pentru a face problema mai generală. Din punctul de vedere al raţionamentului nu se va produce nici o complicaţie. Deci în continuare vom presupune că avem o mulţime formată din n elemente şi căutăm permutările (ce se pot face cu aceste elemente) cu proprietatea că nici un element nu este lăsat invariant.

Fie $M = \{1, 2, ..., n\}$ mulţimea noastră cu n elemente. Mulţimea tuturor permutărilor este descrisă astfel:

$$\Omega = \{(x_1, x_2, ..., x_n) \in M^n / \ x_i \neq x_j, \forall i \neq j\}.$$

Numărul elementelor din Ω este $n!$. Mai întâi vom analiza mulţimile de permutări care lasă pe loc un element. Fie A_j mulţimea permutărilor ce lasă pe loc elementul j :

$$A_j = \{(x_1, x_2, ..., x_n) \in \Omega / \ x_j = j\}.$$

Mulţimea permutărilor care lasă cel puţin un element pe loc este reuniunea $\bigcup_{j=1}^{n} A_j$. Mulţimea permutărilor care nu lasă pe loc nici un element al lui M, este cea care ne interesează şi o notăm cu B. Ea este complementara reuniunii anterioare:

$$B = \left(\bigcup_{j=1}^{n} A_j\right)^c.$$

Prin aplicarea formulei lui Poincaré, de mai jos, avem

$$P(B) = 1 - \sum_{i=1}^{n} P(A_i) + \sum_{i<j} P(A_i \cap A_j) -$$

$$- \sum_{i<j<k} P(A_i \cap A_j \cap A_k) + ... + (-1)^n P\left(\bigcap_{i=1}^{n} A_i\right).$$

Dar pentru mulţimile A_j ştim să calculăm probabilitatea: mulţimea A_j se pune în corespondenţă bijectivă cu mulţimea permutărilor ce se pot face cu obiectele mulţimii $M \setminus \{j\}$, deci are $(n-1)!$ elemente. Probabilitatea este atunci $P(A_j) = \frac{(n-1)!}{n!}$, pentru fiecare indice j. Pentru doi indici diferiţi $i \neq j$ putem calcula la fel numărul elementelor din intersecţia $A_i \cap A_j$: această mulţime se pune în bijecţie cu permutările ce se pot face cu elementele mulţimii $M \setminus \{i, j\}$, deci are $(n-2)!$ elemente. Probabilitatea respectivă este $P(A_i \cap A_j) = \frac{(n-2)!}{n!}$. La fel se găseşte şi pentru trei indici diferiţi: $P(A_i \cap A_j \cap A_k) = \frac{(n-3)!}{n!}$, etc. La prima sumă din formula de mai sus participă n termeni, la cea de a doua participă C_n^2 termeni, la cea de a treia C_n^3, etc. Ţinând cont de toate acestea putem calcula fiecare termen din formula de mai sus:

$$P(B) = 1 - n \times \frac{(n-1)!}{n!} + \frac{n!}{2!(n-2)!} \times \frac{(n-2)!}{n!} - \frac{n!}{3!(n-3)!} \times \frac{(n-3)!}{n!} + ...$$

$$+ (-1)^n \times \frac{1}{n!} = 1 - \frac{1}{1} + \frac{1}{2!} - \frac{1}{3!} + ... + (-1)^n \frac{1}{n!}.$$

Se observă că dezvoltând funcţia exponenţială în serie Taylor în jurul lui 0 şi apoi calculând-o în -1, obţinem seria infinită

$$\exp(-1) = 1 - \frac{1}{1} + \frac{1}{2!} - \frac{1}{3!} + ... + (-1)^n \frac{1}{n!} ...$$

Rezultă că putem evalua diferenţa $e^{-1}-P(B)$ cu restul Taylor de ordin n al funcţiei $\exp x$:

$$e^{-1}-P(B)=\frac{1}{(n+1)!}e^{-\xi}$$

cu un număr $\xi\in(-1,0)$. Deci $P(B)\approx e^{-1}\approx 0,3678$, cu o eroare de ordinul lui $\frac{1}{(n+1)!}$. Pentru $n=4$ avem $\frac{1}{5!}=\frac{1}{120}\approx 0,00833$, pentru $n=5$ avem $\frac{1}{6!}=\frac{1}{720}\approx 0,00139$ iar pentru $n=6$ avem $\frac{1}{7!}=\frac{1}{5040}<0,0002$. Din punct de vedere practic, pentru $n\geq 6$, probabilitatea ca să nu avem coincidenţe nu depinde de n. $\square$

Lema 1.2. *(Formula lui Poincaré) Fie* $(\Omega,\mathcal{F},P)$ *un spaţiu probabilizat şi* $A_1,...,A_n$ *evenimente arbitrare. Atunci are loc următoarea formulă*

$$P\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}P(A_i)-\sum_{i<j}P(A_i\cap A_j)+$$
$$+\sum_{i<j<k}P(A_i\cap A_j\cap A_k)+...+(-1)^{n-1}P\left(\bigcap_{i=1}^{n}A_i\right).$$

Demonstraţie. Pentru $n=2$ formula devine

$$P(A_1\cup A_2)=P(A_1)+P(A_2)-P(A_1\cap A_2).$$

Verificarea ei se face pornind cu următoarele descompuneri în mulţimi disjuncte

$$A_1=(A_1\backslash A_2)\cup(A_1\cap A_2),\ A_2=(A_2\backslash A_1)\cup(A_1\cap A_2),$$
$$A_1\cup A_2=A_1=(A_1\backslash A_2)\cup(A_2\backslash A_1)\cup(A_1\cap A_2).$$

Utilizând aditivitatea măsurii de probabilitate se ajunge la concluzia dorită. În continuare se procedează prin inducţie. Presupunem că formula este adevărată pentru n şi vom demonstra că ea este valabilă şi pentru $n+1$. Notăm $B=\bigcup_{i=1}^{n}A_i$ şi scriem

$$P\left(\bigcup_{i=1}^{n+1}A_i\right)=P(A_{n+1})+P(B)-P(A_{n+1}\cap B)=$$
$$=\sum_{i=1}^{n+1}P(A_i)-\sum_{i<j}^{\leq n}P(A_i\cap A_j)+\sum_{i<j<k}^{\leq n}P(A_i\cap A_j\cap A_k)+...$$
$$+(-1)^{n-1}P\left(\bigcap_{i=1}^{n}A_i\right)-P(A_{n+1}\cap B).\quad(*)$$

Ultimul termen din această expresie se exprimă din nou prin utilizarea formulei Poincaré cu n mulţimi:

$$P(A_{n+1}\cap B)=P\left(\bigcup_{i=1}^{n}(A_{n+1}\cap A_i)\right)=$$

$$=\sum_{i=1}^{n} P\left(A_{n+1}\cap A_i\right)-\sum_{i<j}^{\leq n} P\left(A_{n+1}\cap A_i\cap A_j\right)+...+(-1)^{n-1} P\left(\bigcap_{i=1}^{n+1} A_i\right).$$

Ţinând cont de relaţiile

$$\sum_{i<j}^{\leq n} P\left(A_i\cap A_j\right)+\sum_{i=1}^{n} P\left(A_{n+1}\cap A_i\right)=\sum_{i<j}^{\leq n+1} P\left(A_i\cap A_j\right),$$

$$\sum_{i<j<k}^{\leq n} P\left(A_i\cap A_j\cap A_k\right)+\sum_{i<j}^{\leq n} P\left(A_{n+1}\cap A_i\cap A_j\right)=\sum_{i<j<k}^{\leq n+1} P\left(A_i\cap A_j\cap A_k\right),$$

şi aşa mai departe ..., rezultă că partea dreaptă a egalităţii (*) devine

$$=\sum_{i=1}^{n+1} P\left(A_i\right)-\sum_{i<j}^{\leq n+1} P\left(A_i\cap A_j\right)+\sum_{i<j<k}^{\leq n+1} P\left(A_i\cap A_j\cap A_k\right)+...$$
$$+(-1)^n P\left(\bigcap_{i=1}^{n+1} A_i\right),$$

care este exact membrul drept al formulei ce dorim să probăm cu $n+1$ mulţimi.$\square$

4.3. Problema cu zilele de naştere. Ne propunem să calculăm probabilitatea ca dintr-un grup de n persoane, aflate împreună întâmplător, toate să aibă zile de naştere diferite. Vom arăta că pentru $n=23$ această probabilitate devine mai mică decât $0,5$.

Soluţie. Pentru a modela probabilistic problema facem o comparaţie a zilei de naştere pe care o are o persoană cu extragerea din urna cu 365 de bile. Pentru n persoane va corespunde o serie de n extrageri din urnă cu revenire. (Vezi exemplul 3. din paragraful despre repartiţia unei variabile pentru o discuţie mai riguroasă a modelării, care însă conduce la aceeaşi formulă.) Definim spaţiul $M=\{1,2,...,365\}$ constând din numerele de la 1 la 365 şi reprezentând zilele anului. Mulţimea ce descrie toate posibilităţile este

$$\Omega=\{\omega=(x_1,...,x_n)\mid x_i\in M \quad,\ i=1,...,n\}=M^n.$$

Avem $card\,(\Omega)=365^n$ şi definim o probabilitate pe Ω care să facă echiprobabile toate evenimentele elementare, deci punem

$$P\left(\{\omega\}\right)=\frac{1}{365^n}.$$

Mulţimea care ne interesează o notăm cu Λ şi constă din evenimentele elementare pentru care toate componentele sunt distincte:

$$\Lambda=\left\{\omega=(x_1,...,x_n)\in\Omega\mid x_i\neq x_j \quad,\ \forall\ i\neq j\right\}.$$

Bineînţeles că trebuie să presupunem $n \leq 365$ pentru ca Λ să nu fie vidă. Mulţimea Λ are un număr de A_{365}^{n} elemente. Atunci putem exprima probabilitatea

$$P(\Lambda) = \frac{365(365-1)...(365-n+1)}{365^n}.$$

Este clar că valoarea acestei expresii scade cu cât creşte n. Calculul exact al acestei probabilităţi pentru toate valorile $n = 1, ..., 365$ se poate face utilizând calculatorul. Mai jos vom face o estimare cu calcule „de mână” suficient de fină a acestei probabilităţi astfel încât să putem stabili că pentru $n = 22$ probabilitatea în cauză este mai mare decât $0,5$, iar pentru $n = 23$ acea probabilitate este mai mică decât $0,5$.

Aspecte numerice pentru problema anterioară.

Mai departe ne vom ocupa cu estimarea funcţiei definită prin produsul

$$f(n) = \left(1 - \frac{1}{365}\right) ... \left(1 - \frac{n-1}{365}\right),$$

care reprezintă probabilitatea căutată în problema zilelor de naştere. Vom aplica acestui produs *lema* 1.1. A doua inegalitate din primul rând de la punctul (iii) al lemei ne dă

$$f(n) \leq \exp - \sum_{k=1}^{n-1} \frac{k}{365} = \exp \frac{-(n-1)n}{730}.$$

Vom nota $g(n) = \exp - \frac{(n-1)n}{730}$ astfel că inegalitatea anterioară se scrie $f(n) \leq g(n)$. Presupunând că $n \leq 74$ vom putea să aplicăm şi a doua inegalitate de la punctul (iii) al lemei, care conduce la

$$g(n) - f(n) \leq g(n) \frac{7}{12} \sum_{k=1}^{n-1} \frac{k^2}{365^2}.$$

Suma din membrul drept se majorează prin

$$\sum_{k=1}^{n-1} \frac{k^2}{365^2} \leq \frac{n(n-1)(2n-1)}{6 \times 365^2},$$

astfel că notând $r_n = \frac{7}{12} \frac{n(n-1)(2n-1)}{6 \times 365^2}$ putem scrie

$$g(n) - f(n) \leq g(n) r_n.$$

Calculăm $g(22) \approx 0,531$ şi $g(22) r_{22} \approx 0,0077$, ceea ce conduce la $f(22) \geq 0,523$. În continuare calculăm $g(23) \approx 0,499998$, ceea ce implică $f(23) \leq 0,49999 < 0,5$. Cu aceasta am încheiat verificarea afirmaţiei făcute la început.

Este semnificativ de notat că $f(46) \leq g(46) \approx 0,058$ şi $f(69) \leq g(69) \approx 0,0016$. De aceea, dacă se face experimentul de a lista datele de naştere a 23 de studenţi, în situaţia că nu se nimeresc doi cu aceeaşi zi de naştere, se poate continua cu listarea zilelor de naştere ale părinţilor. În mod practic, atunci vom putea fi siguri că cel puţin două date coincid.□

5. Anexă: noţiuni de combinatorică

În această secţiune vom face o recapitulare a principalelor noţiuni de combinatorică ce sunt învăţate în liceu. Cu această ocazie vom stabili şi o notaţie ce o vom păstra în restul cărţii.

O mulţime finită cu m elemente va fi notată M. Pentru a fi în ton cu modelele noastre probabiliste putem să gândim că M este formată din bile numerotate de la 1 la m. Formal vom scrie $M = \{1, ..., m\}$.

1. Dacă avem două mulţimi distincte $M_i = \{1, ..., m_i\}, i = 1, 2$, pentru a determina cardinalul mulţimii produs $M_1 \times M_2$ facem numărătoarea elementelor înşirându-le în ordine lexicografică:

$$\begin{aligned} M_1 \times M_2 &= \{(1,1), (1,2), ..., (1, m_2), \\ &\quad (2,1), (2,2), ..., (2, m_2), ..., (m_1, 1), (m_1, 2), ..., (m_1, m_2)\}. \end{aligned}$$

Putem astfel constata că avem $card\,(M_1 \times M_2) = m_1 m_2$.

Pentru k mulţimi vom avea

$$card\,(M_1 \times ... \times M_k) = m_1 ... m_k.$$

În particular, dacă considerăm produsul de k ori al aceleiaşi mulţimi M din m elemente vom avea $card\left(M^k\right) = m^k$.

2. Numărul modalităţilor distincte în care pot fi înşiruite elementele mulţimii $M = \{1, ..., m\}$ este numit „permutări de m obiecte" şi este notat P_m. Vom demonstra în continuare, prin inducţie, că acest număr este egal cu $P_m = 1 \cdot 2 \cdot 3 \cdot ... \cdot m = m!$.

O înşiruire a elementelor din M poate fi descrisă prin sistemul de elemente $(x_1, ..., x_m) \in M^m$, unde se presupune că $x_i \neq x_j$ pentru $i \neq j$. Mulţimea tuturor înşiruirilor ce se pot face cu elementele din M o putem deci descrie prin

$$\mathcal{M}_m = \{(x_1, ..., x_m) \in M^m / x_i \neq x_j, \forall i \neq j\}.$$

Şi vrem să arătăm că $card\,(\mathcal{M}_m) = m!$. Presupunem că este adevărată relaţia pentru $m-1$, adică $card\,(\mathcal{M}_{m-1}) = (m-1)!$, unde

$$\mathcal{M}_{m-1} = \left\{(x_1, ..., x_{m-1}) \in (M')^{m-1} / x_i \neq x_j, \forall i \neq j\right\},$$

iar $M' = \{1, ..., m-1\}$. Pentru fiecare element $x \in M$, fixat, observăm că mulţimea

$$\mathcal{U}(x) = \{(x_1, ..., x_m) \in \mathcal{M}_m / x_m = x\}$$

este izomorfă cu $\mathcal{M}_{m-1}$. Pe de altă parte, are loc egalitatea

$$\mathcal{M}_m = \bigcup_{x \in M} \mathcal{U}(x),$$

care permite numărarea elementelor - obţinem

$$card\,(\mathcal{M}_m) = m \cdot card\,(\mathcal{M}_{m-1}) = m!. \ \square$$

3. Fie $1 \leq k \leq m$ şi să presupunem în continuare că M este mulţimea anterioară. Numărul grupurilor ordonate formate cu câte k obiecte din M este numit „aranjamente de m obiecte luate câte k ” şi este notat A_m^k. Dorim să arătăm că acest număr este egal cu $A_m^k = m\,(m-1)\ldots(m-k+1)$. Notăm

$$\mathcal{M}_m^k = \left\{(x_1, ..., x_k) \in M^k / x_i \neq x_j, \forall i \neq j\right\}$$

mulţimea tuturor acestor grupuri ordonate. Din nou raţionăm prin inducţie şi presupunem că ştim formula

$$card\left(\mathcal{M}_{m-1}^{k-1}\right) = (m-1)\,(m-2)\ldots(m-k+1)\,.$$

Pentru fiecare element fixat, $x \in M$, notăm

$$\mathcal{V}\,(x) = \left\{(x_1, ..., x_k) \in \mathcal{M}_m^k / x_k = x\right\}.$$

Vom avea relaţia

$$\mathcal{M}_m^k = \bigcup_{x \in M} \mathcal{V}\,(x)\,.$$

Pe de altă parte, este uşor de văzut că $\mathcal{V}\,(x)$ este izomorfă cu $\mathcal{M}_{m-1}^{k-1}$ şi de aceea deducem

$$card\left(\mathcal{M}_m^k\right) = m \cdot card\left(\mathcal{M}_{m-1}^{k-1}\right) = m\,(m-1)\ldots(m-k+1)\,. \ \square$$

4. Fiind dată aceeaşi mulţime M cu m elemente, cardinalul mulţimii părţilor lui M care au k elemente este numit „combinări de m obiecte luate câte k ”. Acest număr este notat cu C_m^k şi vom arăta, din nou prin inducţie după m, că are loc formula $C_m^k = \frac{m!}{k!(m-k)!}$. Pentru aceasta notăm

$$\Gamma = \{A \subset M / cardA = k\}\,.$$

Descompunem această mulţime în două $\Gamma = \Gamma_1 \cup \Gamma_2$, unde am notat

$$\Gamma_1 = \{A \in \Gamma / m \in A\}\,, \quad \Gamma_2 = \{A \in \Gamma / m \notin A\}\,.$$

Dacă utilizăm şi notaţia $M' = \{1, ..., m-1\}$, putem să scriem

$$\Gamma_2 = \{A \subset M' / cardA = k\}$$

şi pe baza ipotezei de inducţie avem $card\,(\Gamma_2) = C_{m-1}^k = \frac{(m-1)!}{k!(m-1-k)!}$. Pe de altă parte, mulţimea Γ_1 este izomorfă cu $\{A' \subset M' / cardA' = k-1\}$. De aceea avem $card\,(\Gamma_1) = C_{m-1}^{k-1} = \frac{(m-1)!}{(k-1)!(m-1-k)!}$. Rezultă

$$\begin{aligned} C_m^k &= card\,(\Gamma) = card\,(\Gamma_1) + card\,(\Gamma_2) = \\ &= \frac{(m-1)!}{k!\,(m-1-k)!} + \frac{(m-1)!}{(k-1)!\,(m-k)!} = \frac{m!}{k!\,(m-k)!}. \ \square \end{aligned}$$

6. Exerciţii

Exerciţiul 1.1. *Biletele de la o tombolă sunt numerotate de la 1 la 10.000. Ce proporţie dintre ele se termină cu cifra 1, dar cu 2, ..., dar cu 9? Ce proporţie dintre ele încep cu cifra 1, dar cu 2, ..., dar cu 9?*

Exerciţiul 1.2. *Se consideră zilele de 13 din lunile anului 2006. Ce proporţie dintre ele sunt luni, dar marţi, ..., dar duminică? Aceeaşi problemă pentru zilele de 13 din anii 2006 şi 2007. Câţi ani consecutivi trebuiesc luaţi în calcul pentru a avea aceeaşi proporţie de zile de 13 pentru fiecare zi din săptămână? (Se va ţine cont de anii bisecţi.)*

Exerciţiul 1.3. *Fie Ω o mulţime şi $A, B \subset \Omega$. Scrieţi $A \cup B$ ca o reuniune de mulţimi disjuncte care sunt obţinute din A şi B prin operaţiile de algebră de părţi (intersecţie şi complementară) şi care sunt total incluse sau în A sau în B. În câte feluri se poate face acest lucru? Aceeaşi problemă pentru reuniunea a trei mulţimi $A \cup B \cup C$.*

Exerciţiul 1.4. *Fiind dat un spaţiu probabilizat $(\Omega, \mathcal{F}, P)$ şi $A, B, C \in \mathcal{F}$ să se arate că are loc relaţia*

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

Exerciţiul 1.5. *Un pion este pus pe poziţia A în şirul de litere de mai jos.*

$$A \qquad B \qquad C \qquad D \qquad E$$

Se aruncă cu banul de 4 ori la rând şi după fiecare aruncare se acţionează asupra pionului în felul următor: dacă iese cifra, acesta se mută cu o poziţie la dreapta; dacă iese stema, rămâne pe loc. Să se afle probabilitatea ca pionul să: 1) rămână în poziţia A, 2) depăşească B, 3) ajungă în poziţia E.

Exerciţiul 1.6. *Într-o familie 4 fete se ocupă de spălatul vaselor pe rând. Într-o anumită perioadă de timp s-au spart 4 farfurii, dintre care 3 le-a spart mezina. Ea spune că întâmplarea a făcut să se nimerească la ea aceste incidente. Ce sens probabilistic se poate da afirmaţiei făcute de mezină? Presupunând că incidentele se produc aleator, care este probabilitatea ca să se nimerească cel puţin 3 la una din cele 4 fete? Care este probabilitatea ca acestea să se producă la mezină? Cum se schimbă problema dacă după un interval de timp se mai sparg două farfurii, din care una este spartă tot de mezină? Cum interpretam problema dacă de la început era vorba doar de trei farfurii sparte, din care două erau sparte de mazină?*

(Indicaţie. Pentru a modela problema ne imaginăm că toate fetele participă la spălatul vaselor în acelaşi timp. Dacă se sparge o farfurie şi hazardul alege la care din ele să se petreacă incidentul, acest fenomen este analog cu extragerea unei bile dintr-o urnă cu patru bile. Repetarea

de încă trei ori a fenomenului, este analoagă cu repetarea extragerilor cu întoarcere de încă trei ori.)

Exerciţiul 1.7. *Trei studenţi, A,B,C, folosesc împreună o maşină. În decursul a doi ani maşina a fost lovită (uşor) de patru ori, după cum urmează: o dată a fost lovită de A şi de trei ori de C. Se poate trage concluzia că C este mai neîndemânatic? Pentru a lămuri situaţia este relevant să se presupună că accidentele se produc aleator şi se petrec la oricare din cei trei studenţi cu aceeaşi probabilitate. Sub această ipoteză, să se calculeze care ar fi probabilitatea ca din patru accidente petrecute, trei să cadă la un student, iar al patrulea la un alt student?*

Exerciţiul 1.8. *Jucătorul A aruncă cu patru zaruri şi câştigă un premiu dacă obţine cel puţin un as. Jucătorul B aruncă de* 24 *de ori cu două zaruri şi câştigă un premiu dacă obţine la cel puţin una din aruncări o dublă de aşi. Care din jucători are mai multe şanse de a câştiga? (Problemă cunoscută sub numele de problema lui Mère).*

Exerciţiul 1.9. *Jucătorul A aruncă de* 6 *ori cu un zar şi câştigă un premiu dacă obţine cel puţin un as. Jucătorul B aruncă de* 12 *ori cu zarul şi câştigă un premiu dacă obţine cel puţin doi aşi. Care din jucători are probabilitatea de a câştiga mai mare? (Problemă pusă lui Newton, care i-a dat pe loc o soluţie.)*

Exerciţiul 1.10. *La o uşă există două broaşte. De obicei aveţi* 6 *chei asemănătoare asupra dumneavoastră şi printre ele sunt şi cele de la uşa cu pricina. 1) Care este probabilitatea ca primele două chei încercate să deschidă cele două broaşte? 2) Aţi pierdut una din cele şase chei, nu se ştie care. Ce probabilitate există ca să puteţi totuşi deschide uşa respectivă?*

Exerciţiul 1.11. *Într-o parcare sunt* 12 *locuri aşezate unul lângă altul în linie.* 8 *dintre aceste locuri sunt ocupate, iar* 4 *sunt libere. Cineva observă că cele* 4 *locuri libere sunt unul lângă celălalt şi îşi pune întrebarea dacă aceasta este o întâmplare. Care este probabilitatea ca* 4 *din* 12 *poziţii aflate în linie să fie consecutive?*

Exerciţiul 1.12. *Un profesor primeşte* 12 *bilete de amendă pentru parcare pe trotuar lângă Universitate. Dacă toate biletele sunt primite marţea sau joia, se poate trage concluzia că profesorul vine numai în zilele respective cu maşina la universitate? Care este probabilitatea ca din* 12 *evenimente care pot avea loc în oricare din primele* 5 *zile ale săptămânii, toate să se petreacă în zilele de marţi sau joi? Care este probabilitatea ca din* 12 *evenimente petrecute întâmplător în oricare din* 5 *zile, să se nimerească toate în numai* 2 *zile?*

Exerciţiul 1.13. *Din* 15 *bilete de amendă pentru parcare pe trotuar nici unul nu este dat vinerea. Se poate trage concluzia, din această* *informaţie, că vinerea nu sunt probleme cu parcarea? Care este probabilitatea ca din* 5 *evenimente petrecute întâmplător în oricare din cele* 5 *zile lucrătoare, toate să aibă loc în* 4 *?*

Exerciţiul 1.14. *** 15 *elevi nou sosiţi într-o şcoală sunt repartizaţi în mod egal în* 3 *clase paralele. Printre noii veniţi se află şi trei băieţi mai înalţi. Ei prezintă interes pentru echipele de baschet ale celor trei clase. De aceea se pune următoarea problemă. Presupunând că repartizarea lor se face aleator, care este probabilitatea ca cei trei elevi înalţi să nimerească în aceeaşi clasă? Care este probabilitatea ca ei să nimerească în clase diferite? Dar să nimerească* 2 *în aceeaşi clasă şi al treilea în altă clasă?*

Exerciţiul 1.15. *Se aruncă 6 zaruri. Care este probabilitatea ca să se obţină trei numere distincte, fiecare repetat de două ori?*

Exerciţiul 1.16. ** 6 persoane din aceeaşi localitate fac naveta cu un tren care are 3 vagoane. Fiecare se urcă la întâmplare într-un vagon. Care este probabilitatea să nimerească toţi în acelaşi vagon? Care este probabilitatea ca să nimerească câte doi în fiecare vagon? Fiecare vagon are câte 10 compartimente şi se presupune că fiecare persoană nimereşte aleator într-un compartiment. Care este probabilitatea ca să nimerească fiecare în alt compartiment?*

CAPITOLUL 2

Câteva noţiuni de bază

În acest capitol vom prezenta câteva din noţiunile de bază din teoria probabilităţilor care iau în considerare numai un număr finit de evenimente. Ele pot fi discutate cu toată semnificaţia în cadrul unui spaţiu probabilizat finit. Cele mai multe din exemplele pe care le discutăm vor fi tocmai de acest tip. Totuşi vom utiliza drept cadru de bază un spaţiu probabilizat general $(\Omega, \mathcal{F}, P)$, pentru a nu fi obligaţi mai târziu să revenim asupra acestor noţiuni.

1. Probabilităţi condiţionate

1.1. Noţiunea de probabilitate condiţionată. Fiind dat spaţiul probabilizat $(\Omega, \mathcal{F}, P)$ şi mulţimea $A \in \mathcal{F}$ astfel ca $P(A) > 0$, vom nota

$$P(B/A) = \frac{P(B \cap A)}{P(A)},$$

pentru orice mulţime $B \in \mathcal{F}$ şi vom denumi această expresie probabilitatea lui B *condiţionată* de A. Se verifică uşor că $B \to P(B/A)$ este o măsură de probabilitate pe $(\Omega, \mathcal{F})$ şi că $P(A^c/A) = 0, P(A/A) = 1$, deci $P(\cdot/A)$ este suportată de A. Următoarele exemple dau o idee despre semnificaţia acestei definiţii.

Exemplul 1.

Să presupunem că într-o sală de curs se află un număr 60 de studenţi dintre care 35 de fete şi 25 de băieţi. Presupunem că 10 dintre fete şi 10 dintre băieţi au înălţimea mai mare de 1,70. Se alege la întâmplare numele unui student. Să se determine:

a) probabilitatea ca persoana al cărei nume a fost ales la întâmplare din catalog să fie mai înaltă de 1,70.

b) probabilitatea ca o studentă să fie mai înaltă de 1,70.

c) probabilitatea ca un student băiat să fie mai înalt de 1,70.

Soluţie. Vom nota cu Ω mulţimea tuturor studenţilor, cu A mulţimea fetelor, iar cu B mulţimea studenţilor de înălţime mai mare ca 1,70. Probabilitatea este dată de $P(\{\omega\}) = \frac{1}{card\Omega}$, adică este probabilitatea ce acordă aceaşi valoare fiecărui individ.

a) Avem $P(B) = \frac{20}{60} = 0,33$.

b) Avem de calculat probabilitatea condiţionată $P(B/A) = \frac{P(B \cap A)}{P(A)} = \frac{10}{35} = 0,28$.

c) Calculăm probabilitatea condiţionată $P(B/A^c) = \frac{P(B\cap A^c)}{P(A^c)} = \frac{10}{25} = 0,4$. □

Exemplul 2.

Tratamentul aplicat unui număr de 700 pacienţi care sufereau de piatră la rinichi a dat rezultatele următoare: pentru 562 tratamentul a avut succes, iar pentru 138 a avut eşec. Notăm cu Ω mulţimea pacienţilor şi cu A mulţimea pacienţilor care au fost vindecaţi. Deci probabilitatea de succes a tratamentului este $P(A) = \frac{562}{700} = 0,802$. Analizând rezultatele se constată că, împărţind pietrele după dimensiunea lor, comparată cu dimensiunea vaselor renale, în două grupe, pietre mici şi pietre mari, rezultatele pot fi clasate astfel: au fost 357 pietre mici şi pentru 315 dintre ele tratamentul a condus la succes, iar pentru 42 a eşuat; au fost 343 pietre mari, pentru 247 din ele tratamentul a avut succes, iar pentru 96 a eşuat. Putem trage concluzia că avem alte probabilităţi care sunt semnificative. Notăm cu B mulţimea pacienţilor cu piatră mică. Probabilitatea ca tratamentul să reuşească în cazul unei pietre mici este probabilitatea condiţionată $P(A/B) = \frac{315}{357} = 0,882$. Probabilitatea ca tratamentul să reuşească în cazul unei pietre mari este $P(A/B^c) = \frac{247}{342} = 0,722$, astfel că rezultatele tratamentului sunt mai nuanţat caracterizate prin intermediul probabilităţilor condiţionate.

Exemplul 3. (un model din actuariat)

Pentru a modela probabilistic statistica vârstei de deces a unui grup de indivizi (este vorba de o colectivitate mare, de ordinul unei ţări, de exemplu) se consideră o măsură de probabilitate P pe intervalul de timp $[0,100]$, cu unitatea de măsură 1 an. Probabilitatea de deces în intervalul (s,t) este exprimată printr-o formulă de tipul

$$P((s,t)) = \int_s^t \alpha(r)\,dr,$$

unde $\alpha(r)$ este o funcţie de densitate (sau de intensitate a deceselor). Altfel spus, numărul $P((s,t))$ reprezintă raportul dintre numărul deceselor petrecute între vârsta de s ani şi vârsta de t ani faţă de numărul total de decese. Adică raportat la numărul total al indivizilor din populaţia luată în considerare, pentru că se presupune că nici unul nu trăieşte mai mult de o sută de ani. Funcţia α are o expresie analitică ce se determină prin metode speciale după analiza datelor statistice înregistrate oficial.

Pentru a face calcule, în continuare vom considera următoarea expresie simplificată pentru densitate:

$$\alpha(s) = 3\cdot 10^{-9} s^2 (100-s)^2, \quad s\in(0,100).$$

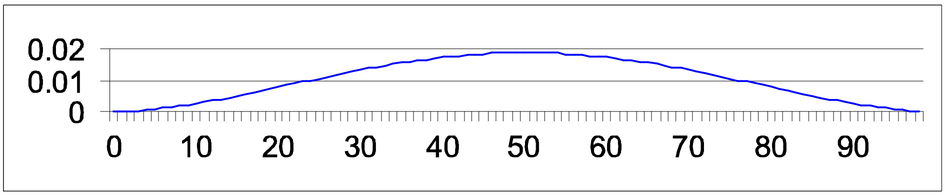


FIGURA 1. Graficul funcţiei α

Se verifică faptul că această funcţie conduce la o probabilitate:

$$\int_0^{100} \alpha(s)\, ds = 1.$$

Graficul funcţiei α arată cum se vede în figura 1. Este un grafic în formă de farfurie întoarsă, care, desigur, nu are decât o asemănare aproximativă cu situaţiile reale.

Vom nota cu A evenimentul deceselor între 60 şi 70 de ani. Probabilitatea acestui eveniment se poate calcula exact, dacă utilizăm formula anterioară pentru α,

$$P(A) = P((60, 70)) = \int_{60}^{70} \alpha(s)\, ds = 0,154.$$

Această valoare reprezintă raportul dintre numărul deceselor ce au loc între 60 şi 70 de ani şi numărul total al deceselor. Dacă notăm cu B evenimentul deceselor peste 60 de ani putem calcula probabilitatea

$$P(B) = \int_{60}^{100} \alpha(s)\, ds = 0,316,$$

după care putem calcula şi probabilitatea condiţionată

$$P(A/B) = \frac{P(A)}{P(B)} = \frac{0,154}{0,316} = 0,486.$$

Acest număr reprezintă raportul dintre numărul deceselor între 60 şi 70 şi numărul deceselor peste 60 de ani, sau altfel spus, este probabilitatea de a deceda până în 70 de ani condiţionată de faptul că am ajuns la 60 de ani. După cum se vede, cele două numere au semnificaţie şi valori diferite: $P(A) \neq P(A/B)$.

Rezultă că probabilitatea de a trăi peste 70 de ani condiţionată de faptul că am ajuns la 60 de ani este $P(C/B) = 1 - 0,486 = 0,514$, unde am notat cu $C = (70, 100)$ evenimentul deceselor peste 70 de ani. Pentru comparaţie, putem calcula şi $P(C) = 0,316 - 0,154 = 0,162$. Desigur, funcţia α cu care lucrăm nu este decât o aproximaţie grosieră pentru datele reale. $\square$

Formula probabilităţii condiţionate se utilizează şi sub forma

$$P(A\cap B)=P(B)\,P(A/B),$$

care permite calculul probabilităţii unei intersecţii. Sub această formă mai este numită şi *regula înmulţirii* probabilităţii condiţionate. Prin iterarea acestei formule se obţine relaţia de înmulţire succesivă

$$P(B_1\cap...\cap B_n)=$$
$$=P(B_1)\,P(B_2/B_1)\,P(B_3/B_1\cap B_2)...P(B_n/B_1\cap...\cap B_{n-1}).$$

În cazul în care spaţiul Ω este partiţionat de evenimentele $B_1,...,B_n$, atunci putem scrie $A=(A\cap B_1)\cup...\cup(A\cap B_n)$ şi deducem

$$\begin{aligned}P(A)&=P(A\cap B_1)+...+P(A\cap B_n)=\\&=P(B_1)\,P(A/B_1)+...+P(B_n)\,P(A/B_n),\end{aligned}$$

relaţie cunoscută sub numele de formula *mediei probabilităţilor condiţionate.* Această denumire va fi înţeleasă după ce vom introduce noţiunea de medie a unei variabile aleatoare. De asemenea, formula mai este denumită şi *formula probabilităţii totale.*

Exemplul 4.

Anumite companii de asigurări încheie poliţe de asigurare auto luând în considerare şi vârsta conducătorului auto. Datele statistice arată că, pentru o anumită perioadă, frecvenţa accidentelor este de 4 la mie printre tinerii sub 25 de ani şi de numai 2 la mie printre persoanele ce au depăşit vârsta de 25 de ani. Se mai ştie că tinerii sub 25 de ani reprezintă 15 la sută dintre conducătorii auto. Care este probabilitatea ca, oprind un autoturism, în el să se afle un tânăr care să fi avut accident? Să se afle frecvenţa accidentelor pe ansamblul populaţiei.

Soluţie. Notăm cu Ω mulţimea tuturor conducătorilor auto, cu A mulţimea celor ce au avut accident în perioada respectivă şi cu B mulţimea tinerilor sub 25 de ani. Datele problemei ne spun că

$$P(A/B)=\frac{4}{1000},\ P(A/B^c)=\frac{2}{1000},\ P(B)=\frac{15}{100}.$$

Pentru a răspunde la prima întrebare trebuie să calculăm

$$P(A\cap B)=P(B)\,P(A/B)=\frac{15\times 4}{100\times 1000}=\frac{6}{10000}.$$

Pentru a răspunde la a doua întrebare calculăm

$$P(A)=P(A\cap B)+P(A\cap B^c)=\tfrac{6}{10000}+P(B^c)\,P(A/B^c)=$$
$$=\tfrac{6}{10000}+\tfrac{85\times 2}{100\times 1000}=\tfrac{23}{10000}.\ \square$$

Exemplu 5.*

La alegerile pentru o anumită funcţie sunt înscrişi doi candidaţi. În urma alegerilor, candidatul A primeşte n voturi, iar candidatul B un număr mai mic $m<n$ voturi. Ne propunem să demonstrăm că probabilitatea ca, pe parcursul numărării voturilor, candidatul A să

fie mereu cotat cu un număr de voturi superior lui B este egală cu $(n-m)/(n+m)$.

Soluţie. Vom nota cu $M = \{1, ..., n+m\}$ mulţimea ce reprezintă voturile exprimate şi cu $N = \{1, ..., n\}$ mulţimea ce reprezintă voturile favorabile candidatului A. $M \backslash N$ este mulţimea ce reprezintă voturile acordate lui B. Mulţimea care reprezintă toate posibilităţile de înşiruire pentru numărarea voturilor este mulţimea permutărilor lui M şi o putem descrie în felul următor:

$$\Omega = \left\{(x_1, ..., x_{n+m}) \in M^{n+m} / x_i \neq x_j, \forall i \neq j\right\}.$$

Cardinalul acestei mulţimi este $(n+m)!$ şi bineînţeles că probabilitatea care se impune pe acest spaţiu este probabilitatea şanselor egale. Evenimentul a cărui probabilitate dorim să o calculăm poate fi descris astfel:

$$A_{m,n} = \left\{(x_1, ..., x_{n+m}) \in \Omega / card\left(\{x_1, ..., x_k\} \cap N\right) > \frac{k}{2}, \forall k \leq n+m\right\}.$$

Pentru a exprima probabilitatea acestui eveniment vom face o discuţie după rezultatul ultimului vot numărat. Vom nota deci

$$B_1 = \{(x_1, ..., x_{n+m}) \in \Omega / x_{n+m} \in N\},$$
$$B_2 = \{(x_1, ..., x_{n+m}) \in \Omega / x_{n+m} \in M \backslash N\}.$$

Dorim să calculăm probabilităţile $P(A_{n,m} \cap B_i), i = 1, 2$. Pentru aceasta vom fixa ultimul vot numărat definind, pentru $x \in N$,

$$B_x = \{(x_1, ..., x_{n+m}) \in \Omega / x_{n+m} = x\},$$

astfel că avem $B_1 = \cup_{x \in N} B_x$ şi $P(A_{n,m} \cap B_1) = \sum_{x \in N} P(A_{n,m} \cap B_x)$. Acum, analizând mulţimile B_x şi $A_{n,m} \cap B_x$, putem stabili o recurenţă, raportându-ne la experimentul în care avem $n-1$ voturi exprimate pentru A şi m voturi exprimate pentru B. Anume, mulţimea B_x este izomorfă cu mulţimea permutărilor mulţimii $M \backslash \{x\}$, iar mulţimea $A_{n,m} \cap B_x$ este izomorfă cu mulţimea $A_{n-1,m}$, care descrie similar numărătoarea voturilor ce îl dau pe A câştigător în cazul aceleiaşi probleme asociate numerelor $n-1$ şi m. Atunci putem scrie $P(A_{n,m}/B_x) = P(A_{n-1,m})$, unde cea de a doua probabilitate este, binînţeles, calculată în raport cu modelul ce corespunde perechii $n-1, m$. Dacă apelăm la un raţionament de inducţie matematică, putem presupune că această ultimă probabilitate are expresia din enunţ, adică $P(A_{n-1,m}) = (n-1-m)/(n-1+m)$. Putem atunci calcula

$$\begin{aligned} P(A_{n,m} \cap B_x) = P(B_x) P(A_{n,m}/B_x) &= \tfrac{(n-1+m)!}{(n+m)!} \tfrac{n-1-m}{n-1+m} \\ &= \tfrac{n-1-m}{(n+m)(n-1+m)}, \end{aligned}$$

ţinând cont că mulţimea B_x are cardinalul $(n-1+m)!$. Rezultă apoi

$$P(A_{n,m} \cap B_1) = \frac{n(n-1-m)}{(n+m)(n-1+m)}.$$

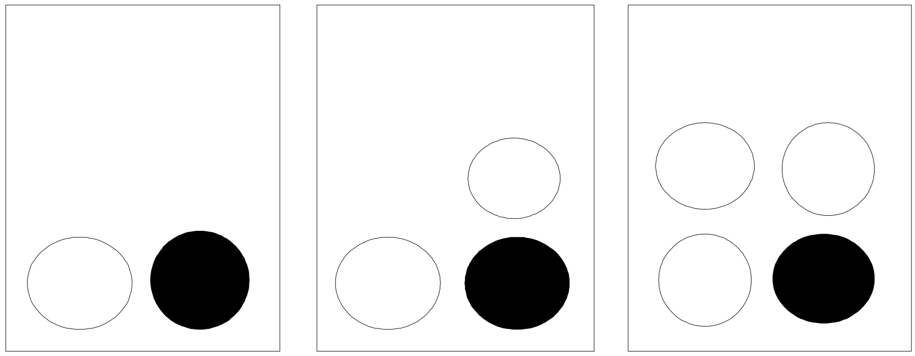


FIGURA 2. Cele trei urne cu bilele vizualizate

Un raţionament absolut similar se face pentru cazul în care $x \in M \backslash N$, obţinând $P(A_{n,m}/B_x) = P(A_{n,m-1})$, ceea ce conduce la

$$P(A_{n,m} \cap B_x) = P(B_x) P(A_{n,m}/B_x) = \frac{n-m+1}{(n+m)(n+m-1)}.$$

Rezultă că

$$P(A_{n,m} \cap B_2) = \frac{m(n-m+1)}{(n+m)(n+m-1)}$$

şi, prin urmare,

$$P(A_{n,m}) = \frac{n(n-1-m)}{(n+m)(n-1+m)} + \frac{m(n-m+1)}{(n+m)(n+m-1)} = \frac{n-m}{n+m}. \square$$

1.2. Formula lui Bayes. Vom examina mai întâi un exemplu.

Extragerea din trei urne.

Presupunem că trei urne conţinând bile se află pe o masă: prima urnă conţine o bilă albă şi o bilă neagră; a doua urnă conţine două bile albe şi una neagră; a treia urnă conţine trei bile albe şi una neagră. Aspectul exterior al urnelor este identic şi nu se vede conţinutul lor. (În figura 2 am desenat totuşi şi bilele, în mod schematic, pentru a vizualiza experimentul.)

Se alege la întâmplare o urnă şi se extrage o bilă din ea. Ştiind că bila extrasă este albă, ne punem problema de a determina probabilitatea ca bila să fi fost extrasă din urna a treia.

Soluţie.Pentru a modela această problemă vom considera mulţimea Ω ca fiind dreptunghiul mare desenat în figura 3.

Pentru diverse submulţimi $D \subset \Omega$ considerăm măsura de probabilitate dată de $P(D) = \frac{ariaD}{aria\Omega}$. Acest dreptunghi este împărţit în trei dreptunghiuri de arii egale B_1, B_2, B_3, care corespund evenimentelor "urna aleasă a fost urna i", $i = 1, 2, 3$. Avem $P(B_i) = \frac{1}{3}, i = 1, 2, 3$. Fiecare din aceste mulţimi va fi împărţită în două submulţimi: $B_i = N_i \cup A_i$, unde N_i corespunde evenimentului "din urna i s-a extras o bilă neagră" şi A_i corespunde evenimentului "din urna i s-a extras o bilă albă".

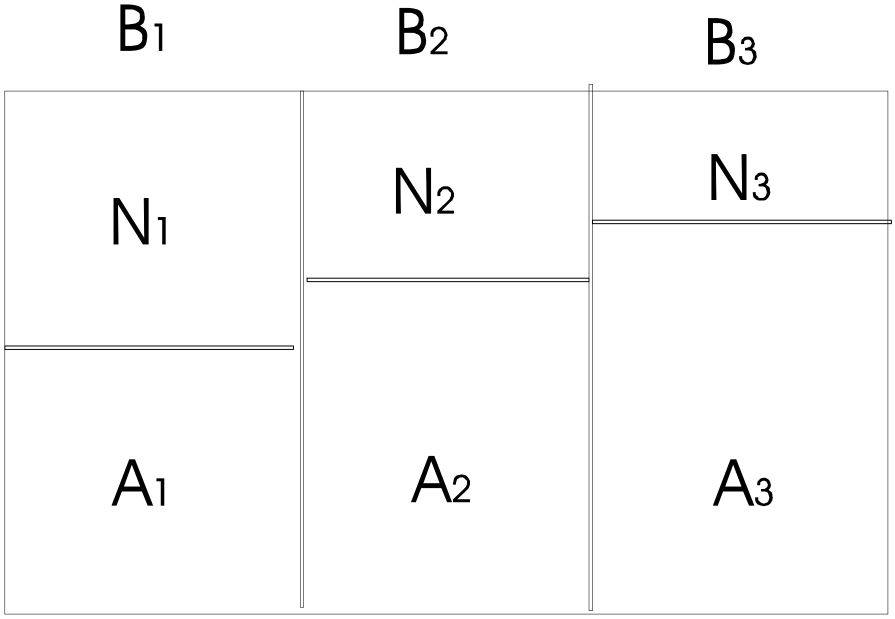


FIGURA 3. Un model probabilist corespunzător celor trei urne

Trebuie să avem

$$P(A_1) = P(N_1),$$
$$P(A_2) = 2P(N_2),$$
$$P(A_3) = 3P(N_3),$$

pentru a reflecta raportul dintre bilele albe şi cele negre din fiecare urnă. Acest raport se traduce în şansele corespunzătoare extragerii unei bile albe sau negre. Atunci putem calcula precis

$$P(A_1) = \frac{1}{3}\frac{1}{2} = \frac{1}{6},\ P(A_2) = \frac{1}{3}\frac{2}{3} = \frac{2}{9},\ P(A_3) = \frac{1}{3}\frac{3}{4} = \frac{1}{4}.$$

Rezultă că, notând A evenimentul "s-a extras o bilă albă", vom avea $A = A_1 \cup A_2 \cup A_3$ şi $P(A) = P(A_1) + P(A_2) + P(A_3) = \frac{1}{6} + \frac{2}{9} + \frac{1}{4} = \frac{6+8+9}{36} = \frac{23}{36}$. Atunci răspunsul la problema pusă este

$$P(B_3 \mid A) = \frac{P(A_3)}{P(A)} = \frac{\frac{1}{4}}{\frac{23}{36}} = \frac{9}{23}. \quad \square$$

Observaţia 2.1. *În modelul construit, nu am pus în evidenţă explicit $\sigma-$algebra $\mathcal{F}$. Ar putea fi $\mathcal{B}(\Omega)$, familia părţilor boreliene din pătratul Ω. Dar pentru nevoile stricte ale problemei noastre putem, la fel de bine, defini $\mathcal{F} = a(A_1, N_1, A_2, N_2, A_3, N_3)$, algebra de părţi generată de partiţia lui Ω formată din mulţimile $A_1, N_1, A_2, N_2, A_3, N_3$. Strict vorbind, noi am neglijat a preciza cui aparţine fiecare frontieră de mulţime. O dată ce avem imaginea, am putea să schimbăm puţin definiţia mulţimilor în felul următor: definim $A_1, N_1, A_2, N_2, A_3, N_3$*

drept mulţimile deschise în plan, reprezentate în figură, fără frontiere. Apoi definim Ω drept reuniunea acestor mulţimi.

Formula lui Bayes.

Problema anterioară a pus în evidenţă următoarea situaţie: spaţiul Ω este împărţit în trei mulţimi disjuncte $\Omega = B_1 \cup B_2 \cup B_3$ ale căror probabilităţi sunt cunoscute. Pentru o mulţime $A \subset \Omega$ se ştiu probabilităţile $P(A \mid B_i)$, $i = 1, 2, 3$. Se cere a se determina $P(B_3 \mid A)$. Într-un cadru general, această problemă capătă răspuns prin aşa numita formulă a lui Bayes cuprinsă în lema care urmează.

Lema 2.1. *Fie $(\Omega, \mathcal{F}, P)$ un spaţiu probabilizat şi $\{B_1, ..., B_n\} \subset \mathcal{F}$ o partiţie a lui Ω, astfel că $P(B_i) > 0$, pentru orice $i = 1, ..., n$. Dacă $A \in \mathcal{F}$ este o mulţime astfel încât $P(A) > 0$, atunci are loc formula*

$$P(B_i \mid A) = \frac{P(A \mid B_i)\, P(B_i)}{\sum\limits_{j=1}^{n} P(B_j)\, P(A \mid B_j)}.$$

Demonstraţie. Este suficient să constatăm că numărătorul şi numitorul din partea dreaptă reprezintă următoarele probabilităţi:

$$P(A \cap B_i) = P(A \mid B_i)\, P(B_i),$$
$$P(A) = \sum_{j=1}^{n} P(A \cap B_j) = \sum_{j=1}^{n} P(B_j)\, P(A \mid B_j). \;\square$$

Un exemplu din medicină.

Se ştie că 1% din persoanele dintr-un grup suferă de o anumită boală. Există un test de analiză a sângelui care este semnificativ pentru boala respectivă dând următoarele rezultate: aplicat unor bolnavi, 95% dintre ei obţin rezultatul pozitiv, iar aplicat unor persoane sănătoase, numai 2% din ele obţin rezultatul pozitiv.

1^0 Presupunem că toţi indivizii din grupul dat ar fi supuşi testului. Se pune problema de a se determina probabilitatea ca un individ găsit pozitiv în urma testului să fie într-adevăr bolnav.

2^0 Presupunem că doctorul face în prealabil o examinare a pacienţilor şi separă un grup de suspecţi dintre care 30% sunt bolnavi. Se supun testului persoanele din acest grup şi se cere să se determine probabilitatea ca o persoană ce iese pozitiv la test, să fie într-adevăr bolnavă.

Soluţie. 1^0 Vom aplica formula lui Bayes. Considerăm spaţiul probabilizat $(\Omega, \mathcal{F}, P)$ în care Ω reprezintă mulţimea tuturor indivizilor din grup. Notăm cu B_1 mulţimea indivizilor sănătoşi şi cu B_2 mulţimea indivizilor bolnavi. Aceste două mulţimi formează o partiţie a lui Ω şi cunoaştem că $P(B_1) = 0,99$ şi $P(B_2) = 0,01$. Notăm apoi cu A mulţimea indivizilor ce obţin testul pozitiv. Avem $P(A \mid B_1) = 0,02$

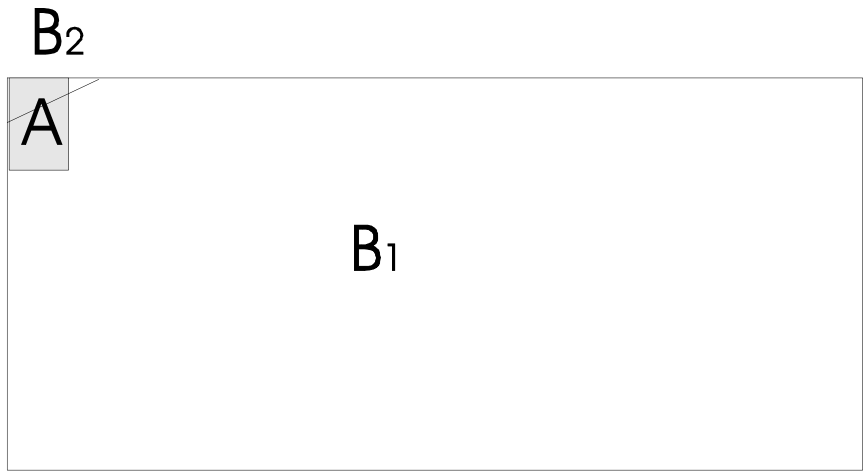


FIGURA 4. Modelul probabilist corespunzător cazului 1^0

şi $P\left(A \mid B_2\right) = 0,95$. Formula lui Bayes ne dă

$$\begin{aligned} P\left(B_2 \mid A\right) &= \tfrac{P(A|B_2)\cdot P(B2)}{P(B_1)P(A|B_1)+P(B_2)P(A|B_2)} = \\ &= \tfrac{0,95\cdot 0,01}{0,99\cdot 0,02+0,01\cdot 0,95} = \tfrac{95}{198+95} = 0,32. \end{aligned}$$

După cum se vede testul nu are relevanţă în contextul dat.

2^0 Pentru analiza grupului selectat de medic se construieşte un model similar celui anterior. De data aceasta vom avea $P\left(B_1\right) = 0,70$ şi $P\left(B_2\right) = 0,30$. Testul fiind acelaşi avem în continuare $P\left(A \mid B_1\right) = 0,02$ şi $P\left(A \mid B_2\right) = 0,95$. De data aceasta rezultatul este

$$P\left(B_2 \mid A\right) = \frac{0,95 \cdot 0,30}{0,70 \cdot 0,02 + 0,30 \cdot 0,95} = \frac{285}{14 + 285} = 0,953.$$

În aceste condiţii testul este semnificativ şi poate ajuta la diagnosticarea bolnavilor. $\square$

Pentru a înţelege de ce în cazul 1^0 nu este relevant testul, am reprezentat mulţimea Ω prin dreptunghiul mare din figura 4. Mulţimea B_2 este reprezentată de triunghiul din colţul de sus stânga. Ea este despărţită de mulţimea B_1, ce constitue restul dreptunghiului printr-o linie. Mulţimea A este reprezentată de dreptunghiul mic, cenuşiu. Ariile acestor mlţimi sunt proporţionale cu probabilităţile lor. Mulţimea B_2 reprezintă $\frac{1}{100}$ din dreptunghiul mare iar mulţimea $A \cap B_1$ reprezintă $\frac{2}{100} \cdot \frac{99}{100} \approx \frac{2}{100}$ din dreptunghiul mare. Deci aria lui $A \cap B_1$ este de două ori mai mare ca aria lui B_2. Cum se vede din figură, cea mai mare parte a mulţimii A se află inclusă în B_1. Aceasta reflectă faptul că printre persoanele ce obţin un rezultat pozitiv la test, cei mai mulţi sunt sănătoşi.

În cazul 2^0 mulţimea B_2 fiind mult mai mare, rezultă că şi mulţimea $A \cap B_2$, care reprezintă 95% din B_2, va fi mare. Deci, de această dată cea mai mare parte din mulţimea A se află inclusă în B_2. Am reprezentat în

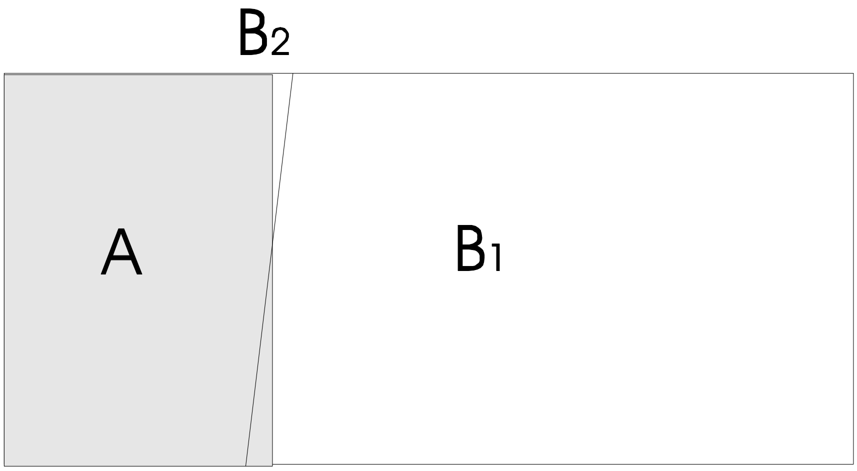


FIGURA 5. Modelul probabilist corespunzător cazului 2^0

figura 5 mulţimea A tot printr-un dreptunghi cenuşiu, iar B_1, B_2 prin trapeze.

2. Independenţă

In această secţiune $(\Omega, \mathcal{F}, P)$ va fi un spaţiu probabilizat fix.

2.1. Independenţa a două evenimente.

Definiţia 2.1. *Vom spune că două evenimente distincte* $A, B \in \mathcal{F}$ *sunt* independente *dacă are loc relaţia* $P(A \cap B) = P(A) P(B)$.

Înainte de a comenta această definiţie vom preciza o denumire. Se numesc *evenimente triviale* acele evenimente care sau au probabilitate nulă sau complementara lor este de probabilitate nulă. Astfel de evenimente nu joacă un rol important în cazul modelelor discrete de care ne ocupăm în prima parte a acestui curs. Ele apar însă în mod natural atunci când formalizăm noţiunea de independenţă. De asemenea au un rol esenţial în cazul spaţiilor probabilizate continue.

1. Observăm că dacă $A \in \mathcal{F}$ are proprietatea că A şi A^c sunt independente, atunci avem $0 = P(A \cap A^c) = P(A) P(A^c)$, ceea ce arată că A este trivial.

2. Un eveniment trivial este independent de orice alt eveniment. Mai precis, dacă $A, B \in \mathcal{F}, A \neq B$ şi A este trivial, atunci A şi B sunt independente.

DEMONSTRAŢIE. Dacă $P(A) = 0$, evident avem şi $P(A \cap B) = 0$, ceea ce conduce la verificarea relaţiei de independenţă. Dacă $P(A) = 1$, înseamnă că $P(A^c) = 0$ şi atunci putem scrie

$$P(B) = P(B \cap A) + P(B \cap A^c) = P(B \cap A),$$

care conduce iarăşi la relaţia de independenţă. $\square$

3. Se observă că dacă A şi B sunt independente atunci putem face calculul

$$P(A \cap B^c) = P(A) - P(A \cap B) = P(A) - P(A)P(B) = P(A)P(B^c),$$

care arată că A şi B^c sunt şi ele independente. La fel se deduce că fiecare din perechile (A^c, B) şi (A^c, B^c) sunt independente.

4. Dacă $P(A) > 0$, atunci relaţia de independenţă se poate scrie sub forma $P(B/A) = P(B)$. Aceasta este forma utilă în aplicaţii. Ea exprimă faptul că evenimentul B se petrece cu aceeaşi probabilitate în prezenţa lui A ca şi în general. Dacă sunt satisfăcute fiecare din relaţiile $P(A) > 0$, $P(A^c) > 0$, $P(B) > 0$, $P(B^c) > 0$, atunci relaţia de independenţă se poate scrie în mod simetric sub fiecare din formele:

$$\begin{gathered} P(B/A) = P(B), P(A/B) = P(A), \\ P(B/A^c) = P(B), P(A^c/B) = P(A^c), \\ P(B^c/A) = P(B^c), P(A/B^c) = P(A), \\ P(B^c/A^c) = P(B^c), P(A^c/B^c) = P(A^c). \end{gathered}$$

5. De asemenea, relaţia $P(B/A) = P(B/A^c)$ este şi ea echivalentă cu independenţa celor două mulţimi. De fapt în aplicaţii concrete această formă de exprimare matematică a independenţei este cea mai sugestivă, pentru că ea prinde ideea limbajului comun privitor la independenţă.

Exemplul 1.

Fie Ω o mulţime care reprezintă o colectivitate umană mare şi pe care introducem probabilitatea şanselor egale $P(\omega) = \frac{1}{card\Omega}$. Fie A mulţimea indivizilor cu ochii albaştri şi B mulţimea indivizilor mai înalţi de 1,70. O ipoteză pe care simţul comun ne-o sugerează este că nu există nici o legătură între culoarea ochilor şi înălţimea individului. Deci putem scrie că proporţia indivizilor mai înalţi de 1,70 printre cei cu ochii albaştri este aceeaşi ca printre cei cu o altă culoare a ochilor:

$$\frac{card(B \cap A)}{cardA} = \frac{card(B \cap A^c)}{cardA^c},$$

adică $P(B/A) = P(B/A^c)$. Cu alte cuvinte, faptul că nu există legătură între culoarea ochilor şi înălţime se exprimă prin independenţa evenimentelor A şi B. Dacă ipoteza făcută este corectă, relaţia aceasta ar trebui să fie verificată pentru orice colectivitate mare.

Exemplul 2.

Dintr-un pachet complet de 52 cărţi de joc se scoate o carte la întâmplare. Să verificăm că evenimentele „a ieşit o carte de treflă” şi „a ieşit un as” sunt independente. Primul eveniment are probabilitatea $\frac{13}{52}$, în timp ce al doilea are probabilitatea $\frac{4}{52}$, iar intersecţia lor are probabilitatea $\frac{1}{52}$. Evident avem $\frac{13}{52} \times \frac{4}{52} = \frac{1}{52}$. Este un rezultat pe care îl aşteptam: nu există legătură între culoarea cărţii şi numărul ei.

Exemplul 3.

O centrală termică se compune în esenţă din două componente, un boiler şi o pompă. Probabilitatea ca să se defecteze biolerul în trei ani de funcţionare este de $0,005$, iar probabilitatea ca să se defecteze pompa în acelaşi interval de timp este de $0,01$. Care este probabilitatea ca centrala să funţioneze neîntrerupt trei ani?

Soluţie. Defectarea boilerului şi defectarea pompei constituie evenimente independente. Pentru a modela problema şi a vedea clar acest lucru, considerăm că avem un număr n de centrale pentru care imaginăm următorul experiment.

Centralele sunt puse în funcţiune pentru trei ani şi la fiecare defecţiune este înlocuită imediat componenta defectă. La sfârşitul celor trei ani vom avea o mulţime A constituită din centralele care au schimbat boilerul şi mulţimea B a centralelor care au schimbat pompa. Notăm cu Ω mulţimea centralelor şi dacă n este foarte mare trebuie să constatăm că are loc egalitatea

$$\frac{card\,(A\cap B)}{card\,(B)}=\frac{card\,(A\cap B^c)}{card\,(B^c)}=\frac{card\,(A)}{card\,(\Omega)}=0,005,$$

adică proporţia de centrale cărora li s-a defectat boilerul printre centralele cărora li s-a defectat pompa este aceeaşi cu proporţia de centrale cărora li s-a defectat boilerul printre toate centralele. Această relaţie exprimă faptul că nu există nici o legătură între defectarea unui boiler şi defectarea unei pompe. Înţelegem prin defectare orice disfuncţionalitate care scoate componenta din parametrii de funcţionare normală. Presupunem că o supraveghere atentă permite depistarea imediată a oricărei defecţiuni. (De exemplu, excludem cazurile de tipul: pompa este defectă puţin, fără a se observa şi aceasta duce la defectarea boilerului.) De aceea funcţionarea unui boiler se face la fel în prezenţa oricărei pompe şi funcţionarea oricărei pompe este la fel în prezenţa oricărui boiler.

Atunci, probabilitatea ca să funcţioneze atât boilerul cât şi pompa este $P\left(A^c\cap B^c\right)=P\left(A^c\right)P\left(B^c\right)=0,995\times 0,99=0,985.\square$

2.2. Independenţa mai multor evenimente.

Definiţia 2.2. *Fie $\mathcal{U}\subset\mathcal{F}$ o familie de evenimente. Spunem că evenimentele din această familie sunt* independente *dacă pentru orice familie finită $\mathcal{K}\subset\mathcal{U}$ este verificată relaţia*

$$P\left(\bigcap_{A\in\mathcal{K}}A\right)=\prod_{A\in\mathcal{K}}P\left(A\right).$$

Noţiunea de familie de evenimente independente este o noţiune ce se referă global la ansamblul familiei.

1. Se observă că fiind dată o familie de evenimente independente, atunci orice subfamilie este constituită tot din evenimente independente.

2. Ţinând cont că în definiţia de mai sus apar doar relaţii ce privesc un număr finit de evenimente, deducem că independenţa evenimentelor dintr-o familie este caracterizată prin independenţa evenimentelor oricărei subfamilii finite.

3. Fie $\mathcal{U}$ o familie de evenimente independente şi $\mathcal{V}$ o familie de evenimente triviale. Atunci $\mathcal{W} = \mathcal{U} \cup \mathcal{V}$ este o familie de evenimente independente.

4. Fie $\mathcal{U} \subset \mathcal{F}$ şi să notăm $\mathcal{U}'$ familia evenimentelor netriviale din $\mathcal{U}$. Atunci evenimentele din $\mathcal{U}$ sunt independente dacă şi numai dacă evenimentele din $\mathcal{U}'$ sunt independente.

Următoarea propoziţie arată un gen de transformare pe care o putem face cu o familie de evenimente independente, păstrând independenţa.

Propoziţia 2.1. *Fie $\mathcal{U}$ o familie de evenimente independente. Dacă $\mathcal{V} \subset \mathcal{U}$ este o subfamilie arbitrară şi definim $\mathcal{V}' = \{A^c / A \in \mathcal{U} \backslash \mathcal{V}\}$, atunci evenimentele din familia $\mathcal{V} \cup \mathcal{V}'$ sunt de asemenea independente.*

Demonstraţie. Mai întâi vom examina cazul în care mulţimea $\mathcal{U} \backslash \mathcal{V}$ se reduce la un singur eveniment, să zicem $\mathcal{U} \backslash \mathcal{V} = \{A_0\}$, ceea ce înseamnă că $\mathcal{V}' = \{A_0^c\}$. Avem de verificat că pentru orice familie finită, $\mathcal{K} \subset \mathcal{V} \cup \mathcal{V}'$, are loc relaţia

$$P\left(\bigcap_{A \in \mathcal{K}} A\right) = \prod_{A \in \mathcal{K}} P(A).$$

Dacă $A_0^c \notin \mathcal{K}$, relaţia este evidentă, pentru că se reduce la relaţia corespunzătoare familiei iniţiale. Să vedem acum cazul $A_0^c \in \mathcal{K}$. Notăm $\mathcal{L} := \mathcal{K} \backslash \{A_0^c\}, \mathcal{I} := \mathcal{L} \cup \{A_0\}$ şi

$$C = \bigcap_{A \in \mathcal{L}} A.$$

Datorită independenţei familiei iniţiale putem scrie

$$P(C) = \prod_{A \in \mathcal{L}} P(A).$$

Apoi se deduce

$$P(A_0 \cap C) = \prod_{A \in \mathcal{I}} P(A) = P(A_0) \prod_{A \in \mathcal{L}} P(A) = P(A_0) P(C).$$

Această relaţie spune că A_0 şi C sunt independente, prin urmare A_0^c şi C sunt de asemenea independente. Deci putem scrie

$$P\left(\bigcap_{A \in \mathcal{K}} A\right) = P(A_0^c \cap C) = P(A_0^c) P(C) = P(A_0^c) \prod_{A \in \mathcal{L}} P(A) = \prod_{A \in \mathcal{K}} P(A),$$

ceea ce încheie verificarea cazului în care mulţimea $\mathcal{U} \backslash \mathcal{V}$ se reduce la un punct.

Pentru situaţia în care mulţimea $\mathcal{U} \backslash \mathcal{V}$ este finită este suficient să utilizăm cazul anterior împreună cu un argument de inducţie. În fine, cazul general este obţinut ţinând cont că de fapt independenţa se verifică printr-o relaţie ce se scrie pentru familii finite. $\square$

Independenţa extragerilor cu întoarcere.

Vom reveni la cazul unei urne cu m bile numerotate din care se fac n extrageri cu revenire. Modelul constă din mulţimea $M = \{1, ..., m\}$, care reprezintă bilele, mulţimea tuturor rezultatelor posibile este descrisă de $\Omega = M^n$, iar probabilitatea ce se introduce este cea a şanselor egale, în care $P(\{(x_1, ..., x_n)\}) = \frac{1}{m^n}$. Evenimentul "la a $l-$ a extragerea a ieşit bila cu numărul i " este descris de mulţimea $A_{l,i} = \{(x_1, ..., x_n) \in \Omega / x_l = i\}$. Vom nota $\mathcal{A}_l = \{A_{l,i} / i = 1, ..., m\}$. Pentru fiecare $l = 1, ..., n$, familia $\mathcal{A}_l$ constituie o partiţie a lui Ω.

Vom arăta că partiţiile $\mathcal{A}_1, ..., \mathcal{A}_n$ sunt independente, în sensul că oricum am lua elementele $i_l \in M, l = 1, ..., n$, familia $\{A_{l,i_l} / l = 1, ...n\}$ este formată din evenimente independente. Pentru aceasta, calculăm mai întâi cardinalul unei mulţimii $A_{l,i}$. Deoarece numai componenta l este fixată pentru punctele din mulţimea $A_{l,i}$, în timp ce celelalte $n-1$ componente parcurg liber pe M, rezultă că mulţimea $A_{l,i}$ poate fi pusă în bijecţie cu mulţimea M^{n-1}. Anume, un punct generic din $A_{l,i}$ este de tipul $(x_1, ..., x_{l-1}, i, x_{l+1}, ..., x_n)$ şi lui îi corespunde $(x_1, ..., x_{l-1}, x_{l+1}, ..., x_n)$, care este un punct general din M^{n-1}.

Tragem concluzia că mulţimea $A_{l,i}$ conţine exact m^{n-1} puncte. Deci $P(A_{l,i}) = \frac{m^{n-1}}{m^n} = \frac{1}{m}$, pentru orice l şi i. Să considerăm acum un număr $k \leq n$ şi k mulţimi din partiţii diferite: $A_{l_1,i_1}, ..., A_{l_k,i_k}$. Intersecţia lor poate fi descrisă astfel

$$A_{l_1,i_1} \cap ... \cap A_{l_k,i_k} = \{(x_1, ..., x_n) \in \Omega / x_{l_1} = i_1, ..., x_{l_k} = i_k\}.$$

După cum se vede punctele din această mulţime au k componente fixate, iar restul de $n-k$ sunt libere în M. Rezultă că această mulţime poate fi pusă în bijecţie cu M^{n-k}, deci cardinalul ei este m^{n-k}. Prin urmare

$$P(A_{l_1,i_1} \cap ... \cap A_{l_k,i_k}) = \frac{m^{n-k}}{m^n} = \frac{1}{m^k} = P(A_{l_1,i_1}) ... P(A_{l_k,i_k}),$$

ceea ce probează independenţa partiţiilor considerate.

În *propoziţia* 4.5 ne vom ocupa în mod special de partiţii independente demonstrând proprietăţi de independenţă suplimentare.

Regula jucătorului.

Vom presupune că şansele de reuşită într-un anumit joc de noroc sunt de una din N, unde N este un număr natural. Aşa numita " regulă a jucătorului" spune că dacă cineva participă la joc de un număr de ori mai mare sau egal decât $\frac{2}{3}N$, atunci probabilitatea de a câştiga cel puţin o dată devine mai mare sau egală ca $\frac{1}{2}$. Ne punem problema de a examina cât de corectă este regula, adică de a afla care este eroarea pentru "regula jucătorului".

Souţie. Vom presupune că $(\Omega, \mathcal{F}, P)$ este un spaţiu probabilizat pe care există n evenimente independente $A_1, ..., A_n$, cu aceeaşi probabilitate $P(A_1) = ... = P(A_n) = \frac{1}{N}$. Presupunem că A_i modelează reuşita la cel de-al $i-$lea joc. Construcţia concretă a unui astfel de spaţiu probabilizat se poate realiza făcând o analogie între participarea la joc şi extragerea unei bile dintr-o urnă cu N bile. Presupunem că din cele N bile una este roşie iar restul sunt negre. Şansele de a extrage bila roşie sunt de una la N. Deci putem echivala reuşita la joc cu faptul că în urma unei extrageri din urnă a ieşit bila roşie. O serie de n extrageri cu revenire corespunde la n participări la jocul dat.

Identificăm mulţimea bilelor cu mulţimea primelor N numere naturale, pe care o notăm $M = \{1, ..., N\}$ şi presupunem că bila roşie corespunde numărului 1. Mulţimea rezultatelor posibil a fi obţinute după n jocuri este descrisă de $\Omega = M^n$. Evenimentul „la jocul al $l-$lea jucătorul a câştigat", sau spus altfel "la cea de a $l-$a extragere a ieşit bila roşie", este descris de mulţimea $A_l = \{(x_1, ..., x_n) \in \Omega / x_l = 1\}$. Ţinând cont de cele demonstrate în paragraful precedent privind independenţa extragerilor cu întoarcere, rezultă că evenimentele $A_1, ..., A_n$ sunt independente şi fiecare are probabilitatea $P(A_l) = \frac{1}{N}$.

Reuşita la cel puţin un joc din n este reprezentată de evenimentul

$$A_1 \cup ... \cup A_n.$$

Nereuşita la nici un joc din cele n este reprezentată de evenimentul complementar

$$A_1^c \cap ... \cap A_n^c,$$

şi pentru acesta avem formula

$$P(A_1^c \cap ... \cap A_n^c) = P(A_1^c) \cdot ... \cdot P(A_n^c) = \left(1 - \frac{1}{N}\right)^n.$$

Funcţia care apare în această exprimare, $f(x) = \left(1 - \frac{1}{N}\right)^x$, este descrescătoare pe $[1, \infty)$ şi $\lim_{x \to \infty} f(x) = 0$. Ne interesează să determinăm cea mai mică valoare a lui $n \in \mathbf{N}^*$ pentru care această probabilitate devine mai mică dacât $\frac{1}{2}$, adică acel număr care satisface condiţiile

$$\left(1 - \frac{1}{N}\right)^n \leq \frac{1}{2} < \left(1 - \frac{1}{N}\right)^{n-1}.$$

Scriem $\frac{1}{2} = \exp - \ln 2$ şi aplicăm apoi *lema* **??** pentru a minora, respectiv majora, cele două expresii care încadrează acest număr. Obţinem

$$\exp - \frac{n}{N}\left(1 + \frac{7}{12N}\right) \leq \exp - \ln 2 \leq \exp - \frac{(n-1)}{N}.$$

Aceste inegalităţi se mai scriu şi sub forma

$$N \ln 2 - \frac{7n}{12N} \leq n < N \ln 2 + 1,$$

sau, notând partea întreagă a lui $N \ln 2$ cu $n_0 = [N \ln 2]$, putem scrie

$$n_0 - \frac{7n}{12N} \leq n \leq n_0 + 1,$$

ceea ce implică $n_0 \leq n \leq n_0+1$. Putem trage concluzia că numărul căutat este aproximat cu o eroare de o unitate de $n_0 = [N \ln 2]$. Deoarece diferenţa $N \ln 2 - \frac{2}{3}N$ este mică raportată la N,

$$\frac{N \ln 2 - \frac{2}{3}N}{N} = \ln 2 - \frac{2}{3} \approx 0,0264,$$

putem spune că eroarea pe care o dă regula jucătorului este mică raportată la N. În valoare absolută ea poate fi mare - de ordinul $0,0264N$.□

Pentru o ruletă standard avem 38 de sectoare, deci şansele de a ieşi un număr ales sunt $\frac{1}{38}$. Cu $N = 38$ eroarea este $0,0264 \times 38 \approx 1$. Regula spune că pentru $38 \times \frac{2}{3} = 25$, eventual 26, de participări la joc probabilitatea de a câştiga măcar o dată devine mai mare decât $0,5$.

Pentru un zar, şansele de a ieşi o anumită faţă sunt $\frac{1}{6}$. Cu $N = 6$ eroarea este $0,0264 \times 6 \approx 0,15$. Putem aplica regula jucătorului şi deducem că probabilitatea de a ieşi o faţă anume din $6 \times \frac{2}{3} = 4$ aruncări este mai mare de $0,5$.

Observăm că probabilităţile ce se cer calculate în problema lui Mère se pretează la aplicarea regulei jucătorului.

2.3. Independenţa a trei evenimente. Conform definiţiei, independenţa a trei evenimente $A_1, A_2, A_3 \in \mathcal{F}$ este asigurată dacă sunt îndeplinite condiţiile următoare:

$$\begin{array}{rl}
P(A_1 \cap A_2) = P(A_1) P(A_2), & (*) \\
P(A_1 \cap A_3) = P(A_1) P(A_3), & (**) \\
P(A_2 \cap A_3) = P(A_2) P(A_3), & (***) \\
P(A_1 \cap A_2 \cap A_3) = P(A_1) P(A_2) P(A_3). & (\#)
\end{array}$$

Se poate pune întrebarea dacă nu cumva un număr mai mic dintre aceste relaţii sunt suficiente pentru a implica valabilitatea tuturor. Următoarele două contraexemple arată că, în general, nu este posibil aşa ceva.

Exemplul 1.

Presupunem că $\Omega = \{1, 2, 3, 4\}$ şi $P(1) = P(2) = P(3) = P(4) = \frac{1}{4}$. Mulţimile următoare

$$A_1 = \{1, 2\}, A_2 = \{1, 3\}, A_3 = \{1, 4\},$$

verifică relaţiile $(*), (**), (***)$, dar nu verifică relaţia $(\#)$.

Exemplul 2.

Presupunem că $\Omega = \{1, 2, 3, 4, 5, 6, 7\}$ şi P este definită de

$$\begin{gathered}
P(1) = P(2) = \tfrac{1}{4}, \\
P(3) = P(4) = P(5) = \tfrac{1}{8}, \\
P(6) = \tfrac{1}{12}, \quad P(7) = \tfrac{1}{24}.
\end{gathered}$$

Notăm $A_1 = \{1, 2, 3, 4\}, A_2 = \{3, 4, 6\}, A_3 = \{2, 4, 5\}$. Aceste mulţimi verifică relaţiile $(*), (**)$ şi $(\#)$, dar nu verifică relaţia $(***)$.

2.4. Independenţa familiilor indexate. Independenţa aşa cum am introdus-o mai sus este în forma utilizată curent în aplicaţiile elementare. În literatură apare de obicei o definiţie mai abstractă, dar şi mai generală, care se referă la familii indexate de evenimente, pe care o prezentăm mai jos.

Definiţia 2.3. *Fiind dată o familie $(A_i)_{i\in I}$ de evenimente din $\mathcal{F}$ vom spune că venimentele familiei sunt independente dacă este satisfăcută relaţia*

$$P\left(\bigcap_{i\in J} A_i\right) = \prod_{i\in J} P(A_i),$$

pentru orice mulţime finită $J \subset I$.

Pentru a face deosebire, atunci când este cazul, vom spune că independenţa dată de definiţia ... este în sens restrâns iar independenţa dată de definiţia ... este în sens extins. Principala deosebire dintre cele două definiţii constă în faptul că la definiţia în sens extins putem avea o aceeaşi muţime pusă în familie cu doi indici diferiţi. Această posibilitate lasă loc la a discuta despre evenimente independente de ele însele. Dacă în familia de evenimente independente $(A_i)_{i\in I}$ avem doi indici distincţi $i \neq j$ pentru care $A_i = A_j$, atunci putem scrie $P(A_i) = P(A_i \cap A_j) = P(A_i)P(A_j) = P(A_i)^2$ şi prin urmare evenimentul corespunzător, A_i, trebuie să fie trivial. Următoarea lemă arată mai precis care este raportul dintre cele două noţiuni de independenţă.

Lema 2.2. *Fie $(A_i)_{i\in I}$ o familie de evenimente şi $\mathcal{U} = \{A_i / i \in I\}$ mulţimea evenimentelor din familie în care uităm indexarea. Atunci independenţa evenimentelor familiei $(A_i)_{i\in I}$ este echivalentă cu următoarele două proprietăţi: (i) evenimentele din $\mathcal{U}$ sunt independente în sens restrâns, (ii) dacă există doi indici $i \neq j$ astfel ca $A_i = A_j$, atunci evenimentul A_i este trivial.*

Demonstraţie. Independenţa evenimentelor familiei $(A_i)_{i\in I}$ implică evident proprietăţile (i) şi (ii). Reciproc, presupunem că sunt satisfăcute proprietăţile (i) şi (ii). Fie $J \subset I$, finită, pentru care vrem să verificăm relaţia

$$P\left(\bigcap_{i\in J} A_i\right) = \prod_{i\in J} P(A_i). \qquad (*)$$

Pentru aceasta notăm cu J_0 mulţimea indicilor care corespund unor evenimente triviale, adică

$$J_0 = \{i \in J / P(A_i) P(A_i^c) = 0\},$$

şi punem $J_1 = J \backslash J_0$. Rezultă că evenimentele $\{A_i / i \in J_1\}$ sunt toate distincte, ca urmare a proprietăţii (ii). Proprietatea (i) ne dă atunci

$$P\left(\bigcap_{i\in J_1} A_i\right) = \prod_{i\in J_1} P\left(A_i\right). \qquad (**)$$

Dacă mulţimea J_0 este vidă atunci această relaţie devine (*) şi s-a încheiat demonstraţia. Dacă J_0 este nevidă, continuăm raţionamentul despărţit în două cazuri.

Cazul 1. Presupunem că există un indice $i_0 \in J_0$ astfel ca $P\left(A_{i_0}\right) = 0$. Atunci incluziunea $\bigcap_{i\in J} A_i \subset A_{i_0}$ ne spune că $P\left(\bigcap_{i\in J} A_i\right) = 0$ şi de aici se deduce imediat relaţia (*).

Cazul 2. Presupunem că $P\left(A_i\right) = 1$ pentru orice $i \in J_0$. Atunci utilizăm repetat proprietatea 2. de la secţiunea „Independenţa a două evenimente" obţinând

$$P\left(\bigcap_{i\in J} A_i\right) = P\left(\bigcap_{i\in J_1} A_i\right).$$

Cum în cazul de faţă avem $\prod_{i\in J_0} P\left(A_i\right) = 1$, din relaţia (**) obţinem imediat relaţia (*). $\square$

Avantajul definiţiei independenţei cu familii de evenimente indexate apare în faptul că este mai apropiată de formularea independenţei pentru $\sigma-$algebre, care se face întotdeauna în termeni de familii indexate. Fenomenul este pregnant în cazul spaţiilor probabilizate nediscrete. De exemplu atunci când se discută despre evenimentele $\sigma-$algebrelor de la infinit. Acestea sunt toate triviale şi aceasta rezultă după ce se demonstrează că sunt independente de ele însele. Pentru modelele discrete în schimb, evenimentele triviale sunt lipsite de orice interes.

2.5. Independenţa $\sigma-$algebrelor. În aceast paragraf vom introduce importanta noţiune de independenţă $\sigma-$algebrelor.

Definiţia 2.4. *Fiind date $\mathcal{F}_1, ..., \mathcal{F}_n$, $\sigma-$algebre incluse în $\mathcal{F}$, vom spune că ele sunt independente dacă pentru orice alegere de evenimente $A_i \in \mathcal{F}_i$, $i = 1, ..., n$, acestea sunt independente (în sens extins). Spunem că $\sigma-$algebrele dintr-o familie arbitrară $\left(\mathcal{F}_i\right)_{i\in I}$ sunt independente dacă $\sigma-$algebrele din orice subfamilie finită sunt independente.*

Dacă $\mathcal{G}$ şi $\mathcal{H}$ sunt două $\sigma-$algebre independente, observăm că atunci intersecţia lor $\mathcal{G} \cap \mathcal{H}$ constă numai din evenimente triviale.

Observaţia 2.2. *Definiţia independenţei $\sigma-$algebrelor se poate formula şi utilizând noţiunea de independenţă în sens restrâns pentru familii de evenimente. Într-adevăr, dacă presupunem că $\sigma-$algebrele $\mathcal{F}_1, ..., \mathcal{F}_n$ au proprietatea că pentru orice alegere de evenimente $A_i \in \mathcal{F}_i$, $i = 1, ..., n$, acestea sunt independente în sens restrâns, în sensul* *că evenimentele din familia* $\mathcal{U} = \{A_i/i = 1, ..., n\}$, *în care se uită indexarea, sunt independente, atunci* $\sigma-$*algebrele sunt independente în sensul definiţiei de mai sus.*

DEMONSTRAŢIE. Vom remarca mai întâi că rămâne valabilă proprietatea menţionată imediat după definiţia de mai sus: anume intersecţia oricăror două algebre, $\mathcal{F}_i \cap \mathcal{F}_j$, este formată din evenimente triviale. Într-adevăr, fie $A \in \mathcal{F}_i \cap \mathcal{F}_j$. În acest caz evenimentele A şi A^c sunt distincte, $A \in \mathcal{F}_i$ şi $A^c \in \mathcal{F}_j$, deci, ca urmare a independenţei în sens restâns, trebuie să satisfacă relaţia $P(A)P(A^c) = P(\emptyset) = 0$. Aceasta înseamnă că A este trivial.

Atunci se poate utiliza lema anterioară pentru a deduce că orice evenimente $A_i \in \mathcal{F}_i$, $i = 1, ..., n$, care sunt independente în sens restrâns sunt de fapt chiar independente în sens extins. □

După cum se vede din însăşi definiţia dată, independenţa este o noţiune ce priveşte ansamblul evenimentelor celor n $\sigma-$algebre, structurate în grupurile formate de fiecare $\sigma-$algebră. Se observă că fiind dată o familie $(\mathcal{F}_i)_{i \in I}$ de $\sigma-$algebre independente şi sub- $\sigma-$algebrele $\mathcal{H}_i \subset \mathcal{F}_i, i \in I$, rezultă că şi familia $(\mathcal{H}_i)_{i \in I}$ constă din $\sigma-$algebre independente. De asemenea, notăm că dacă avem o familie $(\mathcal{F}_i)_{i \in I}$ de $\sigma-$algebre independente, atunci $\sigma-$algebrele din orice subfamilie, $(\mathcal{F}_i)_{i \in J}$, unde $J \subset I$, sunt tot independente.

Observăm că fiecare eveniment $A \in \mathcal{F}$ generează o algebră, anume $\sigma(A) = \{A, A^c, \emptyset, \Omega\}$. Dacă evenimentele $A_1, ..., A_n$ sunt independente (în sens extins), atunci $\sigma-$algebrele generate de fiecare din aceste evenimente, $\sigma(A_1), ..., \sigma(A_n)$, sunt independente, după cum se poate verifica prin utilizarea *propoziţiei* 2.1. Având această idee în minte putem introduce definiţia următoare: un eveniment A este independent de o $\sigma-$algebră $\mathcal{F}' \subset \mathcal{F}$ dacă A este independent de fiecare element din $\mathcal{F}'$. Aceasta revine la independenţa $\sigma-$algebrelor $\sigma(A)$ şi $\mathcal{F}'$.

Lema 2.3. *Independenţa unor* $\sigma-$*algebre* $\mathcal{F}_1, ..., \mathcal{F}_n$ *este echivalentă cu faptul că, pentru orice alegere de elemente* $A_i \in \mathcal{F}_i$, $i = 1, ..., n$, *este satisfăcută relaţia*

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1) \cdot ... \cdot P(A_n).$$

DEMONSTRAŢIE. Dacă $\sigma-$algebrele sunt independente, rezultă imediat relaţia din enunţ. Reciproc, presupunem că relaţia este satisfăcută pentru orice alegere de evenimente. Fiind date evenimentele $A_i \in \mathcal{F}_i$, $i = 1, ..., n$, vrem să arătăm că ele sunt independente. Fie $J = \{1 \leq i_1 < ... < i_k \leq n\} \subset \{1, ..., n\}$ o submulţime de indici. Trebuie să demonstrăm relaţia

$$P\left(\bigcap_{l=1}^{k} A_{i_l}\right) = P(A_{i_1}) ... P(A_{i_k}).$$

Dar această relaţie este exact relaţia din enunţ scrisă pentru sistemul de evenimente $A'_1, ..., A'_n$, definite în felul următor: $A'_i = A_i$, când $i \in J$ şi $A'_i = \Omega$, când $i \notin J$. Anume, se ţine cont că

$$\bigcap_{l=1}^{k} A_{i_l} = \bigcap_{i=1}^{n} A'_i$$

şi că $P(A'_i) = 1$, dacă $i \notin J$. $\square$

3. Exerciţii

Exerciţiul 2.1. *Într-o pălărie se află trei cartoane: unul are ambele feţe albe, altul are ambele feţe negre şi al treilea are o faţă albă şi una neagră. Se ia unul şi se pune pe masă. Constatând că faţa de deasupra cartonului este neagră se pune problema de a determina probabilitatea ca dosul să fie alb.*

Exerciţiul 2.2. *Se ştie că pentru un examen sunt alcătuite* 27 *de bilete, din care se aprciază ca dificile* 10*. Un student intrat la examen trece să îşi pregătească subiectul în bancă, luând biletul pe care la tras cu sine. Următorul student va trage un bilet dintre cele rămase. Cine are mai multe şanse de a nimeri un bilet dintre cele* 10 *dificile, primul intrat la examen sau al doilea?*

Exerciţiul 2.3. *Pe un spaţiu probabilizat se cunosc cunosc probabilităţile* $P(A_i), i = 1, 2, 3,$ *a trei evenimente disjuncte* A_1, A_2, A_3*. Se mai cunosc probabilităţile* $P(B/A_i), i = 1, 2, 3,$ *ale unui alt eveniment* B *şi se notează* $A = A_1 \cup A_2 \cup A_3$*. Să se exprime* $P(B/A)$ *în funcţie de probabilităţile cunoscute.*

Exerciţiul 2.4. *Doi vânători, A şi B, trag simultan asupra unui urs. Ei găsesc ursul ucis cu un singur glonţ. După vânzarea pielei îşi pun problema cum trebuie împărţită suma obţinută. Ei ştiu că la distanţa şi în condiţiile date A avea şansele de a nimeri dintr-un foc de* 80%*, iar B avea şansele de a nimeri dintr-un foc de* 40%*. Propuneţi o formulă raţională de împărţire a banilor.*

(Indicaţie. Nu este raţional ca banii să fie împărţiţi proporţional cu probabilităţile fiecăruia de a nimeri la ţintă. Un astfel de raţionament, în cazul în care am şti că A ar avea şansele de a nimeri 100%*, iar B de* 50%*, ar conduce la raportul de împărţire a banilor de* 2 *la* 1*. Ori într-o astfel de situaţie, cum ştim că doar un glonţ a nimerit ţinta, este clar că autorul nu poate fi decât A. Rezultă că metoda nu este acceptabilă. O judecată raţională se face considerând probabilităţile condiţionate de evenimentul că s-a găsit un singur glonţ în urs.)*

Exerciţiul 2.5. *Există două legături identice, fiecare a câte* 6 *chei. Fiecare legătură deci deschide aceleaşi broaşte. O uşă are două broaşte ce se deschid cu două chei care sunt în fiecare legătură. La un moment* *dat se pierde o cheie din prima legătură şi două chei din a doua legătură. 1) Care este probabilitatea de a deschide uşa respectivă cu cheile rămase în prima legătură? 2) Care este probabilitatea de a deschide uşa cu cheile rămase în a doua legătură? 3) Care este probabilitatea de a deschide uşa având ambele legături? 4) Care este probabilitatea de a deschide uşa luând la întâmplare una din legăturile de chei?*

Exerciţiul 2.6. *O clădire are două lifturi identice. Probabilitatea de defectare a unui lift în decursul unei zile este de* $0,03$. *După defectarea unui lift celălalt este supus la o sarcină dublă şi prin urmare probabilitatea sa de defectare în decursul zilei următoare este* $0,06$. *Pentru simplificare presupunem că defectările au loc la sfârşitul zilei. Mai presupunem că serviciul de întreţinere nu intervine. Care este atunci probabilitatea ca în decursul unei săptămâni clădirea să rămână fără nici un lift în funcţiune?*

Exerciţiul 2.7. *Aruncăm cu un zar până când iese de două ori la rând aceeaşi faţă. Notăm cu* p_n *probabilitatea ca acest lucru să se petreacă exact după* n *aruncări. Să se determine o formulă pentru* p_n. *Să se determine suma* $p_2 + ... + p_{10}$.

Exerciţiul 2.8. *Trei maşini A,B,C produc respectiv* $60\%, 30\%$ *şi* 10% *din piesele fabricate de o uzină. Maşina A produce* 2% *din piese cu defect. Maşina B produce* 3% *cu defect iar maşina C* 4%. *Se alege la întâmplare o piesă la ieşirea din uzină. 1) Care este probabilitatea ca piesa să fie defectă ? 2) Constatând că piesa are defect, care este probabilitatea ca ea să provină de la maşina C ?*

Exerciţiul 2.9. *O întreprindere produce aparate electronice care sunt perfect funcţionabile cu probabilitatea* $\frac{9}{10}$. *Fiecare aparat este supus testării înainte de livrare. Se constată că orice aparat bun trece testul. Dar din cele cu defecte, numai* $\frac{10}{11}$ *dintre ele sunt depistate ca atare; restul de* $\frac{1}{11}$ *scapă şi sunt acceptate. Să se calculeze probabilitatea evenimentului: 1) „aparatul a trecut testul şi el funcţionează", 2) „aparatul a trecut testul dar nu funcţionează". 3) Să se calculeze probabilitatea ca aparatul să funcţioneze ştiind că el a trecut testul.*

Exerciţiul 2.10. *O uzină fabrică piese, din care* $1,8\%$ *sunt cu defecte. Controlul de calitate are următoarele caracteristici: piesele fără defect sunt acceptate în proporţie de* 97%, *iar cele defecte sunt refuzate în proporţie de* 99%. *1) Care este probabilitatea ca o piesă să fie defectă deşi a fost acceptată? 2) Care este probabilitatea ca să fie o eroare în control? 3) Care este probabilitatea ca în cinci controale consecutive să se producă exact două erori?*

Exerciţiul 2.11. *Studenţii se prepară pentru un examen al cărui subiect va fi propus de unul din membrii unei comisii formate din 3 profesori: A,B,C. Analizând datele din anii trecuţi studenţii evaluează* *cu* $0,35$ *probabilitatea ca subiectul să fie propus de* A, *cu* $0,4$ *probabilitatea ca subiectul să fie propus de* B *şi cu* $0,25$ *probabilitatea ca* C *să propună subiectul. Pe de altă parte, studenţii se tem de un anumit subiect. Ei apreciază că* A *ar propune la examen subiectul în cauză cu probabilitatea* $\frac{1}{10}$, *că* B *ar propune subiectul respectiv cu o probabilitate de* $\frac{2}{5}$, *iar* C *l-ar propune cu probabilitatea* $\frac{41}{50}$. *1) Care este probabilitatea ca subiectul temut să fie dat la examen? 2) Ştiind că s-a dat subiectul nedorit la examen, să se afle probabilitatea ca cel ce a alcătuit subiectele să fie* C.

Exerciţiul 2.12. *Într-un oraş,* 2% *din populaţie este atinsă de o boală contagioasă. Se ştie că în cazul unei întâlniri între o persoană contagioasă şi una sănătoasă există un risc de contaminare de* $0,7$. *Care este riscul de contaminare la care este supusă o persoană sănătoasă ce vine în oraş şi vizitează* 3 *persoane arbitrare?*

(Indicaţie. Mai întâi se va analiza cazul unei singure vizite. Dificultatea modelării provine din faptul că, propriu-zis, sunt două fenomene aleatoare în succesiune: mai întâi nimerirea aleatoare a unui sănătos sau contagios, iar apoi, în cazul vizitării unui contagios, este îmbolnăvirea aleatoare. Pentru a construi un model în acest caz, se imaginează o mulţime mare de vizitatori, de exemplu, persoane venite din alte localităţi. Rezultatele acestor vizite se pot termina în trei feluri: sau a fost vizitată o persoană sănătoasă, sau a fost vizitată o persoană contagioasă dar nu s-a produs contaminarea, sau a fost vizitată o persoană contagioasă şi s-a produs contaminarea. Reprezentăm cu o regiune Ω *din plan mulţimea tuturor acestor rezultate. O submulţime* $A \subset \Omega$ *va reprezenta rezultatele care au corespuns vizitelor la persoane contagioase. O altă submulţime* $B \subset A$ *va reprezenta rezultatele care constau în vizite la persoane contagioase terminate cu contaminare. Mulţimea* A^c *va reprezenta rezultatele care constau în vizite la persoane sănătoase, iar* $A \backslash B$ *va reprezenta rezultatele care constau în vizitele la persoane contagioase dar care nu au dus la contaminare.*

Pentru a modela cazul a 3 vizite se construieşte un alt model în care rezultatele cu care se termină cele trei vizite sunt considerate ca evenimente independente. O astfel de abordare poate fi făcută în cazul unei populaţii foarte mari, caz în care se poate spune cu o bună aproximare că, indiferent care este persoana ce a fost vizitată mai întâi, la cea de a doua vizită există aceleaşi şanse de contaminare. Vezi problema care urmează.)

Exerciţiul 2.13. ** Fie o urnă cu* m *bile, din care* r *sunt roşii. Se fac două extrageri fără revenire. Fie* A_1 *evenimentul „la prima extragere a ieşit obilă roşie" şi* A_2 *evenimentul „la a doua extragere a ieşit o bilă roşie". (i) Arătaţi că cele două evenimente nu sunt independente. (ii) Arătaţi că dacă* $m \to \infty$ *şi* $r \to \infty$, *astfel încât raportul* $\frac{r}{m} = p$ *este constant, atunci la limită se obţine* $P(A_1 \cap A_2) = P(A_1)\, P(A_2)$.

Exerciţiul 2.14. *Trei persoane joacă un joc ce constă în realizarea pe rând, de către fiecare, a următorului experiment: dintr-un pachet complet de cărţi de joc, bine amestecate, se extrage o carte. Dacă cartea extrasă nu este o treflă se consideră că jucătorul respectiv a eşuat. Dacă cartea este o treflă, jucătorul aruncă cu zarul. Dacă îi iese un as el ridică potul. Dacă îi iese alt număr el este eşuat. Care este probabilitatea ca după ce fiecare jucător şi -a jucat experimentul, niciunul să nu fi ridicat potul?*

Exerciţiul 2.15. *O fabrică produce în serie ceasuri cu un proces tehnologic în două faze. Prima fază de fabricaţie produce defectul A la 2% dintre produse, iar faza a doua produce defectul B la 3% din produse.*

(i) Se ia la întâmplare un ceas. Presupunând că defectele în cele două etape se produc independent, se cere probabilitatea ca: a) ceasul să aibă ambele defecte, b) ceasul să nu aibă niciunul din defecte, c) ceasul să aibă exact un defect.

(ii) Se iau pe rând 5 mostre de produs finit. Se presupune că numărul total de ceasuri este aşa de mare încât la fiecare extragere de mostră proporţia de ceasuri defecte este neschimbată. Determinaţi probabilitatea ca cel puţin 4 din cele 5 ceasuri să fie fără defecte.

Exerciţiul 2.16. ** Sultanul îi spune lui Ali Baba: în faţa ta se află două vase identice. Unul conţine r bile roşii, iar celălalt conţine n bile negre. Ai voie să amesteci bilele în cele două vase dar să laşi cel puţin câte o bilă în fiecare vas. Eu voi veni apoi, mă voi apropia de un vas şi voi scoate o bilă. Dacă bila este roşie, vei primi un coş cu rodii, iar dacă bila va fi neagră, vei primi un coş cu mărăcini. Cum amestecă Ali Baba bilele pentru a-şi mări şansele de a obţine rodii? Demonstraţi că există o metodă optimă.*

(Indicaţie. Să zicem că urnele sunt U_1 şi U_2. Se va analiza mai întâi ce se întâmplă cu o urnă, o dată aleasă. Fie U_1 urna aleasă. Probabilitatea de a extrage o bilă roşie este maximă şi egală cu 1 dacă în urna U_1 sunt numai bile roşii. Într-o astfel de situaţie, pentru a mări şansele de extragere a unei bile roşii şi pentru urna U_2, este optim a se pune cât mai multe bile roşii şi în ea. Deci se lasă o singură bilă roşie în urna U_1 şi se pune restul de $r-1$ bile roşii împreună cu cele n bile negre în urna U_2. Pe ansamblul experimentului (alegere urnă şi extragere bilă), probabilitatea de a fi extrasă o bilă roşie în cazul acestei repartizări a bilelor este $p_0 = \frac{1}{2}\left(1 + \frac{r-1}{r-1+n}\right)$.

Raţionamentul de mai sus nu este complet. Trebuie arătat că pentru o altă repartizare a bilelor în urne se obţine o probabilitate mai mică. Să zicem că în urna U_1 s-ar afla r_1 bile roşii şi n_1 bile negre cu $n_1 \geq 1$. Atunci, în urma experimentului, probabilitatea de a fi extrasă o bilă roşie devine $p = \frac{1}{2}\left(\frac{r_1}{r_1+n_1} + \frac{r_2}{r_2+n_2}\right)$, unde am notat $n_2 = n - n_1$ şi $r_2 = r - r_1$ numerele bilelor ce se află în urna U_2. Trebuie arătat *că* $p_0 \geq p$. *Pentru a demonstra această inegalitate este util a se ţine evidenţa raportului dintre bilele negre şi cele roşii. Mai întâi se observă că inegalitatea* $\frac{r_1}{n_1} \geq \frac{r}{n}$ *este echivalentă cu* $\frac{r_2}{n_2} \leq \frac{r}{n}$. *Apoi se face raţionamentul în cazul în care este adevărată inegalitatea* $\frac{r_1}{n_1} \geq \frac{r}{n}$. *Atunci avem* $\frac{r_1}{r_1+n_1} \geq \frac{r}{r+n} \geq \frac{r_2}{r_2+n_2}$. *Calculăm diferenţa*

$$p_0 - p = \frac{1}{2}\left(1 - \frac{r_1}{r_1+n_1} - \frac{r_{21}}{r_2+n_2} + \frac{r-1}{r-1+n}\right) \geq$$

$$\geq \frac{1}{2}\left(\frac{n_1}{r_1+n_1} - \frac{r}{r+n} + \frac{r-1}{r-1+n}\right) > \frac{1}{2}\left(\frac{n_1}{r+n} - \frac{n}{(r+n)(r-1+n)}\right) \geq$$

$$\geq \frac{1}{2}\left(\frac{n_1}{r+n} - \frac{1}{(r+n)}\right) \geq 0.$$

Se obţine deci inegalitatea strictă $p_0 > p$. *)*

Exerciţiul 2.17. *Avem trei persoane* $1, 2, 3$. *Notăm cu* B_{ij} *evenimentul: persoanele* i *şi* j *au aceeaşi zi de naştere. Sunt independente evenimentele* B_{12} *şi* B_{23} *? Dar evenimentele* B_{12}, B_{23}, B_{13} *? Sunt ele independente câte două?*

Exerciţiul 2.18. *Presupunem că sunt aruncate trei monede. Fie* A *evenimentul „toate monezile cad cu aceiaşi faţă deasupra" şi fie* B *evenimentul „cel puţin o monedă a căzut cu stema deasupra". Sunt aceste evenimente independente?*

Exerciţiul 2.19. *O vitrină este luminată pe timpul nopţii cu două becuri* b_1 *şi* b_2. *Presupunem că probabilitatea ca becul* b_1 *să funcţioneze toată noaptea este de* $0,95$, *iar probabilitatea ca becul* b_2 *să funcţioneze toată noaptea este* $0,98$. *Să se determine probabilitatea ca cel puţin unul din becuri să funcţioneze toată noaptea în fiecare din cazurile: (i) becurile sunt montate în serie, (ii) becurile sunt montate în paralel.*

Exerciţiul 2.20. *Un sportiv trage patru focuri la ţintă. Se ştie că probabilitatea de a lovi ţinta cu un foc este de 0,3. a) Care este probabilitatea de a rata toate primele 3 focuri şi de a lovi ţinta cu al patrulea? b) Care este probabilitatea de a lovi ţinta exact de două ori?*

Exerciţiul 2.21. *Trei trăgători trag fiecare câte un foc asupra unei ţinte. Probabilităţile de succes ale fiecăruia sunt:* p_1, p_2, p_3. *După tragere s-a constatat că ţinta a fost atinsă o singură dată. Care este probabilitatea ca primul trăgător să fie cel ce a nimerit ţinta?*

Exerciţiul 2.22. *Într-o urnă se află* m *bile din care* r *sunt roşii. Se extrag la întâmplare fără revenire* n *bile care se înseamnă cu semnul* $+$. *Se pun la loc bilele, după care se extrage o bilă. Numim* A *evenimentul „bila este roşie" şi numim* B *evenimentul „bila are semnul* $+$ *". Să se arate că evenimentele* A *şi* B *sunt independente.*

(Indicaţie: Se poate arăta că $P(B/A) = P(B/A^c)$. *)*

CAPITOLUL 3

Partiţii finite sau numărabile

În toate modelele elementare familia evenimentelor corespunzătoare este asociată unei partiţii. Prezentăm în acest capitol noţiunile de bază despre partiţii şi generarea σ -algebrlor.

1. Proprietatea de σ -aditivitate

Mai întâi vom discuta în jurul proprietăţii de $\sigma-$ aditivitate, definită în capitolul introductiv, care va interveni incidental în această secţiune.

Lema 3.1. *Fie $(E, \mathcal{E})$ un spaţiu măsurabil şi $\mu : \mathcal{E} \to [0,1]$ o aplicaţie care verifică proprietăţile următoare:*
(i) $\mu(E) = 1$,
(ii) $\mu(A \cup B) = \mu(A) + \mu(B)$, pentru orice două mulţimi, $A, B \in \mathcal{E}$, care sunt disjuncte.
Atunci funcţia de mulţime μ este $\sigma-$ aditivă dacă şi numai dacă satisface relaţia

$$\lim_{n\to\infty} \mu(A_n) = 0,$$

pentru orice şir descrescător $(A_n)_{n\in\mathbf{N}}$ de mulţimi din $\mathcal{E}$ pentru care avem $\cap_n A_n = \emptyset$.

Demonstraţie. Observăm că μ are prin ipoteză proprietăţile unei măsuri de probabilitate pe un spaţiu măsurabil finit. Presupunem că μ este $\sigma-$ aditivă şi că $(A_n)_{n\in\mathbf{N}}$ este un şir descrescător de mulţimi din $\mathcal{E}$ astfel că $\cap_n A_n = \emptyset$. Notăm

$$B_n = A_n \backslash A_{n+1},\ n\in \mathbf{N}.$$

Pentru $m > n$ se verifică relaţia

$$B_n \subset A_{n+1}^c \subset A_m^c,$$

(deoarece $A_m \subset A_n$). Aceasta arată că $B_n \cap B_m = \emptyset$. Deci mulţimile şirului $(B_n)_{n\in\mathbf{N}}$ sunt disjuncte între ele. De asemenea se verifică direct că are loc relaţia

$$\bigcup_{n\in\mathbf{N}} B_n = A_0.$$

Proprietatea de $\sigma-$ aditivitate ne permite să scriem

$$\mu(A_0) = \sum_{n\in\mathbf{N}} \mu(B_n).$$

Dar avem $\mu(B_n) = \mu(A_n) - \mu(A_{n+1})$ şi atunci deducem că sumele parţiale ale seriei pot fi exprimate astfel

$$\sum_{k=0}^{n} \mu(B_k) = \mu(A_0) - \mu(A_{n+1}).$$

Cum aceste sume au limita $\mu(A_0)$, rezultă că $\mu(A_{n+1}) \to 0$.

Să arătăm implicaţia inversă. Presupunem că μ satisface proprietatea din enunţ şi vom arăta că este $\sigma-$ aditivă. Vom alege notaţia care pune în evidenţă faptul că raţionamentul de la prima implicaţie este pur şi simplu repetat în ordine inversă. Fie deci $(B_n)_{n\in\mathbf{N}}$ un şir de mulţimi disjuncte din $\mathcal{E}$. Notăm $A_n = \bigcup_{k\geq n} B_k$ şi vom avea $A_0 = \bigcup_{n\in\mathbf{N}} B_n$. Se verifică din construcţie că şirul $(A_n)_{n\in\mathbf{N}}$ este descrescător şi că $\bigcap_{n\in\mathbf{N}} A_n = \emptyset$. Proprietatea din enunţ implică

$$\lim_{n\to\infty} \mu(A_n) = 0.$$

Pe de altă parte avem

$$\begin{aligned}
\mu(A_n) &= \mu\left(A_0 \setminus \bigcup_{k\leq n-1} B_k\right) = \mu(A_0) - \mu\left(\bigcup_{k\leq n-1} B_k\right) = \\
&= \mu(A_0) - \sum_{k\leq n-1} \mu(B_k).
\end{aligned}$$

Trecând la limită după n se obţine $0 = \mu(A_0) - \sum_{k=0}^{\infty} \mu(B_k)$, ceea ce este acelaşi lucru cu relaţia dorită: $\sum_{k=0}^{\infty} \mu(B_k) = \mu\left(\bigcup_{n\in\mathbf{N}} B_n\right)$. $\square$

Enunţăm acum trei proprietăţi ale măsurilor de probabilitate.

Lema 3.2. *Fie $(\Omega, \mathcal{F}, P)$ un spaţiu probabilizat.*

(i) Pentru orice şir descrescător de evenimente, $(A_n)_{n\in\mathbf{N}} \subset \mathcal{F}$, are loc relaţia

$$\lim_{n\to\infty} P(A_n) = P(\cap_n A_n).$$

(ii) Pentru orice şir crescător de evenimente, $(A_n)_{n\in\mathbf{N}} \subset \mathcal{F}$, are loc relaţia

$$\lim_{n\to\infty} P(A_n) = P(\cup_n A_n).$$

(iii) Pentru un şir arbitrar de evenimente, $(A_n)_{n\in\mathbf{N}} \subset \mathcal{F}$, are loc relaţia

$$P(\cup_n A_n) \leq \sum_{n\in\mathbf{N}} P(A_n).$$

Demonstraţie. (i) Pentru a demonstra prima relaţie din enunţ, să notăm $A = \cap_n A_n$ şi $C_n = A_n \setminus A$. Şirul $(C_n)_{n\in\mathbf{N}}$ este descrescător, cu intersecţia vidă, după cum se poate uşor verifica. Conform lemei de mai înainte, rezultă

$$\lim_{n\to\infty} P(C_n) = 0.$$

Cum $P(A_n) = P(A) + P(C_n)$, deducem relaţia din enunţ.

(ii) A doua relaţie din enunţ rezultă din prima prin trecere la complementară.

(iii) Pentru a treia relaţie pornim de la inegalitatea valabilă pentru un număr finit de mulţimi (vezi punctul 4. din secţiunea introductivă despre modelul probabilist):

$$P\left(\cup_{k\leq n}A_k\right)\leq\sum_{k\leq n}P\left(A_k\right)\leq\sum_{n\in\mathbf{N}}P\left(A_n\right).$$

Apoi aplicăm punctul (i) şirului crescător $\cup_{k\leq n}A_k, n\in\mathbf{N}$, pentru a deduce

$$P\left(\cup_n A_n\right)=\lim_{n\to\infty}P\left(\cup_{k\leq n}A_k\right)\leq\sum_{n\in\mathbf{N}}P\left(A_n\right),$$

adică relaţia de demonstrat. □

2. Generarea algebrelor şi $\sigma-$algebrelor

Vom presupune că Ω este o mulţime arbitrară. Fiind dată o familie arbitrară de părţi, $\mathcal{U}\subset\mathcal{P}\left(\Omega\right)$ vom nota cu $ma\left(\mathcal{U}\right)$ mulţimea tuturor algebrelor de părţi ale lui Ω care conţin pe $\mathcal{U}$. Această mulţime nu este vidă deoarece, în mod evident, $\mathcal{P}\left(\Omega\right)\in ma\left(\mathcal{U}\right).$ Vom nota cu $a\left(\mathcal{U}\right)$ intersecţia tuturor algebrelor din $ma\left(\mathcal{U}\right):$

$$a\left(\mathcal{U}\right)=\bigcap_{\mathcal{G}\in ma(\mathcal{U})}\mathcal{G}.$$

Se verifică fără dificultate că $a\left(\mathcal{U}\right)$ este o algebră, şi că $\mathcal{U}\subset a\left(\mathcal{U}\right)\subset\mathcal{F}$. De asemenea, este evident că orice element $\mathcal{G}$ din $ma\left(\mathcal{U}\right)$ conţine $a\left(\mathcal{U}\right)$ şi de aceea se poate zice că $a\left(\mathcal{U}\right)$ este cea mai mică algebră ce conţine $\mathcal{U}$. Vom spune că $a\left(\mathcal{U}\right)$ este *algebra generată* de familia $\mathcal{U}$.

Se poate de asemenea verifica uşor că în cazul a două familii de părţi astfel încât $\mathcal{U}\subset\mathcal{U}'$, rezultă $ma\left(\mathcal{U}'\right)\subset ma\left(\mathcal{U}\right)$ şi $a\left(\mathcal{U}\right)\subset a\left(\mathcal{U}'\right).$

În mod absolut similar notăm cu $msa\left(\mathcal{U}\right)$ mulţimea tuturor $\sigma-$algebrelor de părţi din Ω care conţin pe $\mathcal{U}$. Vom nota cu $\sigma\left(\mathcal{U}\right)$ intersecţia tuturor $\sigma-$algebrelor din $msa\left(\mathcal{U}\right):$

$$\sigma\left(\mathcal{U}\right)=\bigcap_{\mathcal{E}\in msa(\mathcal{U})}\mathcal{E}.$$

Aceasta este cea mai mică $\sigma-$algebră care conţine pe $\mathcal{U}$ şi se numeşte *$\sigma-$algebra generată* de $\mathcal{U}$. Pentru două familii de evenimente astfel încât $\mathcal{U}\subset\mathcal{U}'$, vom avea $msa\left(\mathcal{U}'\right)\subset msa\left(\mathcal{U}\right)$ şi în consecinţă $\sigma\left(\mathcal{U}\right)\subset\sigma\left(\mathcal{U}'\right).$

Este uşor de văzut că dacă $\mathcal{U}$ este o algebră, atunci $a\left(\mathcal{U}\right)=\mathcal{U}$, iar dacă $\mathcal{U}$ este o σ-algebră atunci $\sigma\left(\mathcal{U}\right)=\mathcal{U}$. În general, pentru orice familie $\mathcal{U}$ are loc egalitatea $\sigma\left(\mathcal{U}\right)=\sigma\left(a\left(\mathcal{U}\right)\right).$

Dacă $\Lambda\subset\Omega$ este o parte arbitrară şi $\mathcal{U}\subset\mathcal{P}\left(\Omega\right)$ vom utiliza notaţia

$$\mathcal{U}_{|\Lambda}=\left\{A\subset\Lambda/\ \exists\ A'\in\mathcal{U}, A=A'\cap\Lambda\right\},$$

pentru a desemna urma familiei $\mathcal{U}$ pe Λ. Atunci, după cum afirmă următorul enunţ, urma σ-algebrei generate de $\mathcal{U}$ şi σ-algebra generată de urma lui $\mathcal{U}$ coincid.

Lema 3.3. $\sigma\left(\mathcal{U}_{|\Lambda}\right) = \sigma(\mathcal{U})_{|\Lambda}$

DEMONSTRAŢIE. Mai întâi afirmăm că $\sigma(\mathcal{U})_{|\Lambda}$ este o σ-algebră. Aceasta se verifică direct, fără probleme. Pe de altă parte este clar că are loc incluziunea $\mathcal{U}_{|\Lambda} \subset \sigma(\mathcal{U})_{|\Lambda}$. Atunci rezultă incluziunea $\sigma\left(\mathcal{U}_{|\Lambda}\right) \subset \sigma(\mathcal{U})_{|\Lambda}$. Pentru a proba incluziunea inversă notăm

$$\mathcal{M} = \left\{A \in \sigma(\mathcal{U}) \,/\, A \cap \Lambda \in \sigma\left(\mathcal{U}_{|G}\right)\right\}.$$

Se verifică direct că această clasă de mulţimi este o σ-algebră. Cum, evident, are loc incluziunea $\mathcal{U} \subset \mathcal{M}$, rezultă $\sigma(\mathcal{U}) \subset \mathcal{M}$. Această incluziune spune exact că $\sigma(\mathcal{U})_{|\Lambda} \subset \sigma\left(\mathcal{U}_{|\Lambda}\right)$, ceea ce încheie demonstraţia. $\square$

3. Mulţimile măsurabile Borel*

Clasa *mulţimilor boreliene* din spaţiul euclidian $\mathbf{R}^n$ este notată $\mathcal{B}(\mathbf{R}^n)$ şi este definită prin $\mathcal{B}(\mathbf{R}^n) = \sigma(\mathcal{D})$, unde $\mathcal{D}$ reprezintă familia mulţimilor deschise din $\mathbf{R}^n$. Poartă numele matematicianului francez Emile Borel (1871 -1956). Mulţimile boreliene sunt obţinute şi în felul următor.

Lema 3.4. *Să notăm prin $\mathcal{K}$ familia mulţimilor compacte şi prin $\mathcal{K}'$ familia mulţimilor închise din $\mathbf{R}^n$. Are loc egalitatea $\sigma(\mathcal{K}) = \sigma(\mathcal{K}') = \mathcal{B}(\mathbf{R}^n)$.*

DEMONSTRAŢIE. Orice mulţime compactă, $E \in \mathcal{K}$, este închisă şi de aceea avem $\mathcal{K} \subset \mathcal{K}'$. Rezultă $\sigma(\mathcal{K}) \subset \sigma(\mathcal{K}')$. Pe de altă parte, orice mulţime închisă $F \in \mathcal{K}$ se scrie sub forma $F = \bigcup_{n\in\mathbf{N}} F_n$, unde $F_n = F \cap \{|x| \leq n\}, n \in \mathbf{N}$. Mulţimile $F_n, n \in \mathbf{N}$, sunt închise şi mărginite, deci compacte. Prin urmare $\mathcal{K} \subset \sigma(\mathcal{K}')$, ceea ce implică $\sigma(\mathcal{K}) \subset \sigma(\mathcal{K}')$. Am verificat astfel egalitatea $\sigma(\mathcal{K}) = \sigma(\mathcal{K}')$.

Cum orice mulţime este deschisă este complementara unei mulţimi închise, rezultă imedia egalitatea $\sigma(\mathcal{D}) = \sigma(\mathcal{K}')$. $\square$

Dacă $G \subset \mathbf{R}^n$ este o mulţime arbitrară, atunci, după cum se ştie, ea este înzestrată cu topologia urmă $\mathcal{D}_{|G}$ în care mulţimile deschise sunt urmele pe G ale deschişilor din $\mathbf{R}^n$. Ţinând cont de lema anterioară rezultă că $\mathcal{B}(\mathbf{R}^n)_{|G} = \sigma\left(\mathcal{D}_{|G}\right)$. De obicei se utilizează notaţia $\mathcal{B}(G) = \mathcal{B}(\mathbf{R}^n)_{|G}$ şi se spune că aceasta este clasa mulţimilor boreliene din G.

În cazul dreptei reale, $\mathbf{R}$, există mai multe familii de intervale care generează $\sigma-$algebra mulţimilor boreliene. Acestea sunt descrise prin

$$\mathcal{I} = \{(a,b)\,/a < b\}, \; \mathcal{I}' = \{[a,b]\,/a \leq b\},$$

$$\mathcal{I}'' = \{[a,b)/a < b\}, \; \mathcal{I}''' = \{(a,b]/a < b\},$$

$$\mathcal{J} = \{(a,\infty)/a \in \mathbf{R}\}, \ \mathcal{J}' = \{[a,\infty)/a \in \mathbf{R}\},$$
$$\mathcal{J}'' = \{(\infty,a)/a \in \mathbf{R}\}, \ \mathcal{J}''' = \{(\infty,a]/a \in \mathbf{R}\}.$$

Se poate verifica că au loc relaţiile

$$a(\mathcal{I}) = a(\mathcal{I}') = a(\mathcal{I}'') = a(\mathcal{I}''') = a(\mathcal{J}) = a(\mathcal{J}') = a(\mathcal{J}'') = a(\mathcal{J}''').$$

$$\begin{aligned}\mathcal{B}(\mathbf{R}^n) &= \sigma(\mathcal{I}) = \sigma(\mathcal{I}') = \sigma(\mathcal{I}'') = \sigma(\mathcal{I}''') = \\ &= \sigma(\mathcal{J}) = \sigma(\mathcal{J}') = \sigma(\mathcal{J}'') = \sigma(\mathcal{J}''').\end{aligned}$$

Rămâne ca exerciţiu pentru cititor verificarea acestor egalităţi.

Se poate spune că printre mulţimile boreliene se află toate mulţimile ce apar prin operaţiile uzuale din analiză. Mulţimile care nu sunt boreliene sunt de fapt patologice, iar construcţia lor este de obicei dificilă.

4. Partiţii

Pentru binecunoscuta noţiune de partiţie vom pune următoarea definiţie, care stabileşte câteva aspecte mai precise ca de obicei. Mai întâi vom preciza termenul de *mulţime cel mult numărabilă*: prin aceasta înţelegem două posibilităţi, sau mulţimea este finită şi prin urmare este izomorfă cu o mulţime de forma $\{1,...,n\}$ cu un anumit număr $n \in \mathbf{N}^*$, sau mulţimea este numărabilă şi atunci poate fi pusă în bijecţie cu mulţimea numerelor naturale. Dacă I este o mulţime numărabilă şi $f : \mathbf{N} \to I$ este o bijecţie spunem că I este numerotată prin această bijecţie şi se obişnuieşte notaţia $I = \{i_0, i_1, i_2, ...\}$, unde $i_k = f(k), k \in \mathbf{N}$.

Definiţia 3.1. *O familie de părţi nevide $\mathcal{A} = (A_i)_{i\in I} \subset \mathcal{P}(\Omega)$, indexată după o mulţime cel mult numărabilă I, este numită* partiţie cel mult numărabilă *a lui Ω, dacă sunt satisfăcute condiţiile următoare:*
(i) dacă $i \neq j$, atunci $A_i \cap Aj = \emptyset$,
(ii) $\bigcup_{i\in I} A_i = \Omega$.

Facem convenţia ca în acest capitol să utilizăm cuvântul *partiţie* pentru a prescurta termenul de partiţie cel mult numărabilă. O astfel de partiţie mai este numită şi *desfacere*. Un eveniment care aparţine unei partiţii este numit *atom* al partiţiei. Un exemplu tipic de partiţie finită a apărut în paragraful dedicat exemplului controlului calităţii.

Fie $\mathcal{A} = \{A_i \mid i \in I\}$ o partiţie cel mult numărabilă. Pentru o parte $J \subset I$, vom utiliza notaţia $A_J = \bigcup_{i\in J} A_i$, cu convenţia $A_\emptyset = \emptyset$. Deoarece atomii sunt părţi nevide, rezultă că o mulţime A_J este nevidă dacă $J \neq \emptyset$. Mai notăm cu $\mathcal{K}(I)$ familia submulţimilor finite din I. Următoarea lemă descrie algebra şi $\sigma-$algebra asociate partiţiei.

Lema 3.5. *(i) Algebra generată de $\mathcal{A}$ are următoarea descriere*

$$a(\mathcal{A}) = \{A_J/J \in \mathcal{K}(I), \ sau \ J^c \in \mathcal{K}(I)\}.$$

Dacă partiţia $\mathcal{A}$ este finită atunci $a(\mathcal{A})$ este tot finită. În particular în acest caz avem $a(\mathcal{A}) = \sigma(\mathcal{A})$.

(ii) $\sigma-$algebra generată de $\mathcal{A}$ are următoarea descriere

$$\sigma(\mathcal{A}) = \{A_J \mid J \subset I\}.$$

(iii) Dacă $\mathcal{A}'$ este o a doua partiţie şi $\sigma(\mathcal{A}) = \sigma(\mathcal{A}')$, atunci $\mathcal{A}$ şi $\mathcal{A}'$ coincid ca mulţimi de părţi, modulo felul în care sunt indexate.

(iv) O măsură de probabilitate P pe $\sigma(\mathcal{A})$ este complet determinată de numerele $P(A_i), i \in I$, prin formula

$$P(A_J) = \sum_{i \in J} P(A_i).$$

Demonstraţie. Mai întâi să observăm că au loc următoarele egalităţi:

$$A_{J \cup J'} = A_J \cup A_{J'}, \quad A_{J \cap J'} = A_J \cap A_{J'}, \quad (*)$$

pentru orice părţi $J, J' \subset I$. În particular rezultă

$$A_J \cup A_{J^c} = A_{J \cup J^c} = \Omega, \quad A_J \cap A_{J^c} = A_\emptyset = \emptyset,$$

ceea ce implică

$$(A_J)^c = A_{J^c}, \quad A_J = (A_{J^c})^c \qquad (**)$$

(i) Vom nota cu $\mathcal{G}$ familia de părţi reprezentată de membrul drept al egalităţii ce descrie $a(\mathcal{A})$. Dacă $J \in \mathcal{K}(I)$, rezultă că mulţimea A_J reprezintă o reuniune finită de atomi, deci este în $a(\mathcal{A})$. Fie acum o mulţime de indici J astfel încât $J^c \in \mathcal{K}(I)$. Atunci avem $A_{J^c} \in a(\mathcal{A})$ şi din a doua relaţie din (**) rezultă că şi $A \in a(\mathcal{A})$. Deci am verificat incluziunea $\mathcal{G} \subset a(\mathcal{A})$.

Egalităţile (*) şi (**) de mai sus, permit a se proba proprietăţile de algebră pentru $\mathcal{G}$. În concluzie, rezultă că $\mathcal{G} = a(\mathcal{A})$.

Dacă mulţimea de indici I este finită, atunci mulţimea părţilor $\mathcal{P}(I)$ este tot finită şi prin urmare $\mathcal{G}$, la rândul ei, este tot finită.

(ii) Faptul că $\sigma(\mathcal{A})$ are descrierea afirmată în enunţ se verifică cu argumente similare celor de la punctul precedent.

(iii) Fie $\mathcal{A}' = (A'_i)_{i \in I'}$ o a doua partiţie care generează aceeaşi $\sigma-$algebră. O mulţime din prima partiţie, A_i, fiind în algebra $\sigma(\mathcal{A}')$, se scrie, conform punctului (i), ca o reuniune de atomi ai celei de a doua partiţii,

$$A_i = \bigcup_{j \in J} A'_j,$$

unde $J \subset I'$. Deoarece A_i este nevidă rezultă că J trebuie să fie nevidă. Să analizăm o mulţime care participă la această reuniune, fie ea A'_j cu $j \in J$. Ţinând cont că A'_j este în $\sigma(\mathcal{A})$, această mulţime se va exprima ca o reuniune de atomi din partiţia $\mathcal{A}$,

$$A'_j = \bigcup_{l \in L} A_l,$$

unde $L \subset I$. La fel ca şi A'_j, fiecare din mulţimile care participă la această reuniune este inclusă în A_i. Dar atomii distincţi ai partiţiei $\mathcal{A}$ trebuie să fie disjuncţi. Rezultă că mulţimea L se reduce la un singur punct $L = \{i\}$. În particular, am dedus că $A'_j = A_i$, ceea ce înseamnă că $A_i \in \mathcal{A}'$. Dar mulţimea A_i a fost o mulţime arbitrară din $\mathcal{A}$, şi atunci putem trage concluzia că $\mathcal{A} \subset \mathcal{A}'$. În mod simetric avem şi incluziunea inversă, ceea ce probează egalitatea $\mathcal{A} = \mathcal{A}'$.

(iv) Pentru a stabili formula din enunţ pornim mai întâi cu mulţimea de indici finită $J \subset I$, care defineşte mulţimea $A_J = \bigcup_{i \in J} A_i \in \sigma(\mathcal{A})$. În acest caz formula rezultă imediat din proprietatea de finit aditivitate a măsurii.

Să presupunem acum că mulţimea de indici J este infinită. Trebuie să menţionăm că atunci suma care apare în formula din enunţ trebuie înţeleasă în sensul dat în paragraful despre sumabilitate din capitolul următor. Cum J este numărabilă alegem o numerotare $J = \{i_0, i_1, ...\}$şi exprimăm $A_J = \bigcup_{n \in \mathbf{N}} A_{i_n}$. Atunci relaţia de σ-aditivitate ne dă

$$P(A_J) = \sum_{n \in \mathbf{N}} P(A_{i_n}) = \sum_{i \in J} P(A_i).$$

La ultimul semn egal am utilizat punctul doi din *lema* 4.1. Cu aceasta am încheiat demonstraţia. □

Observaţia 3.1. *Din descrierea făcută algebrei generate de o partiţie reiese că aceasta este şi ea cel mult numărabilă. Din contră, σ -algebra generată de o partiţie numărabilă este de acelaşi cardinal cu mulţimea părţilor lui* $\mathbf{N}$, *adică de puterea continuului.*

Observaţia 3.2. *Din lema precedentă rezultă că două partiţii $\mathcal{A}, \mathcal{A}'$, sunt legate prin relaţia $\mathcal{A}' \subset a(\mathcal{A})$ dacă şi numai dacă $\mathcal{A}$ este mai fină decât $\mathcal{A}'$, în sensul că orice element $B \in \mathcal{A}'$ se reprezintă ca o reuniune de elemente din $\mathcal{A}$:*

$$B = \bigcup_{A \in \mathcal{A}, A \subset B} A.$$

Lema anterioară mai arată că în cazul în care o $\sigma-$algebră este generată de o partiţie, partiţia respectivă este unic determinată. În cazul partiţiilor finite, se poate completa relaţia dintre partiţia finită şi algebra generată. Anume, vom arăta în continuare că orice algebră finită este generată de o partiţie finită, ceea ce pune în corespondenţă bijectivă cele două clase de obiecte. Mai întâi stabilim că orice familie finită dă naştere la o partiţie care generează aceeaşi algebră ca şi familia dată.

Lema 3.6. *Fie $\mathcal{U} = \{A_1, ..., A_n\}$ o familie finită arbitrară de părţi. Notăm $A_{l,1} = A_l$, $A_{l,-1} = A_l^c$ şi apoi*

$$A^\lambda = A_{1,\lambda_1} \cap ... \cap A_{n,\lambda_n},$$

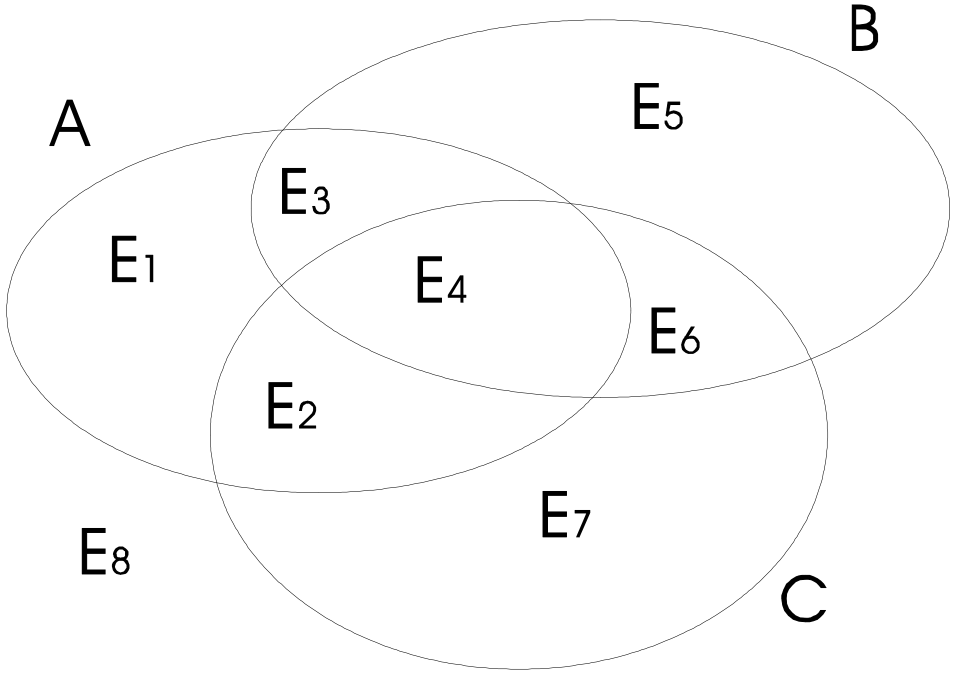


FIGURA 1. Diagrama Venn a mulţimilor A,B,C şi partiţia generată de ele

unde $\lambda = (\lambda_1, ..., \lambda_n) \in \{-1, 1\}^n$. *Mai notăm*

$$\Lambda = \left\{\lambda \in \{-1, 1\}^n \; / \; A^\lambda \neq \emptyset\right\}.$$

Atunci $\mathcal{A} = \left\{A^\lambda / \; \lambda \in \Lambda\right\}$ *este o partiţie (finită) şi* $a(\mathcal{U}) = a(\mathcal{A})$.

DEMONSTRAŢIE. Să verificăm că elementele lui $\mathcal{A}$ sunt disjuncte. Fie $\lambda \neq \gamma$ doi multi -indici distincţi din Λ. Atunci există $1 \leq i \leq n$ astfel ca $\lambda_i \neq \gamma_i$, ceea ce implică $A_{i,\lambda_i} \cap A_{i,\gamma_i} = \emptyset$. Rezultă $A^\lambda \cap A^\gamma = \emptyset$.

Să verificăm că $\mathcal{A}$ acoperă Ω. Fie $x \in \Omega$. Pentru fiecare $1 \leq l \leq n$ este adevărată una din relaţiile $x \in A_l$ sau $x \in A_l^c$; în prima situaţie definim $\lambda_l = 1$, iar în a doua $\lambda_l = -1$. Obţinem astfel multi -indicele $\lambda = (\lambda_1, ..., \lambda_n)$ pentru care avem $x \in A^\lambda$. Am verificat deci că $\mathcal{A}$ este o partiţie.

Este clar din definiţia lui A^λ că acest eveniment aparţine lui $a(\mathcal{U})$. Atunci avem $\mathcal{A} \subset a(\mathcal{U})$ şi, prin urmare, $a(\mathcal{A}) \subset a(\mathcal{U})$. Pe de altă parte, se poate verifica următoarea relaţie

$$A_i = \bigcup_{\lambda \in \Lambda, \lambda_i = 1} A^\lambda,$$

pentru fiecare $1 \leq i \leq n$, relaţie care arată că $\mathcal{U} \subset a(\mathcal{A})$ şi prin urmare $a(\mathcal{U}) \subset a(\mathcal{A})$. □

Exemplu.

Diagrama Venn din figura 1 reprezintă mulţimile A, B, C în formă de elipsă.

Presupunem că aceste mulţimi sunt părţi ale mulţimii mai mari Ω care este reprezentată de planul în care se află desenul. Partiţia asociată algebrei generate de evenimentele A, B, C este compusă din mulţimile $E_1, ..., E_8$. Mulţimea E_8 este complementara reuniunii $A \cup B \cup C$.

Propoziţia 3.1. *Fiind dată o algebră finită există o partiţie finită care o generează.*

DEMONSTRAŢIE. Fie $\mathcal{G}$ algebra finită dată. Aplicând lema anterioară relativ la $\mathcal{G}$ pe post de $\mathcal{U}$, rezultă o partiţie finită, $\mathcal{A}$, astfel încât $\mathcal{G} = a(\mathcal{G}) = a(\mathcal{A})$. $\square$

Cum orice algebră finită este chiar $\sigma-$algebră, unicitatea dată de punctul (iii) al *lemei* 3.5 arată că partiţia dată de această propoziţie este unică. Ea va fi numită *partiţia asociată* algebrei $\mathcal{G}$. Cu această denumire putem enunţa următorul corolar al lemelor anterioare.

Corolarul 3.1. *Există o bijecţie între algebrele finite pe Ω şi partiţiile finite, iar aplicaţiile „algebra generată de o partiţie" şi „partiţia asociată la o algebră" sunt inverse una alteia.*

$\sigma-$algebrele care sunt generate de o partiţie cel mult numărabilă sunt de un tip special. Am putea să numim atomice aceste $\sigma-$algebre. În cazul în care mulţimea Ω este cel mult numărabilă, evident, orice $\sigma-$algebră este generată de o partiţie cel mult numărabilă. Descrierea în termeni de atomi pe care o dă *lema* 3.5 (ii) arată că ele sunt în mod esenţial mai simple decât $\sigma-$algebra mulţimilor boreliene din $\mathbf{R}^n$.

Observăm că reuniunea unui număr finit de algebre finite este o familie finită şi, prin urmare, algebra generată de această reuniune este tot finită. În secţiunea următoare ne vom interesa de reuniunile finite de σ -algebre generate de partiţii cel mult numărabile şi.vom arăta că o astfel de reuniune generează o $\sigma-$algebră care este de acelaşi tip, adică şi ea este generată de o partiţie.

În general nu ne putem aştepta ca o $\sigma-$algebră arbitrară să poată fi generată de o partiţie. Se poate chiar să avem o $\sigma-$algebră generată de o algebră de părţi numărabilă, dar să nu existe o partiţie numărabilă asociată acelei σ-algebre. Aceasta o probează următorul exemplu.

Exemplu*.

Fie $\Omega = [0, 1)$. Notăm cu $\mathcal{G}$ clasa mulţimilor de forma $\bigcup_{l=1}^{n} [a_l, b_l)$, unde a_l, b_l sunt numere raţionale, $n \in \mathbf{N}^*$ şi

$$0 \leq a_1 < b_1 < ... < a_n < b_n \leq 1,$$

la care vom adăuga şi mulţimea vidă. Atunci se verifică direct că această clasă de mulţimi este o algebră de părţi numărabilă. Vom arăta că $\sigma(\mathcal{G})$ nu poate fi generată de o partiţie.

Începem prin a remarca că pentru fiecare punct $x \in [0, 1)$ există două şiruri de numere raţionale $(a_n)_{n \in \mathbf{N}}, (b_n)_{n \in \mathbf{N}}$ astfel că primul este crescător iar al doilea este strict descrescător şi ambele au limita x.

Rezultă că are loc relaţia

$$\{x\} = \bigcap_{n\in\mathbf{N}} [a_n, b_n).$$

Aceasta arată că, dacă am presupune că există o partiţie care ar genera aceeaşi σ-algebră ca şi $\mathcal{G}$, atunci fiecare mulţime de tipul $\{x\}, x \in [0,1)$, ar trebui să fie un atom, conform cu ceea ce spune lema de mai jos. Dar atunci atomii ar forma o mulţime nenumărabilă. Cum noi discutăm de partiţii cel mult numărabile, lucrul acesta nu este posibil şi, prin urmare, presupunerea făcută cade.

Pe de altă parte, se poate uşor verifica că $\sigma(\mathcal{G}) = \mathcal{B}([0,1))$. $\square$

Lema 3.7. *Fie $\mathcal{A}$ o partiţie cel mult numărabilă şi $A \in \mathcal{A}$ un atom al ei, iar $x \in A$. Atunci orice mulţime $B \in \sigma(\mathcal{A})$ astfel că $x \in B$ are proprietatea că $A \subset B$.*

DEMONSTRAŢIE. Fie $\mathcal{A} = (A_i)_{i\in I}$ partiţia şi A_{i_0} atomul care conţine punctul x. Mulţimea B va fi de forma $B = A_J$ cu $J \subset I$, datorită descrierii date de punctul (ii) din *lema* 3.5 pentru $\sigma(\mathcal{A})$. Cum $x \in B$, rezultă că va exista un indice $j \in J$ astfel ca $x \in A_j$. Privim acum atomii A_{i_0} şi A_j. Aceştia au în comun punctul x. Rezultă că ei trebuie să coincidă, pentru că, altfel, atomii distincţi sunt disjuncţi. Dar atunci este clar că $A_{i_0} = A_j \subset B$. $\square$

5. $\sigma-$algebra generată de o aplicaţie

Fie $(\Omega, \mathcal{F})$ un spaţiu măsurabil şi E o mulţime cel mult numărabilă. O aplicaţie $X : \Omega \to E$ este numită *aplicaţie măsurabilă* dacă are proprietatea că $X^{-1}(B) \in \mathcal{F}$, pentru orice $B \subset E$. Remarcăm că, pentru ca X să fie măsurabilă este suficient să se verifice că, pentru orice punct $e \in E$, este satisfăcută condiţia $X^{-1}(e) \in \mathcal{F}$.

Să presupunem acum că Ω este o mulţime arbitrară şi $X : \Omega \to E$ este o aplicaţie de la Ω cu valori într-o mulţime cel mult numărabilă E. Acestei aplicaţii îi putem asocia clasa de mulţimi

$$\sigma(X) = \left\{X^{-1}(B) \,/\, B \subset E\right\},$$

care este formată din toate preimaginile întoarse de aplicaţia X. Listăm mai jos câteva din faptele mai importante legate de această clasă de mulţimi.

1. $\sigma(X)$ este o σ -algebră şi anume este cea mai mică σ -algebră care face măsurabilă aplicaţia X. După cum se poate uşor verifica, ea coincide cu $\sigma-$algebra generată de partiţia $\mathcal{A} = \left(X^{-1}(e)\right)_{e\in E'}$, unde $E' = X(\Omega)$ este mulţimea valorilor lui X.

$\sigma(X)$ poartă denumirea de σ -algebra *generată* de X.

2. Fie $E_1, ..., E_n$, mulţimi cel mult numărabile şi să presupunem că avem aplicaţiile $X_i : \Omega \to E_i, i = 1, ..., n$. Notăm prin $\sigma(X_1, ..., X_n)$ cea mai mică σ -algebră în raport cu care sunt măsurabile toate aceste aplicaţiiadică $\sigma(X_1, ..., X_n) = \sigma\left(X^{-1}(B) / B \subset E_i, i = 1, ..., n\right)$. Notăm $\mathcal{A}_i =$

$\left(X_i^{-1}(e)\right)_{e\in E_i'}$ partiţiile corespunzătoare acestor aplicaţii, unde $E_i' = X_i(\Omega), i = 1, ..., n$. Atunci se poate uşor demonstra egalitatea

$$\sigma(X_1, ..., X_n) = \sigma\left(\bigcup_{i=1}^{n} \mathcal{A}_i\right).$$

3. Cu notaţia de la punctul anterior, definim $Z : \Omega \to E = E_1 \times ... \times E_n$, aplicaţia exprimată prin $Z(\omega) = (X_1(\omega), ..., X_n(\omega))$. Atunci are loc egalitatea $\sigma(X_1, ..., X_n) = \sigma(Z)$.

Pentru verificarea acestei afirmaţii observăm mai întâi că pentru orice punct din spaţiul produs $e = (e_1, ..., e_n) \in E$ are loc egalitatea

$$Z^{-1}(e) = X_1^{-1}(e_1) \cap ... \cap X_n^{-1}(e_n),$$

care arată că preimaginea $Z^{-1}(e)$ este în $\sigma(X_1, ..., X_n)$. Pentru o mulţime arbitrară $B \subset E$ putem scrie

$$Z^{-1}(B) = \bigcup_{e\in B} Z^{-1}(e),$$

cu o reuniune numărabilă în dreapta, ceea ce arată că şi $Z^{-1}(B) \in \sigma(X_1, ..., X_n)$. Deci avem $\sigma(Z) \subset \sigma(X_1, ..., X_n)$. Pe de altă parte, pentru un indice $1 \leq i \leq n$ fixat şi orice mulţime $A \subset E_i$ putem scrie

$$X_i^{-1}(A) = Z^{-1}(E_1 \times ... \times E_{i-1} \times A \times E_{i+1} \times ... \times E_n),$$

relaţie ce arată că $\sigma(X_i) \subset \sigma(Z)$. Rezultă $\sigma(X_1, ..., X_n) \subset \sigma(Z)$, ceea ce încheie verificarea afirmaţiei făcute.

4. De asemenea, păstrând notaţia anterioară, dacă $f : E_1 \times ... \times E_n \to F$ este o aplicaţie cu valori în mulţimea cel mult numărabilă F, atunci $Y = f(X_1, ..., X_n)$ este măsurabilă faţă de $\sigma(X_1, ..., X_n)$.

Într-adevăr, pentru orice mulţime $A \subset F$, are loc relaţia

$$Y^{-1}(A) = Z^{-1}(B),$$

unde $B = f^{-1}(A)$, relaţie care pune în evidenţă faptul că $Y^{-1}(A)$ este măsurabilă faţă de $\sigma(Z)$.

Propoziţia 3.2. *Fie* $\mathcal{A}_l = (A_{l,i})_{i\in I_l}$, $l = 1, ..., n$, *partiţii pe mulţimea* Ω. *Atunci există o unică partiţie* $\mathcal{A}$ *astfel încât* $\sigma(\mathcal{A}) = \sigma\left(\bigcup_{l=1}^{n} \mathcal{A}_l\right)$.

Demonstraţie. Considerăm variabilele aleatoare $X_l : \Omega \to I_l$, definite prin $X_l(\omega) = i$, dacă $\omega \in A_{l,i}$. Observăm că are loc egalitatea $\sigma(X_l) = \sigma(\mathcal{A}_l)$. Pe de altă parte, aplicaţia $Z = (X_1, ..., X_n) : \Omega \to I$, cu valori în spaţiul produs $I = I_1 \times ... \times I_n$ ne dă σ -algebra

$$\sigma(Z) = \sigma(X_1, ..., X_n) = \sigma\left(\bigcup_{l=1}^{n} \mathcal{A}_l\right),$$

care este generată de partiţia $\{Z^{-1}(i_1,...,i_n) / (i_1,...,i_n) \in I'\}$, unde am notat cu I' mulţimea imagine $I' = Z(\Omega)$. Unicitatea este asigurată de *lema* 3.5 (iii). $\square$

Observaţia 3.3. *Privitor la propoziţia anterioară, facem observaţia că atomii partiţiei obţinute pot fi exprimaţi în funcţie de cei ai partiţiilor date. Anume, ţinînd cont că* $Z^{-1}(i_1,...,i_n) = X_1^{-1}(i_1) \cap ... \cap X_n^{-1}(i_n)$ *şi de relaţia* $X_l^{-1}(i_l) = A_{l,i_l}$ *rezultă că atomii partiţiei căutate sunt de forma* $A_{1,i_1} \cap ... \cap A_{n,i_n}$, *cu condiţia ca această intersecţie să nu fie vidă.* $\square$

6. Exerciţii

Exerciţiul 3.1. ** Notăm cu* $\mathcal{D}_f$ *muţimea tuturor partiţiilor şi introducem următoarea relaţie de ordine pe această mulţime: dacă* $\mathcal{A}, \mathcal{A}' \in \mathcal{D}_f$ *vom spune că* $\mathcal{A}$ *este mai fină decât* $\mathcal{A}'$ *şi vom scrie* $\mathcal{A} \prec \mathcal{A}'$ *dacă fiecare element* A *din* $\mathcal{A}'$ *poate fi scris ca o reuniune de elemente din* $\mathcal{A}$, *adică dacă are loc relaţia*

$$A = \bigcup_{B \in \mathcal{A}, B \subset A} B.$$

i) Să se arate că $\mathcal{A}$ *este mai fină decât* $\mathcal{A}'$ *dacă şi numai dacă are loc incluziunea* $\mathcal{A}' \subset a(\mathcal{A})$.

ii) Dacă $\mathcal{A} \prec \mathcal{A}'$ *şi* $\mathcal{A}' \prec \mathcal{A}$, *atunci* $\mathcal{A} = \mathcal{A}'$.

iii) Dacă $\mathcal{A} \prec \mathcal{A}'$ *şi* $\mathcal{A}' \prec \mathcal{A}''$, *atunci* $\mathcal{A} \prec \mathcal{A}''$.

iv) Pentru două elemente $\mathcal{A}, \mathcal{A}' \in \mathcal{D}_f$ *există întotdeauna cel mai mare minorant* $\mathcal{A} \wedge \mathcal{A}'$ *în raport cu relaţia de ordine* $\prec$.

v) Fiind date partiţiile $\mathcal{A}_1, ..., \mathcal{A}_n \in \mathcal{D}_f$, *cel mai mare minorant al lor este partiţia* $\mathcal{A}$ *asociată la reuniunea* $\mathcal{A}_1 \cup ... \cup \mathcal{A}_n$, *prin construcţia din* propoziţia *3.2.*

Exerciţiul 3.2. *Presupunem că spaţiul probabilizat* $(\Omega, \mathcal{F}, P)$ *este finit.*

i) Să se arate că există o mulţime finită Ω' *şi o aplicaţie* $f : \Omega \to \Omega'$ *surjectivă astfel încât* $\mathcal{F} = f^{-1}(\mathcal{P}(\Omega'))$.

ii) Perechea (Ω', f) *este unic determinată de aceste proprietăţi, în sensul că dacă* (Ω'', g) *este o altă mulţime finită şi o aplicaţie surjectivă* $g : \Omega \to \Omega''$, *astfel încât* $\mathcal{F} = g^{-1}(\mathcal{P}(\Omega''))$, *atunci există o bijecţie* $h : \Omega' \to \Omega''$ *astfel încât* $g = h \circ f$.

iii) Dacă notăm cu P' *măsura de probabilitate transportată pe mulţimea* Ω' *de la punctul i) prin formula* $P'(A) = P(f^{-1}(A)), A \in \mathcal{P}(\Omega')$, *atunci spaţiul probabilizat* $(\Omega', \mathcal{P}(\Omega'), P')$ *este echivalent cu* $(\Omega, \mathcal{F}, P)$, *ca spaţiu de modelare. Mai precis, arătaţi că aplicaţia* $F : \mathcal{P}(\Omega') \to \mathcal{F}$ *definită prin* $F(A) = f^{-1}(A)$ *este o bijecţie cu proprietăţile următoare: a) dacă* $A \subset B$, *atunci* $F(A) \subset F(B)$; *b)* $P'(A) = P(F(A))$.

Exerciţiul 3.3. ** Daţi exemplu de un şir de partiţii* $\mathcal{A}_1, \mathcal{A}_2, ...$ *astfel încât* $\sigma(\mathcal{A}_1, \mathcal{A}_2, ...)$ *să nu fie generată de o partiţie.*

Exerciţiul 3.4. ** Fie* $(\Omega, \mathcal{F}, P)$ *un spaţiu probabilizat şi* $A_1, ..., A_n$ *evenimente. Arătaţi că aceste evenimente sunt independente dacă şi numai dacă* A_k *este independent de algebra* $a\left(A_1, A_2, ..., A_{k-1}\right)$, *pentru fiecare* $k = 2, 3, ..., n$.

Exerciţiul 3.5. ** Fie* $\mathcal{F}_1, ..., \mathcal{F}_n$ *o familie de* σ *-algebre cu proprietatea că, pentru fiecare* $k = 1, ..., n-1$, $\mathcal{F}_{k+1}$ *este independentă de* $\sigma\left(\mathcal{F}_1, ..., \mathcal{F}_k\right)$. *Arătaţi că, în ansamblu,* $\mathcal{F}_1, ..., \mathcal{F}_n$ *sunt independente.*

CAPITOLUL 4

Spaţiul probabilizat numărabil

1. Mulţimi numărabile

1.1. Sumabilitate*. În continuare vom avea nevoie să sumăm familii numărabile de numere. Fie I o mulţime cel mult numărabilă şi fie $(a_i)_{i\in I}$ o familie de numere indexată după I. Dorim să dăm sens sumei numerelor $(a_i)_{i\in I}$. În cazul în care mulţimea I este finită nu există dubii asupra sensului. Oricum am defini o bijecţie $f:\{1,2,...,n\}\to I$, suma $a_{f(1)}+...+a_{f(n)}$ are aceeaşi valoare şi o notăm $\sum_{i\in I}a_i$. Pentru cazul numărabil există posibilitatea să definim suma prin utilizarea noţiunii de serie trecând prin numerotarea mulţimii I şi apoi se poate demonstra că suma nu depinde de modul în care s-a făcut numerotarea. Pentru a mai face mai clar aceast aspect, dăm o definiţie de la început independentă de modul în care este pusă mulţimea I în bijecţie cu $\mathbf{N}$. Pentru aceasta notăm cu $\mathcal{K}(I)$ familia tuturor submulţimilor finite din I. Vom presupune mai întâi că numerele a_i sunt toate pozitive şi vom spune că familia $(a_i)_{i\in I}$ este *sumabilă* dacă este îndeplinită următoarea condiţie:

$$\sup_{K\in\mathcal{K}(I)}\sum_{i\in K}a_i<\infty.$$

Vom nota $\sum_{i\in I}a_i:=\sup_{K\in\mathcal{K}(I)}\sum_{i\in K}a_i$ şi vom spune că acest număr este suma familiei de numere $(a_i)_{i\in I}$. Se vede imediat că orice familie finită de numere este sumabilă şi că suma are sensul dat anterior. De asemenea, se constată că dacă familia $(a_i)_{i\in I}$ este sumabilă, atunci orice subfamilie $(a_i)_{i\in J}$ rămâne sumabilă, unde $J\subset I$. Următoarea lemă este uşor de verificat.

Lema 4.1. (i) *Dacă familia de numere pozitive* $(a_i)_{i\in I}$ *este sumabilă, atunci pentru orice* $\varepsilon>0$, *există o submulţime de indici* $K\in\mathcal{K}(I)$ *astfel încât are loc relaţia* $\sum_{i\in I\setminus K}a_i\le\varepsilon$.

(ii) *Presupunem că mulţimea de indici este infinită şi fie* $I=\{i_0,i_1,...\}$ *o numerotare a sa. Atunci familia* $(a_i)_{i\in i}$ *este sumabilă dacă şi numai dacă seria* $\sum_{k=0}^{\infty}a_{i_k}$ *este convergentă. În cazul convergenţei are loc egalitatea* $\sum_{i\in I}a_i=\sum_{k=0}^{\infty}a_{i_k}$.

În continuare vom avea nevoie să lucrăm şi cu sume de familii cel mult numărabile de numere care nu sunt neapărat pozitive. Iată cum se extinde noţiunea de sumabilitate în acest caz. Fie $(a_i)_{i\in I}$ o familie de numere reale indexată după o mulţime cel mult numărabilă I. Vom spune că această familie de numere este *absolut sumabilă* dacă familia valorilor absolute, $(|a_i|)_{i\in I}$, este sumabilă.

Lema 4.2. *Presupunem că familia $(a_i)_{i\in I}$ este absolut sumabilă, iar mulţimea de indici I este numărabilă. Atunci există un unic număr $a \in \mathbf{R}$ astfel încât, pentru orice $\varepsilon > 0$, există o mulţime $K \in \mathcal{K}(I)$ cu proprietatea că*

$$\left|\sum_{i\in M} a_i - a\right| \leq \varepsilon,\ \forall M \in \mathcal{K}(I), M \supset K.$$

Dacă $I = \{i_0, i_1, ...\}$ este o numerotare a mulţimii de indici, atunci $a = \sum_{n=0}^{\infty} a_{i_n}$.

Demonstraţie. Vom porni cu o numerotare a mulţimii de indici ca în enunţ. Seria $\sum_{n=0}^{\infty} a_{i_n}$ este absolut convergentă, datorită ipotezei că familia $(a_i)_{i\in I}$ este absolut sumabilă. Notăm $a = \sum_{n=0}^{\infty} a_{i_n}$ şi vom verifica proprietatea afirmată de lemă în felul următor. Datorită faptului că familia $(|a_i|)_{i\in I}$ este sumabilă, prin aplicarea *lemei* 4.1(i), pentru $\varepsilon > 0$, există $K \in \mathcal{K}(I)$ astfel încât $\sum_{i\in L} |a_i| < \frac{\varepsilon}{2}$, pentru orice $L \in \mathcal{K}(I)$ astfel încât $L \cap K = \emptyset$. Rezultă că avem

$$\left|\sum_{i\in M} a_i - \sum_{i\in K} a_i\right| \leq \sum_{i\in M\setminus K} |a_i| < \frac{\varepsilon}{2},\ \forall M \in \mathcal{K}(I), M \supset K.$$

Dar pentru n suficient de mare, mulţimea $\{i_0, ..., i_n\}$ va conţine pe K şi atunci putem scrie

$$\left|\sum_{k=0}^{n} a_{i_k} - \sum_{i\in K} a_i\right| < \frac{\varepsilon}{2}.$$

Trecând la limită cu n obţinem $\left|a - \sum_{i\in K} a_i\right| \leq \frac{\varepsilon}{2}$. În combinaţie cu relaţia de mai înainte se deduce $\left|a - \sum_{i\in M} a_i\right| \leq \left|a - \sum_{i\in K} a_i\right| + \left|\sum_{i\in M} a_i - \sum_{i\in K} a_i\right| \leq \varepsilon$, care dă proprietatea din enunţ.

Unicitatea numărului a se demonstrează la fel ca unicitatea unei limite de şiruri. □

Dacă familia $(a_i)_{i\in I}$ este absolut sumabilă, vom nota cu $\sum_{i\in I} a_i$ numărul a dat de lema anterioară şi vom spune că el reprezintă *suma familiei de numere.*

Se verifică fără probleme că dacă $(a_i)_{i\in I}$ este o familie sumabilă de numere pozitive şi $(b_i)_{i\in I}$ este o altă familie de numere reale, indexată după aceeaşi mulţime de indici şi are loc inegalitatea $|b_i| \leq a_i$, pentru orice indice $i \in I$, atunci familia $(b_i)_{i\in i}$ este absolut sumabilă. De asemenea se poate vedea că pentru orice două familii absolut sumabile $(a_i)_{i\in I}, (b_i)_{i\in I}$, astfel că $a_i \leq b_i$, pentru orice $i \in I$, are loc inegalitatea $\sum_{i\in i} a_i \leq \sum_{i\in i} b_i$.

Mai departe vom enunţa două leme cu privire la calcule cu sume de familii absolut sumabile. Aceste calcule sunt binecunoscute în cazul seriilor de numere. Pentru demonstraţie se poate utiliza numerotarea astfel ca, prin intermediul lemelor de mai sus, să se reducă totul la probleme de serii de numere.

Lema 4.3. *Fie $\{I_l / l \in L\}$ o partiţie cel mult numărabilă a mulţimii de indici I. (Asta înseamnă că L este cel mult numărabilă.) Dacă familia de numere $(a_i)_{i\in I}$ este absolut sumabilă, atunci şi familia $\left(\sum_{i\in I_l} a_i\right)_{l\in L}$ este absolut sumabilă şi are loc următoarea formulă de grupare a sumelor*

$$\sum_{i\in I} a_i = \sum_{l\in L} \sum_{i\in I_l} a_i .$$

Reciproc, dacă $a_i \geq 0$, pentru orice $i \in I$ şi fiecare din familiile de numere $(a_i)_{i\in I_l}$, $l \in L$, este sumabilă iar familia $\left(\sum_{i\in I_l} a_i\right)_{l\in L}$ este şi ea sumabilă, atunci familia iniţială $(a_i)_{i\in I}$ este sumabilă şi este valabilă formula de mai sus.

O situaţie similară apare în cazul unei mulţimi de indici care este produsul altor mulţimi. Să presupunem acum că avem n muţimi de indici $I_1, ..., I_n$, cel mult numărabile, şi o familie de numere reale pozitive $(a_{i_1,...,i_n})_{i_1\in I_1,...,i_n\in I_n}$, indexată după toate mulţimile de indici. Dacă notăm $I = I_1 \times ... \times I_n$, această mulţime produs este cel mult numărabilă şi putem considera că familia de numere este indexată după I, scriind $(a_i)_{i\in I}$, unde cu i desemnăm un multi-indice $i = (i_1, ..., i_n)$. Are loc următoarea caracterizare a sumabilităţii în raport cu aceşti multi-indici.

Lema 4.4. *Familia de numere de mai sus este sumabilă ca familie indexată după I dacă şi numai dacă ea este iterativ sumabilă în sensul următor: pentru fiecare sistem de indici fixaţi $i_1 \in I_1, ..., i_{n-1} \in I_{n-1}$, familia $(a_{i_1,...,i_n})_{i_n\in I_n}$ este sumabilă, de asemenea, pentru fiecare sistem de indici $i_1 \in I_1, ..., i_{n-2} \in I_{n-2}$ familia $\left(\sum_{i_n\in I_n} a_{i_1,...,i_{n-2},i_{n-1},i_n}\right)_{i_{n-1}\in I_{n-1}}$ este sumabilă, şi aşa mai departe, familia*

$$\left(\sum_{i_2\in I_2}\left(\cdots\left(\sum_{i_{n-1}\in I_{n-1}}\left(\sum_{i_n\in I_n} a_{i_1,i_2,...,i_{n-1},i_n}\right)\right)\cdots\right)\right)_{i_1\in I_1}$$

este sumabilă. În plus are loc relaţia

$$\sum_{i\in I} a_i = \sum_{i_1\in I_1}\left(\sum_{i_2\in I_2}\cdots\left(\sum_{i_{n-1}\in I_{n-1}}\left(\sum_{i_n\in I_n} a_{i_1,i_2,...,i_{n-1},i_n}\right)\right)\cdots\right).$$

Ca o consecinţă a acestei leme se observă că ordinea de sumare poate fi făcută în orice fel. Verificarea lemei o lăsăm cititorului. (Cazul când $n = 2$ este de fapt un caz particular al lemei anterioare.)

1.2. Măsuri pe o mulţime cel mult numărabilă. În continuare, vom nota cu E o mulţime cel mult numărabilă. Din punctul de vedere al teoriei măsurii, spaţiul acesta este foarte simplu. Mulţimea părţilor $\mathcal{P}(E)$ reprezintă singura $\sigma-$algebră de interes pe E. Dacă E este infinită, mai există un obiect distinct, anume algebra generată de mulţimile formate dintr-un punct. Această algebră constă din toate mulţimile finite şi cele cu complementara finită. Dar ea nu are nici un rol anume.

Lema următoare arată că o măsură de probabilitate pe această $\sigma-$algebră este complet determinată de valorile pe care le dă mulţimilor cu un singur element. Fie μ o măsură de probabilitate pe $\mathcal{P}(E)$ şi să notăm $\mu_e = \mu(\{e\}), e \in E$ valorile pe care le ia pe punctele lui E.

Lema 4.5. *Fie $(A_i)_{i\in I}$ o familie cel mult numărabilă de părţi ale lui E astfel că $A_i \cap A_j = \emptyset$, dacă $i \neq j$. Atunci are loc relaţia*

$$\mu\left(\bigcup_{i\in I} A_i\right) = \sum_{i\in I} \mu(A_i).$$

Valoarea pe care o dă măsura unei mulţimi arbitrare $A \subset E$ o putem exprima astfel:

$$\mu(A) = \sum_{e\in A} \mu_e. \quad (*)$$

În particular, două măsuri care coincid pe puncte sunt identice.

Demonstraţie. Prima relaţie este trivial verificată în cazul în care I este finită. Dacă I este numărabilă se numerotează $I = \{i_0, i_1, ...\}$ şi se scrie

$$\mu\left(\bigcup_{i\in I} A_i\right) = \mu\left(\bigcup_{k=0}^{\infty} A_{i_k}\right) =$$

$$= \lim_{n\to\infty} \mu\left(\bigcup_{k=0}^{n} A_{i_k}\right) = \lim_{n\to\infty} \sum_{k=0}^{n} \mu(A_{i_k}) = \sum_{i\in I} \mu(A_i),$$

ceea ce demonstrează relaţia. Ultima afirmaţie a lemei rezultă imediat din relaţia demonstrată. □

Construcţia de măsuri pe o mulţime cel mult numărabilă se bazează pe relaţia (*) din lemă.

Lema 4.6. *Fie $(\mu_e)_{e\in E}$ o familie de numere reale care satisface condiţiile următoare:*

1) $\mu_e \in [0,1]$ pentru orice $e \in E$,

2) $\sum_{e\in E} \mu_e = 1$.

Atunci relaţia (*) din lema de mai sus defineşte o măsură de probabilitate pe E.

Demonstraţie. Este uşor de văzut că definiţia dată implică finit aditivitatea măsurii. Pentru a verifica $\sigma-$ aditivitatea, pornim cu un şir de mulţimi disjuncte $(A_n)_{n\in\mathbf{N}}$ şi notăm $A = \cup_n A_n$. Avem

$$\mu(A) = \sum_{e\in A}\mu_e = \sum_{n\in\mathbf{N}}\sum_{e\in A_n}\mu_e = \sum_{n\in\mathbf{N}}\mu(A_n),$$

unde la al doilea semn egal am utilizat *lema* 4.4. □

Măsuri Dirac.

Cele mai simple măsuri sunt cele suportate de câte un singur punct. Fie $e_0 \in E$ fixat şi $A \subset E$, arbitrară. Definim măsura Dirac suportată de punctul e_0, pe care o notăm δ_{e_0}, prin

$$\delta_{e_0}(A) = 1_A(e_0).$$

Altfel spus, avem $\delta_{e_0}(A) = 1$, dacă $e_0 \in A$ şi $\delta_{e_0}(A) = 0$, dacă $e_0 \notin A$. Dacă utilizăm notaţia de mai sus cu $\mu = \delta_{e_0}$ pentru a exprima valorile pe puncte, avem $\mu_e = 0$, pentru orice $e \neq e_0$ şi $\mu_{e_0} = 1$.

Orice măsură de probabilitate μ poate fi scrisă şi ca o serie de măsuri Dirac. Dacă valorile sale pe puncte sunt reprezentate de familia de numere $(\mu_e)_{e\in E}$, atunci putem scrie $\mu = \sum_{e\in E}\mu_e\delta_e$, egalitate care se verifică pentru fiecare mulţime A :

$$\mu(A) = \sum_{e\in A}\mu_e = \sum_{e\in A}\mu_e\delta_e(A) = \sum_{e\in E}\mu_e\delta_e(A).$$

Integrala.

Noţiunea de integrală în raport cu o măsură de probabilitate μ pe $\mathcal{P}(E)$ se defineşte în felul următor. Fiind dată o funcţie $f : E \to R$, se spune că aceasta este *integrabilă* dacă familia $(f(e)\mu_e)_{e\in E}$ este absolut sumabilă. Numărul

$$\int f d\mu := \sum_{e\in E} f(e)\mu_e$$

este numit *integrala* funcţiei f în raport cu măsura μ. Se mai utilizează şi notaţia $\int_E f(e)\mu(de) = \int f d\mu$.

2. Produsul de măsuri discrete

Pentru a construi modele probabiliste ce corespund mai multor experimente ale căror rezultate sunt independente, se apelează la produsele de spaţii probabilizate. Vom considera mai întâi cazul a două mulţimi finite pentru a înţelege mai bine notaţia ce intervine.

Să presupunem că avem date mulţimile finite $E = \{e_1, ..., e_m\}$ şi $F = \{f_1, ..., f_n\}$, cu măsurile de probabilitate $\mu = (\mu_1, ..., \mu_m), \nu = (\nu_1, ..., \nu_n)$, unde am notat $\mu_i = \mu(\{e_i\})$ şi $\nu_j = \nu(\{f_j\})$. Deci au loc relaţiile $\mu_1 + ... + \mu_m = 1$ şi $\nu_1 + ... + \nu_n = 1$. Pe spaţiul produs $E \times F$ se poate defini o măsură λ punând $\lambda_{ij} = \lambda(\{(e_i, f_j)\}) := \mu_i\nu_j$.

Într-adevăr, se verifică imediat condiţia

$$\sum_{i=1,j=1}^{m,n} \lambda_{ij} = \sum_{i=1}^{m} \sum_{j=1}^{n} \mu_i \nu_j = \left(\sum_{i=1}^{m} \mu_i \right) \left(\sum_{j=1}^{n} \nu_j \right) = 1.$$

Măsura λ se numeşte măsura produs a celor două măsuri şi se notează $\lambda = \mu \bigotimes \nu$.

Această construcţie se poate generaliza la produsul a n mulţimi cel mult numărabile. Fie $E_1, ..., E_n$ mulţimi cel mult numărabile pe care avem definite, respectiv, măsurile de probabilitate $\mu_1, ..., \mu_n$. Analog cazului anterior folosim notaţia $\mu_{l,e} = \mu_l (\{e\}), e \in E_l, l = 1, ..., n$ pentru măsura punctelor din spaţiile respective. Pe spaţiul produs $E = \prod_{i=1}^n E_i$, care este de asemenea o mulţime cel mult numărabilă, definim măsura μ, punând valoarea pe puncte

$$\mu (\{(e_1, ..., e_n)\}) := \mu_{1,e_1} ... \mu_{n,e_n}, \ \forall (e_1, ..., e_n) \in E.$$

Se verifică relaţia

$$\begin{aligned} \textstyle\sum_{(e_1,...,e_n)\in E} \mu (\{(e_1, ..., e_n)\}) &= \textstyle\sum_{e_1\in E_1} ... \sum_{e_n\in E_n} \mu_{1,e_1} ... \mu_{n,e_n} \\ &= \left(\textstyle\sum_{e_1\in E_1} \mu_{1,e_1}\right) ... \left(\textstyle\sum_{e_n\in E_n} \mu_{n,e_n}\right) = 1, \end{aligned}$$

care arată că valorile puse pe punctele spaţiului produs definesc o măsură de probabilitate pe E. Această măsură este notată cu $\mu = \mu_1 \bigotimes ... \bigotimes \mu_n$.

Lema 4.7. *Fie $B_l \subset E_l, l = 1, ..., n$ mulţimi arbitrare. Atunci are loc formula*

$$\mu_1 \bigotimes ... \bigotimes \mu_n (B_1 \times ... \times B_n) = \mu_1 (B_1) ... \mu_n (B_n).$$

Demonstraţie. Mulţimea produs $B_1 \times ... \times B_n$ se poate exprima ca reuniunea punctelor sale în felul următor

$$B_1 \times ... \times B_n = \bigcup_{(e_1,...,e_n)\in B_1\times...\times B_n} \{(e_1, ..., e_n)\}.$$

Deoarece mulţimile ce participă la această reuniune sunt disjuncte rezultă

$$\begin{aligned} \mu (B_1 \times ... \times B_n) &= \textstyle\sum_{(e_1,...,e_n)\in B_1\times...\times B_n} \mu (\{(e_1, ..., e_n)\}) = \\ = \textstyle\sum_{e_1\in B_1} ... \sum_{e_n\in B_n} \mu (\{(e_1, ..., e_n)\}) &= \textstyle\sum_{e_1\in B_1} ... \sum_{e_n\in B_n} \mu_{1,e_1} ... \mu_{n,e_n} = \\ &= \left(\textstyle\sum_{e_1\in B_1} \mu_{1,e_1}\right) ... \left(\textstyle\sum_{e_n\in B_n} \mu_{n,e_n}\right) = \mu_1 (B_1) ... \mu_n (B_n), \end{aligned}$$

care este relaţia de demonstrat. □

Pentru fiecare $l = 1, ..., n$ definim pe E câte o partiţie numărabilă $\mathcal{A}_l = (A_{l,e})_{e\in E_l}$ punând

$$A_{l,e} = \{(e_1, ..., e_n) \in E / \ e_l = e\} = E_1 \times ... \times E_{l-1} \times \{e\} \times E_{l+1} \times ... \times E_n, \ \ e \in E_l.$$

Următoarea lemă descrie $\sigma-$ algebrele generate de aceste partiţii.

Lema 4.8. *O mulţime $A \subset E$ aparţine lui $\sigma(\mathcal{A}_l)$ dacă şi numai dacă există o mulţime $B \subset E_l$ astfel ca să aibă loc reprezentarea*

$$A = E_1 \times ... \times E_{l-1} \times B \times E_{l+1} \times ... \times E_n.$$

DEMONSTRAŢIE. Mulţimea care apare în reprezentarea din enunţ se exprimă şi sub forma

$$E_1 \times ... \times E_{l-1} \times B \times E_{l+1} \times ... \times E_n = \bigcup_{e \in B} A_{l,e}.$$

Această egalitate arată clar că o mulţime care se reprezintă ca în enunţ este în $\sigma(\mathcal{A}_l)$. Fie acum $A \in \sigma(\mathcal{A}_l)$. Pe de altă parte, se poate uşor verifica că familia mulţimilor de acest tip,

$$\mathcal{H} = \{E_1 \times ... \times E_{l-1} \times B \times E_{l+1} \times ... \times E_n /\ B \subset E_l\},$$

este o σ-algebră. Aceasta implică $\mathcal{H} = \sigma(\mathcal{A}_l)$. □

Propoziţia 4.1. *$\sigma-$ algebrele $\sigma(\mathcal{A}_1), ..., \sigma(\mathcal{A}_n)$ sunt independente relativ la măsura produs $\mu = \mu_1 \bigotimes ... \bigotimes \mu_n$.*

DEMONSTRAŢIE. Fie $A_l \in \sigma(\mathcal{A}_l), l = 1, ..., n$. Din lema anterioară ştim că pentru fiecare l există mulţimea $B_l \subset E_l$ astfel încât să aibă loc reprezentarea

$$A_l = E_1 \times ... \times E_{l-1} \times B_l \times E_{l+1} \times ... \times E_n.$$

Cu această notaţie putem scrie

$$A_1 \cap ... \cap A_n = B_1 \times ... \times B_n,$$

ceea ce duce la

$$\mu(A_1 \cap ... \cap A_n) = \mu_1(B_1) ... \mu_n(B_n).$$

Pe de altă parte, se poate calcula şi măsura $\mu(A_l) = \mu_l(B_l)$, ceea ce permite să deducem

$$\mu(A_1 \cap ... \cap A_n) = \mu(A_1) ... \mu(A_n),$$

adică exact relaţia care arată independenţa $\sigma-$ algebrelor. □

3. Repartiţia unei variabile aleatoare

În această secţiune vom presupune dat un spaţiu probabilizat $(\Omega, \mathcal{F}, P)$. În prezenţa probabilităţii P denumirea de aplicaţie măsurabilă capătă un nou nume.

Definiţia 4.1. *Fie E o mulţime cel mult numărabilă. O aplicaţie $X : \Omega \to E$ care este măsurabilă în raport cu spaţiul măsurabil $(\Omega, \mathcal{F})$ va fi numită* variabilă aleatoare *cu valori în E.*

Reamintim că aceasta înseamnă că fiecare din mulţimile $X^{-1}(B), B \subset E$, aparţine lui $\mathcal{F}$.

O variabilă aleatoare $X : \Omega \to E$, cu valori în mulţimea cel mult numărabilă E, poate fi utilizată pentru a transporta măsura de probabilitate P de pe Ω pe E. Anume, variabilei X i se asociază următoarea aplicaţie $P_X : \mathcal{P}(E) \to [0,1]$, definită prin $P_X(B) = P(X^{-1}(B))$. Se poate uşor verifica că P_X este o măsură de probabilitate pe E. Ea este numită *repartiţia* lui X. Evident că repartiţia aceasta este suportată de imaginea $E' = X(\Omega)$, în sensul că $P_X(E \backslash E') = 0$. Ea este determinată de valorile pe puncte $P_X(\{e\}) = P(X^{-1}(e)), e \in E$, prin intermediul valorilor lui P pe atomii partiţiei asociate variabilei X, $\mathcal{A} = \{X^{-1}(e) / e \in E'\}$. Pentru $e \in E \backslash E'$, avem $X^{-1}(e) = \emptyset$, deci $P_X(\{e\}) = 0$, în acest caz. Dacă $E' = \{e_1, ..., e_n\}$ este o mulţime finită, repartiţia este reprezentată schematic de un tablou de tipul

$$P_X = \left(\begin{array}{cccc} e_1 & e_2 & ... & e_n \\ p_1 & p_2 & ... & p_n \end{array} \right),$$

în care pe linia de jos sunt înşirate valorile probabilităţilor ce corespund puncelor, $p_i = P(X^{-1}(e_i)), i = 1, ..., n$. Mai putem scrie

$$P_X = p_1 \delta_{e_1} + ... + p_n \delta_{e_n},$$

În general se poate scrie $P_X = \sum_{e \in E'} P(X^{-1}(e)) \delta_e$.

Dacă μ este o măsură de probabilitate pe E, atunci ea poate fi privită ca repartiţia variabilei aleatoare $X : E \to E$, dată de aplicaţia identitate, definită prin $X(e) = e$, pentru orice $e \in E$. În acest caz se consideră ca domeniu de definiţie al variabilei spaţiul probabilizat $(E, \mathcal{P}(E), \mu)$. De aceea se obişnuieşte a numi repartiţie orice măsură de probabilitate pe E. Este important să facem următoarea observaţie.

Observaţia 4.1. *Fie* $(\Omega_i, \mathcal{F}_i, P_i), i = 1, 2$, *două spaţii probabilizate. Dacă variabilele aleatoare* $X_i : \Omega_i \to E, i = 1, 2$, *cu valori în aceeaşi mulţime cel mult numărabilă, au aceeaşi reparţie, iar* $f : E \to F$ *este o aplicaţie arbitrară, atunci variabilele* $Y_i = f(X_i), i = 1, 2$ *au şi ele aceeaşi repartiţie.*

Repartiţiile variabilelor aleatoare au un rol important în legătură cu proprietăţile de independenţă ale acestora.

Definiţia 4.2. *Fie* $X_i : \Omega \to E_i, i = 1, ..., n$, *variabile aleatoare cu valori în mulţimi cel mult numărabile. Vom spune că acestea sunt* independente *dacă* $\sigma-$*algebrele* $\sigma(X_1), ..., \sigma(X_n)$ *sunt independente.*

Conform cu *lema* 2.3, independenţa variabilelor $X_i, i = 1, ..., n$ este echivalentă cu verificarea relaţiei

$$P\left(X_1^{-1}(A_1) \cap ... \cap X_n^{-1}(A_n)\right) = P\left(X_1^{-1}(A_1)\right) ... P\left(X_n^{-1}(A_n)\right),$$

pentru orice mulţimi $A_i \subset E_i, i = 1, ..., n$. Pe aceeaşi linie, facem observaţia că un număr finit de evenimente, $A_1, ..., A_n$, sunt independente dacă şi numai dacă variabilele aleatoare, cu valori în $\{0,1\}$, constituite din funcţiile caracteristice $1_{A_1},...,1_{A_n}$ sunt independente.

Următoarea propoziţie caracterizează independenţa în termenii repartiţiilor.

Propoziţia 4.2. *Presupunem că $X_i : \Omega \to E_i, i = 1,2$ sunt două variabile aleatoare cu valori în spaţiile cel mult numărabile $E_i, i = 1,2$ şi notăm $Z : \Omega \to E_1 \times E_2$, aplicaţia definită prin $Z(\omega) = (X_1(\omega), X_2(\omega))$. Variabilele aleatoare X_1 şi X_2 sunt independente dacă şi numai dacă variabila Z are repartiţia exprimată sub forma produsului*

$$P_Z = P_{X_1} \bigotimes P_{X_2}.$$

Demonstraţie. Să presupunem că variabilele date sunt independente. Rezultă că pentru orice puncte $e_1 \in E_1$ şi $e_2 \in E_2$ are loc relaţia

$$P\left(X_1^{-1}(e_1) \cap X_2^{-1}(e_2)\right) = P\left(X_1^{-1}(e_1)\right) P\left(X_2^{-1}(e_2)\right).$$

Pe de altă parte, ţinând cont că $X_1^{-1}(e_1) \cap X_2^{-1}(e_2) = Z^{-1}(e_1,e_2)$, relaţia aceasta devine

$$P_Z(\{(e_1,e_2)\}) = P_{X_1}(\{e_1\}) P_{X_2}(\{e_2\}) = P_{X_1} \bigotimes P_{X_2}(\{(e_1,e_2)\}),$$

pentru orice pereche $(e_1,e_2) \in E_1 \times E_2$. Rezultă că măsurile P_Z şi $P_{X_1} \bigotimes P_{X_2}$ coincid.

Reciproc, să presupunem că măsurile P_Z şi $P_{X_1} \bigotimes P_{X_2}$ coincid. Fie mulţimile $A_1 \subset E_1, A_1 \subset E_1$ arbitrare. Din relaţia $X_1^{-1}(A_1) \cap X_2^{-1}(A_2) = Z^{-1}(A_1 \times A_2)$ se deduce

$$P\left(X_1^{-1}(A_1) \cap X_2^{-1}(A_2)\right) = P_Z(A_1 \times A_2) = P_{X_1} \bigotimes P_{X_2}(A_1 \times A_2) =$$
$$= P_{X_1}(A_1) P_{X_2}(A_2) = P\left(X_1^{-1}(A_1)\right) P\left(X_2^{-1}(A_2)\right),$$

ceea ce arată că variabilele X_1 şi X_2 sunt independente. La penultimul semn egal am utilizat lema 4.7. $\square$

Propoziţia demonstrată se generalizează, cu aceeaşi demonstraţie, la cazul a n variabile.

Propoziţia 4.3. *Variabilele $X_1,...,X_n$, sunt independente dacă şi numai dacă repartiţia variabilei $Z = (X_1,...,X_n)$ satisface relaţia*

$$P_Z = P_{X_1} \bigotimes ... \bigotimes P_{X_n}.$$

Această propoziţie permite construirea de modele probabiliste pentru variabile aleatoare independente cu repartiţii date. Mai precis, fie $E_i, i = 1,...,n$, mulţimi cel mult numărabile pe care sunt definite măsurile de probabilitate $\mu_i, i = 1,...,n$. Notăm $W = E_1 \times ... \times E_n$ spaţiul produs şi $X_i : W \to E_i$ aplicaţia de proiecţie pe componenta i, definită prin $X_i((e_1,...,e_n)) = e_i$, pentru orice $(e_1,...,e_n) \in W$. Pe spaţiul W, înzestrat cu $\sigma-$ algebra $\mathcal{P}(W)$, introducem măsura produs

$Q = \mu_1 \bigotimes ... \bigotimes \mu_n$. Cu aceste notaţii putem spune că $(W, \mathcal{P}(W), Q)$ este un spaţiu probabilizat şi că $X_1, ..., X_n$ sunt variabile aleatoare. În plus are loc următorul rezultat.

Propoziţia 4.4. *Repartiţia lui X_i este μ_i, pentru orice $i = 1, ..., n$, iar variabilele aleatoare $X_1, ..., X_n$ sunt independente.*

Demonstraţie. Vom calcula mai întâi repartiţia lui X_i. Pentru $e \in E_i$ are loc egalitatea

$$X_i^{-1}(e) = E_1 \times ... \times E_{i-1} \times \{e\} \times E_{i+1} \times ... \times E_n,$$

care se verifică imediat. De aceea avem

$$\begin{aligned} Q_{X_i}(e) &= Q\left(X_i^{-1}(e)\right) = \\ &= \mu_1(E_1)...\mu_{i-1}(E_{i-1})\,\mu_i(e)\,\mu(E_{i+1})...\mu(E_n) = \mu_i(e), \end{aligned}$$

ceea ce probează egalitatea $Q_{X_i} = \mu_i$.

Pe de altă parte, aplicaţia $Z = (X_1, ..., X_n) : W \to E_1 \times ... \times E_n$ este aplicaţia identică, şi de aceea avem $Q_Z = \mu_1 \bigotimes ... \bigotimes \mu_n$. Propoziţia anterioară ne asigură că variabilele noastre sunt independente. □

Exemplul 1.

Un zar măsluit are pusă o greutate în vârful adiacent feţelor cu numerele $1, 3, 5$. Această greutate face ca feţele opuse, adică cele ce poartă numerele $2, 4, 6$, să apară cu o probabilitate mai mare la aruncare. Deci feţele cu numere impare au probabilităţile de a ieşi egale între ele şi la fel feţele cu numere pare. Dar o faţă cu număr par are mai multe şanse de a ieşi decât una cu număr impar. Ne propunem să construim un model probabilist care să descrie seriile de mai multe aruncări cu un astfel de zar.

Să notăm cu r raportul dintre şansele de a ieşi o faţă pară şi una impară la o aruncare. Deci prin ipoteză avem $r > 1$. Vom determina mai întâi în funcţie de r probabilităţile de a ieşi fiecare faţă. Notând p_i probabilitatea de a ieşi faţa i, vom putea scrie $\sum_{i=1}^{6} p_i = 1, p_1 = p_3 = p_5, p_2 = p_4 = p_6$ şi $\frac{p_2}{p_1} = r$. Rezultă $3p_1 + 3rp_1 = 1$, deci $p_1 = \frac{1}{3(1+r)}$ şi $p_2 = \frac{r}{3(1+r)}$.

Mai departe notăm $E = \{1, 2, 3, 4, 5, 6\}$, mulţimea care descrie posibilităţile ce pot apărea în urma unei aruncări. Pe această mulţime definim măsura μ prin numerele $\mu_1 = \mu_3 = \mu_5 = \frac{1}{3(1+r)}, \mu_2 = \mu_4 = \mu_6 = \frac{r}{3(1+r)}$. Dacă dorim să modelăm seriile de n aruncări notăm $\Omega = E^n = \underbrace{E \times ... \times E}_{n}$ şi $X_l : \Omega \to E$ este proiecţia pe componenta $l, l = 1, ..., n$. Spaţiul Ω reprezintă mulţimea tuturor rezultatelor posibil a fi obţinute în urma unei serii de n aruncări cu zarul. Variabila X_l descrie rezultatul de la cea de a l -a aruncare cu zarul, iar, în ansamblu, cele n variabile descriu rezultatele succesive într-o serie de n aruncări. Este clar că diferitele aruncări cu zarul sunt independente. De aceea punem pe spaţiul Ω măsura produs $P = \mu \bigotimes ... \bigotimes \mu$, care face variabilele X_l independente şi identic repartizate cu repartiţia μ. Deci $(\Omega, \mathcal{P}(\Omega), P)$ descrie seriile de n aruncări cu zarul măsluit. $\square$

Exemplul 2.

Revenim acum la modelul bilei întoarse. Am notat $M = \{1, 2, ..., m\}$ şi $\Omega = M \times ... \times M$, produsul cu un număr de n factori care modelează seriile de n extrageri succesive cu revenire. Măsura de probabilitate P, care defineşte şanse egale pentru fiecare element $\omega = (x_i, ..., x_n) \in \Omega$ a fost definită prin expresia $P(\{\omega\}) = \frac{1}{m^n}$. Vom obseva acum că această măsură este o măsură produs. Într-adevăr, dacă notăm cu μ măsura uniformă pe M, definită prin $\mu(\{l\}) = \frac{1}{m}$, pentru orice $l \in M$, atunci se poate constata că are loc egalitatea $P = \mu \bigotimes ... \bigotimes \mu$. Să mai observăm că variabila care reprezintă proiecţia pe componenta i, definită prin $X_i : \Omega \to M, X_i(\{(x_1, ..., x_n)\}) = x_i$ este variabila care indică numărul bilei extrase la cea de a $i-$a extragere. Prin urmare, putem spune că modelul probabilist introdus pentru schema bilei întoarse asigură independenţa variabilelor $X_i, i = 1, ..., n$ şi faptul că acestea sunt uniform repartizate. Este clar că aceste două proprietăţi ale acestor variabile sunt în concordanţă cu ceea ce aşteptăm de la model. Pe de altă parte, se poate verifica că P este singura măsură de probabilitate pe Ω care face din proiecţii variabile independente şi uniform repartizate. Cu alte cuvinte aceste proprietăţi determină modelul.

Exemplul 3*.

Să reflectăm acum la modelarea problemei zilelor de naştere tratată în *capitolul* 2. Am adoptat pentru această problemă modelul extragerilor cu revenire. Dacă aceptăm ca trăsături obligatorii ale acestei probleme faptul că zilele de naştere ale persoanelor din grup sunt reprezentate de variabile aleatoare independente şi uniform repartizate, înseamnă că trebuie să acceptăm ca unic model cel adoptat în capitolul 2.

Vom discuta acum o altă variantă de modelare a problemei zilelor de naştere. Să presupunem că de fapt problema se referă în mod concret la o anumită populaţie, formată dintr-un anumit număr de indivizi, m, pentru care se ştie numărul de persoane ce au zilele de naştere în fiecare zi a anului. Atunci, un grup de n indivizi luaţi la întâmplare din această populaţie ($n \leq 365$) sunt modelaţi de o extragere fără ordine de n bile dintr-o urnă cu m bile. Desigur că acesta este modelul cel mai limpede şi mai uşor de acceptat. Vom examina acest model în cazul în care presupunem că populaţia în cauză este compusă dintr-un număr egal de indivizi născuţi pentru fiecare zi a anului şi facem presupunerea că avem doar 365 de zile în fiecare an. Vom trece apoi la limită cu $m \to \infty$, pentru a vedea că se ajunge la acelaşi model ca în paragraful dedicat problemei din capitolul 2.

Dacă notăm cu l numărul indivizilor ce au aceeaşi zi de naştere, atunci numărul total de indivizi va fi $m = 365l$. Vom face modelarea extragerilor de n bile fără ordine bazându-ne pe schema bilei neîntoarse.

Pentru aceasta identificăm populaţia celor m indivizi cu mulţimea $M = \{1, ..., m\}$ şi notăm cu A_j mulţimea indivizilor născuţi în cea de a $j-$a zi a anului, $j = 1, ...365$. Rezultă că familia $(A_j)_{j=1,...,365}$ formează o partiţie a lui M şi cardinalul fiecăreia din mulţimile A_j este l. Spaţiul tuturor posibilităţilor este cel ce corespunde schemei bilei neîntoarse

$$W = \{(x_1, ..., x_n) \,/\, x_i \in M,\ x_i \neq x_j, \forall i \neq j\}.$$

Probabilitatea ce o considerăm pe W o notăm cu Q şi, aşa cum se ştie, ea acordă ponderea $Q(\{(x_1, ..., x_n)\}) = \frac{1}{cardW} = \frac{1}{A_m^n}$ pentru fiecare punct $w = (x_1, ..., x_n) \in W$. Din punctul de vedere al problemei noastre, ceea ce interesează sunt doar zilele de naştere ale indivizilor dintr-un grup w. Vom nota $E = \{1, ..., 365\}$ mulţimea zilelor anului şi vom defini apoi o aplicaţie $f : M \to E$, în felul următor: $f(x) = k$ dacă şi numai dacă $x \in A_k$. Aceasta este aplicaţia care asociază fiecărui individ ziua sa de naştere. Vom introduce spaţiul $\Omega = E^n$, care înregistrează toate configuraţiile de zile de naştere posibil a corespunde unui grup $w \in W$. Aplicaţia f permite a defini o variabilă aleatoare, $F : W \to \Omega$, prin $F(x_1, ..., x_n) = (f(x_1), ..., f(x_n))$. Deoarece pentru problema ce ne interesează evenimentele elementare ale lui W sunt cercetate doar prin intermediul imaginii lor prin F, rezultă că putem lua ca model spaţiul Ω cu repartiţia variabilei F, care este o probabilitate pe Ω obţinută prin transportul lui Q prin intermediul variabilei F. Deoarece Q depinde de numărul l, ce îl vom presupune variabil, tinzând chiar la infinit, vom pune $P_l = Q_F$ pentru a desemna repartiţia aceasta. Mai departe vom compara probabilitatea P_l cu probabilitetea şanselor egale pe Ω. Mai precis, fie P probabilitatea şanselor egale pe Ω, definită prin $P(\{\omega\}) = \frac{1}{card\Omega} = \frac{1}{365^n}$, pentru orice punct $\omega \in \Omega$. Vom arăta că P_l tinde la P, atunci când l tinde la infinit, în sensul următor:

$$\lim_{l \to \infty} P_l(\{\omega\}) = P(\{\omega\}),\ \forall \omega \in \Omega.$$

Pentru a verifica această relaţie vom fixa un element arbitrar $\omega \in \Omega$ şi vom calcula cardinalul mulţimii $F^{-1}(\omega)$. Să zicem că elementul dat are descrierea pe componente $\omega = (k_1, ..., k_n)$. Componentele nu sunt neapărat toate distincte şi, pentru a clarifica acest lucru vom proceda astfel. Pentru fiecare zi a anului, $k \in E$, vom pune $J_k = \{i \leq n / \ k_i = k\}$, pentru a desemna indicii componentelor ce coincid cu ziua k, iar apoi $j_k = cardJ_k$, pentru a număra de câte ori se repetă ziua k între componentele lui ω. Mulţimile J_k pot fi vide pentru anumite valori $k \in E$, astfel că $j_k = 0$ într-o astfel de situaţie. În orice caz avem $\sum_{k=1}^{365} j_k = n$, numărul total de componente ale lui ω. Fie acum $w = (x_1, ..., x_n) \in W$, astfel ca $F(w) = \omega$. Rezultă că pentru un $k \in E$ fixat are loc relaţia $f(x_i) = k$, pentru orice $i \in J_k$. Dacă un alt element $w' = (x_1', ..., x_n') \in W$ este construit, pornind de la w, prin înlocuirea componentelor de pe poziţiile J_k cu alte puncte din mulţimea A_k rezultă că vom avea $F(w') = \omega$. Dar există $A_l^{j_k} = l(l-1)...(l-j_k+1)$ posibilităţi de a obţine elemente distincte din W prin acest procedeu. Cu convenţia $A_l^0 = 1$, putem exprima numărul total al elementelor din $F^{-1}(\omega)$ sub forma $\prod_{k\in E} A_l^{j_k}$.

Putem concluziona că

$$P_l(\{\omega\}) = Q\left(F^{-1}(\omega)\right) = \frac{cardF^{-1}(\omega)}{cardW} = \frac{\prod_{k\in E} A_l^{j_k}}{A_m^n} =$$

$$= \frac{l(l-1)...(l-j_1+1)...l(l-1)...(l-j_{365}+1)}{365l(365l-1)...(365l-n+1)}.$$

Deoarece, aşa cum am remarcat, are loc relaţia $\sum_{k=1}^{365} j_k = n$, rezultă că în produsul de la numărător sunt n factori. Jos sunt tot atâţia factori şi atunci se poate scoate în factor şi simplifica cu l, astfel că fracţia anterioară devine

$$= \frac{\left(1-\frac{1}{l}\right)...\left(1-\frac{j_1+1}{l}\right)...\left(1-\frac{1}{l}\right)...\left(1-\frac{j_{365}+1}{l}\right)}{365\left(365-\frac{1}{l}\right)...\left(365-\frac{n-1}{l}\right)}.$$

Trecând la limită este clar că obţinem $\lim_{l\to\infty} P_l(\{\omega\}) = \frac{1}{(365)^n} = P(\{\omega\})$. Rezultă că modelul utilizat de noi în paragraful dedicat problemei zilelelor de naştere, din *capitolul* 2, corespunde unei situaţii ideale, în care facem trei presupuneri simplificatoare: 1) presupunem că anul ar avea întotdeauna 365 de zile, 2) numărul de persoane născute în aceeaşi zi este acelaşi pentru fiecare zi a anului, 3) populaţia totală luată în discuţie este infinită, în sensul dat de procesul de trecere la limită descris mai sus.

Dar aproximarea dată de trecerea la limită cu l converge destul de rapid. De exemplu, evaluările pentru mulţimea

$$\Lambda = \{\omega = (x_1, ..., x_n) \in \Omega \mid x_i \neq x_j \ \ , \ \forall \ i \neq j\},$$

făcute cu măsura P_l sau cu P sunt foarte apropiate. Într-adevăr, ştim că $P(\Lambda) = \frac{(365-1)(365-2)...(395-n+1)}{(365)^{n-1}}$. Pe de altă parte, nu este greu de calculat cardinalul mulţimii $F^{-1}(\Lambda)$ pe baza ideilor expuse anterior. Se obţine $cardF^{-1}(\Lambda) = A_{365}^n \times l^n$. Atunci avem

$$P_l(\Lambda) = \frac{l^n 365(365-1)(365-2)...(365-n+1)}{365l(365l-1)...(365l-n+1)} =$$

$$= P(\Lambda)\frac{l^{n-1}(365)^{n-1}}{(365l-1)...(365l-n+1)}.$$

Vom nota $\theta(n,l) = \frac{l^{n-1}(365)^{n-1}}{(365l-1)...(365l-n+1)}$ şi vom evalua acest număr scriindu-l sub forma

$$\theta(n,l) = \left(1+\frac{1}{365l-1}\right)...\left(1+\frac{n-1}{365l-n+1}\right).$$

Pentru $l \geq 100$ avem $\theta(n,l) \leq \left(1+\frac{1}{36478}\right)...\left(1+\frac{n-1}{36478}\right)$, iar ultimul produs se majorează conform *lemei* 1.1 cu $e^{\frac{n(n-1)}{2\times 36478}}$. Pentru $n \leq 23$, avem atunci $\theta(n,l) \leq e^{\frac{23\times 22}{2\times 36478}} \leq 1+\frac{23\times 22}{2\times 36478}+\left(\frac{23\times 22}{2\times 36478}\right)^2$. (am utilizat inegalitatea $e^x \leq 1+x+x^2$, care are loc pentru $x \in [0,1]$). Calculând numeric expresiile din dreapta obţinem $\theta(n,l) \leq 1,007$. Deci tragem concluzia că $P_l(\Lambda) = P(\Lambda)\theta(n,l)$, unde $1 \leq \theta(n,l) \leq 1,007$, pentru $l \geq 100$. □

Independenţa partiţiilor.

Examinând demonstraţia *propoziţiei* 4.2 se constată că ea arată mai mult decât scrie în enunţ. Anume, rezută că independenţa a două variabile $X_i : \Omega \to E_i, i = 1,2$ cu valori în spaţiile cel mult numărabile $E_i, i = 1,2$, este implicată de relaţia de independenţă verificată pe partiţiile asociate, adică de relaţiile:

$$P\left(X_1^{-1}(e_1) \cap X_2^{-1}(e_2)\right) = P\left(X_1^{-1}(e_1)\right) P\left(X_2^{-1}(e_2)\right),$$

cu fiecare din punctele $e_i \in E_i'$,unde $E_i' = X_i(\Omega), i = 1,2$.

Această observaţie o vom generaliza acum. Fie $\mathcal{A}_l = (A_{l,i})_{i\in I_l}, l = 1,...,n$ partiţii ai căror atomi aparţin lui $\mathcal{F}$. Evident că dacă σ -algebrele generate de aceste partiţii sunt independente, atunci putem scrie

$$P(A_{1,i_1} \cap ... \cap A_{n,i_n}) = P(A_{1,i_1})...P(A_{n,i_n}), \qquad (*)$$

relaţie care este valabilă pentru orice alegere de atomi $A_{l,i_l}, i_l \in I_l, l = 1,...,n$.Următoarea propoziţie afirmă că reciproca este şi ea valabilă.

Propoziţia 4.5. *Cu notaţia de mai sus, dacă relaţia (*) este verificată pentru orice alegere de atomi, atunci $\sigma(\mathcal{A}_1),...,\sigma(\mathcal{A}_n)$ sunt independente.*

Demonstraţie. Considerăm variabilele aleatoare $X_l : \Omega \to I_l$, definite prin $X_l(\omega) = i$, dacă $\omega \in A_{l,i}$. Observăm că are loc egalitatea $\sigma(X_l) = \sigma(\mathcal{A}_l)$. Pe de altă parte, aplicaţia $Z = (X_1,...,X_n) : \Omega \to I$, cu valori în spaţiul produs $I = I_1 \times ... \times I_n$ are repartiţia exprimată pe puncte, $(i_1,...,i_n) \in I$, prin

$$P_Z(\{(i_1,...,i_n)\}) = P\left(Z^{-1}(i_1,...,i_n)\right) = P(A_{1,i_1} \cap ... \cap A_{n,i_n}) =$$

$$= P(A_{1,i_1})...P(A_{n,i_n}) = P_{X_1}(\{i_1\})...P_{X_n}(\{i_n\}),$$

ceea ce arată că $P_Z = P_{X_1} \bigotimes ... \bigotimes P_{X_n}$. *Propoziţia* 4.3 ne dă independenţa dorită. □

3.1. Asocierea şi disocierea independenţei. Presupunem că $(\mathcal{F}_l)_{l\in L}$ este o familie arbitrară de sub - σ -algebre ale lui $\mathcal{F}$. Presupunem că mulţimea indicilor este partiţionată în forma $L = \cup_{j\in J} L_j$ astfel că $L_j \cap L_{j'} = \emptyset$, dacă $j \neq j'$. Vom arăta că proprietatea de independenţă se pliază acestei partiţionări a mulţimii de indici prin asociere şi disociere. Pentru fiecare indice $j \in J$, notăm $\mathcal{F}^j = \sigma\left(\cup_{l\in L_j} \mathcal{F}_l\right)$.

Teorema 4.1. *(i) Dacă $\sigma-$algebrele $(\mathcal{F}_l)_{l\in L}$ sunt independente, atunci şi $\sigma-$algebrele $(\mathcal{F}^j)_{j\in J}$ sunt independente.*

(ii) Reciproc, presupunem că $\sigma-$algebrele $(\mathcal{F}^j)_{j\in J}$ sunt independente şi că, pentru fiecare indice $j \in J, \sigma-$ algebrele din familia $(\mathcal{F}_l)_{l\in L_j}$ sunt, de asemenea, independente. Atunci $\sigma-$ algebrele din familia mare, $(\mathcal{F}_l)_{l\in L}$, sunt independente în ansamblu.

Demonstraţie. Într-o primă etapă vom considera cazul în care mulţimea de indici L este o finită, iar fiecare din σ -algebrele familiei iniţiale sunt generate de partiţii. Fie atunci $(\mathcal{A}_l)_{l\in L}$ o familie de partiţii cel mult numărabile, $\mathcal{A}_l = (A_{l,i})_{i\in I_l}$, cu atomi din $\mathcal{F}$, astfel că $\mathcal{F}_l = \sigma(\mathcal{A}_l)$. Notăm $X_l : \Omega \to I_l$ aplicaţia definită prin $X_l(\omega) = i$, dacă $\omega \in A_{l,i}, i \in I_l$. Definim apoi aplicaţia cu valori în spaţiul produs

$$Z : \Omega \to I = \prod_{l\in L} I_l, \qquad Z(\omega) = (X_l(\omega))_{l\in L}$$

şi aplicaţiile cu valori în produsele parţiale

$$Z^j : \Omega \to I^j = \prod_{l\in L^j} I_l, \qquad Z^j(\omega) = (X_l(\omega))_{l\in L^j} \quad .$$

Facem identificarea $I = \prod_{j\in J} I^j$, astfel că Z se exprimă şi sub forma $Z(\omega) = (Z^j(\omega))_{j\in J}$. Vom avea atunci $\mathcal{F}_l = \sigma(X_l)$ şi $\mathcal{F}^j = \sigma(Z^j)$, iar independenţa variabilelor $X_l, l \in L$ este echivalentă cu independenţa σ -algebrelor $\mathcal{F}_l, l \in L$. La fel, independenţa variabilelor $Z^j, j \in J$ este echivalentă cu independenţa σ -algebrelor $\mathcal{F}^j, j \in J$.

(i) Independenţa familiei iniţiale de σ -algebre implică relaţiile

$$P_Z = \bigotimes_{l\in L} P_{X_l}\ , \quad P_{Z^j} = \bigotimes_{l\in L^j} P_{X_l}\ .$$

Pe de altă parte o verificare directă pe puncte arată că aceste relaţii implică egalitatea $P_Z = \bigotimes_{j\in J} P_{Z^j}$. Rezultă că variabilele $Z^j, j \in J$ sunt independente, ceea ce încheie demonstraţia acestui caz.

(ii) În această situaţie avem prin ipoteză că $P_Z = \bigotimes_{j\in J} P_{Z^j}$ şi $P_{Z^j} = \bigotimes_{l\in L^j} P_{X_l}$. Aceasta implică relaţia $P_Z = \bigotimes_{l\in L} P_{X_l}$, care exprimă independenţa dorită.

Pentru cazul general, când σ -algebrele $\mathcal{F}_l, l \in L$, nu mai sunt presupuse generate de partiţii, iar L nu se mai presupune finită, se ţine cont că relaţia de independenţă trebuie verificată cu un număr finit de evenimente. Se consideră atunci algebrele generate de acele evenimente şi se aplică rezultatul verificat mai sus. □

Următoarea propoziţie ne va fi utilă mai departe.

Propoziţia 4.6. *Fie $X_1, ..., X_n$ o familie de variabile aleatoare independente cu valori în spaţiile cel mult numărabile $E_i, i = 1, ..., n$. Să presupunem că $1 \le n_1 < ... < n_{k-1} < n_k = n$ sunt numere* *naturale care împart variabilele în k grupuri. Dacă $f_j : E_{n_{j-1}+1} \times \ldots \times E_{n_j} \to F_j, j = 1,..,k$ sunt aplicaţii ce iau valori în mulţimile cel mult numărabile $F_j, j = 1,...,k$, atunci variabilele aleatoare $Y_j = f_j\left(X_{n_{j-1}+1},...,X_{n_j}\right), j = 1,...,k$ sunt independente.*

Demonstraţie. Pentru a demonstra afirmaţia făcută vom observa mai întâi că variabila Y_j este măsurabilă faţă de $\sigma\left(X_{n_{j-1}+1},...,X_{n_j}\right)$. Apoi, din *teorema* 4.1, se deduce faptul că $\sigma-$algebrele

$$\sigma\left(X_1,...,X_{n_1}\right),...,\sigma\left(X_{n_{k-1}+1},...,X_{n_k}\right)$$

sunt independente, ceea ce implică afirmaţia făcută. □

Extrageri de bile colorate.

Vom analiza o urnă cu m bile colorate. Numărul de culori este k, iar numărul de bile colorate în culoarea c_i este m_i. Prin urmare este satisfăcută relaţia $\sum_{i=1}^{k} m_i = m$. Vom presupune că se fac n extrageri repetate cu întoarcere, iar ca rezultat al fiecărei extrageri, este notată numai culoarea bilei extrase. Modelul ce corespunde acestui experiment este descris în felul următor. Notăm M mulţimea ce corespunde bilelor din urnă. Este o mulţime finită cu m elemente. Mulţimea bilelor de culoarea c_i va fi notată M_i; este o mulţime cu m_i elemente. Rezultă că mulţimile $M_1,...,M_k$ formează o partiţie a lui M. Spaţiul probabilizat este dat de mulţimea

$$\Omega = M^n = \left\{(x_1,...,x_n) \,/\, x_l \in M, l = 1,...,n\right\},$$

în care punctul generic va fi notat $\omega = (x_1,...,x_n)$. Aşa cum ştim, pentru extrageri cu întoarcere, probabilitatea corespunzătoare este probabilitatea şanselor egale. Deci pe mulţimea Ω se pune probabilitatea ce corespunde şanselor egale: $P\left(\{\omega\}\right) = \frac{1}{m^n}$, pentru orice $\omega \in \Omega$.

Ceea ce ne interesează acum este să punem în evidenţă partiţiile ce corespund fiecărei extrageri. Anume, la o extragere se pot manifesta k evenimente distincte. Să zicem că este vorba despre extragerea l; atunci pentru $l = 1,...,k$ avem evenimentul descris prin propoziţia „la extragerea l a ieşit culoarea i ". Acest eveniment este modelat de mulţimea $A_{l,i} = \left\{(x_1,...,x_n) \in \Omega / \, x_l \in M_i\right\}$. Este clar că familia $\mathcal{A}_l = \left(A_{l,i}\right)_{i=1,...k}$ este o partiţie finită. Avem deci n partiţii distincte, fiecare cu câte k atomi. Vom verifica acum că aceste partiţii sunt independente. Pentru a număra punctele dintr-o mulţime $A_{l,i}$ fixată, vom observa că ea este în bijecţie cu mulţimea $M_i \times M^{n-1}$. Rezultă că avem $P\left(A_{l,i}\right) = \frac{m_i m^{n-1}}{m^n} = \frac{m_i}{m}$. Pe de altă parte, fiind date mulţimile $A_{1,i_1},...,A_{n,i_n}$, mulţimea $\cap_{l=1}^{n} A_{l,i_l}$ este în bijecţie cu produsul $M_{i_1} \times \ldots \times M_{i_n}$ şi prin urmare va avea cardinalul $m_{i_1}...m_{i_n}$. Acest calcul conduce la relaţia de independenţă a partiţiilor:

$$P\left(\cap_{l=1}^{n} A_{l,i_l}\right) = \frac{m_{i_1}...m_{i_n}}{m^n} = P\left(A_{1,i_1}\right)...P\left(A_{n,i_n}\right).$$

Propoziţia 4.5 este atunci aplicabilă şi rezultă independenţa algebrelor generate de partiţiile $\mathcal{A}_l, l = 1,...,n$. Aceasta revine la independenţa oricăror evenimente ce se referă la extrageri distincte. Cu alte cuvinte, putem spune că cele n extrageri sunt independente într-un sens mai larg.

Aplicaţie numerică.

Să presupunem că avem $k = 5$, iar culorile sunt: c_1 -alb, c_2 -negru, c_3 -galben, c_4 -roşu, c_5 -albastru. Presupunem că se fac şase extrageri şi doi jucători mizează pe rezultatele extragerilor alternativ. Primul mizează pe culorile roşu şi albastru la extragerile $1, 3, 5$ iar al doilea pe culorile alb şi roşu la extragerile $2, 4, 6$. Ne propunem să determinăm: (i) care este probabilitatea ca primul jucător să fi câştigat la cel puţin o extragere la care a mizat, (ii) care este probabilitatea ca, în acelaşi timp, cel de al doilea jucător să fi câştigat şi el la cel puţin una din extragerile la care a mizat.

Souţie. (i) Sunt şase extrageri, deci $n = 6$. Notăm

$$B_l = A_{l,4} \cup A_{l,5},$$

evenimentul care exprimă faptul că la extragerea l au ieşit culorile pe care a mizat primul jucător. Avem $P(B_l) = P(A_{l,4}) + P(A_{l,5}) = \frac{m_4+m_5}{m}$. Evenimentul care exprimă faptul că primul jucător a câştigat cel puţin o dată este $B = B_1 \cup B_3 \cup B_5$ şi, ţinând cont că evenimentele care-l compun sunt independente, el are probabilitatea

$$P(B) = P(B_1) + P(B_3) + P(B_5) - P(B_1 \cap B_3) - P(B_1 \cap B_5) -$$

$$-P(B_3 \cap B_5) + P(B_1 \cap B_3 \cap B_5) = 3\frac{m_4 + m_5}{m} - 3\left(\frac{m_4 + m_5}{m}\right)^2 +$$

$$+\left(\frac{m_4 + m_5}{m}\right)^3 = \frac{m_4 + m_5}{m^3}\left(3m^2 - 3m(m_4 + m_5) + (m_4 + m_5)^2\right).$$

(ii) Notăm cu D evenimentul ca cel de al doilea jucător să câştige cel puţin o dată. Probabilitatea sa se calculează similar şi este

$$P(D) = \frac{m_1 + m_4}{m^3}\left(3m^2 - 3m(m_1 + m_4) + (m_1 + m_4)^2\right).$$

Dar evenimentele B şi D sunt independente pentru că $B \in \sigma(\mathcal{A}_1, \mathcal{A}_3, \mathcal{A}_5)$ şi $D \in \sigma(\mathcal{A}_2, \mathcal{A}_4, \mathcal{A}_6)$. Prin urmare avem

$$P(B \cap D) = P(B) P(D) = \frac{(m_4 + m_5)(m_1 + m_4)}{m^6} \times$$

$$\times \left(3m^2 - 3m(m_4 + m_5) + (m_4 + m_5)^2\right) \times$$

$$\times \left(3m^2 - 3m(m_1 + m_4) + (m_1 + m_4)^2\right),$$

ceea ce dă rezultatul căutat. □

Modelul lui Lebesgue.*

Pentru a se construi un şir infinit de variabile aleatoare independente cu repartiţie dată se procedează ca în cazul unui număr finit de variabile, cu singura deosebire că trebuie făcut un produs infinit de spaţii şi de măsuri. Nu vom face acest lucru aici, deşi vom avea nevoie de aşa ceva în paragraful despre repartiţia geometrică din capitolul ce urmează. În schimb prezentăm acum construcţia dată de Lebesgue pentru un şir de variabile aleatoare independente, repartizate Bernoulli cu $p = \frac{1}{2}$.

Pe intervalul $[0,1)$, considerăm mulţimile $A_{n,k} = [\frac{k}{2^n}, \frac{k+1}{2^n}), n \in \mathbf{N}^*, k = 0, ..., 2^n - 1$. Pentru fiecare n fixat, familia $\mathcal{A}_n = (A_{n,k})_{k=0}^{2^n-1}$ formează o partiţie a intervalului $[0,1)$. Aceste partiţii se rafinează una pe alta când n creşte. Mai precis, mulţimile din partiţia de rang n se împart în două mulţimi din partiţia de rang $n+1$, în felul următor:

$$A_{n,k} = A_{n+1,2k} \cup A_{n+1,2k+1}.$$

Pentru fiecare partiţie luăm intervalele de ordin impar şi funcţia lor caracteristică o notăm

$$X_n = 1_{\cup_{l=0}^{2^n-1} A_{n,2l+1}} = \sum_{l=0}^{2^n-1} 1_{A_{n,2l+1}}.$$

Este clar că, notând cu μ măsura Lebesgue, au loc relaţiile $\mu(X_n = 0) = \mu(X_n = 1) = \frac{1}{2}$. Cu alte cuvinte, toate au repartiţia $\frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$.

Acum se pune problema de a utiliza spaţiul $[0,1)$ cu măsura Lebesgue ca un spaţiu probabilizat. Până acum am evitat noţiunile mai avansate de teoria măsurii, însă aici trebuie să apelăm la extensia măsurii lungime dată de Lebesgue. Mai precis, vom utiliza faptul că există o unică măsură σ -aditivă, μ, definită pe $\mathcal{B}([0,1))$ care coincide cu măsura lungimii pentru intervale. Astfel $([0,1), \mathcal{B}([0,1)), \mu)$ poate fi considerat ca un spaţiu probabilizat.

Propoziţia 4.7. *Variabilele $X_1, X_2, ...$, construite mai sus pe spaţiul probabilizat $([0,1), \mathcal{B}([0,1)), \mu)$, sunt independente.*

Demonstraţie. Pentru demonstraţie, vom arăta mai întâi următoarea egalitate

$$\{X_1 = a_1, ..., X_n = a_n\} = [a, a + \frac{1}{2^n}), \quad (*)$$

unde $a_1, ..., a_n \in \{0,1\}$ şi $a = \sum_{i=1}^n \frac{a_i}{2^i}$. Observăm că în cazul $n = 1$, verificarea egalităţii revine la verificarea egalităţilor $\{X_1 = 0\} = [0, \frac{1}{2})$, pentru $a_1 = 0$ sau $\{X_1 = 1\} = [\frac{1}{2}, 1)$, pentru $a_1 = 1$, egalităţi evident verificate. Prin inducţie, presupunând că este verificată egalitatea

$$\{X_1 = a_1, ..., X_{n-1} = a_{n-1}\} = [a', a' + \frac{1}{2^{n-1}}),$$

unde $a' = \sum_{i=1}^{n-1} \frac{a_i}{2^i}$, dorim să deducem egalitatea (*). Pentru aceasta vom scrie numărul anterior sub forma $a' = \frac{k}{2^{n-1}}$, unde $k = \sum_{i=1}^{n-1} a_i 2^{n-1-i}$. Deoarece $k \leq \sum_{i=1}^{n-1} 2^{n-1-i} = 2^{n-1} - 1$, rezultă că putem scrie $[a', a' + \frac{1}{2^{n-1}}) = A_{n-1,k} = A_{n,2k} \cup A_{n,2k+1}$. Pe de altă parte, avem

$$\{X_n = 1\} = \cup_{l=0}^{2^n-1} A_{n,2l+1}, \quad \{X_n = 0\} = \cup_{l=0}^{2^n-1} A_{n,2l},$$

ceea ce ne permite să scriem

$$\{X_1 = a_1, ..., X_n = a_n\} = [a', a' + \frac{1}{2^{n-1}}) \cap \{X_n = a_n\} =$$
$$= (A_{n,2k} \cup A_{n,2k+1}) \cap \left(\cup_{l=0}^{2^n-1} A_{n,2l+1}\right) = A_{n,2k+1} \ ,$$

dacă $a_n = 1$, iar în cazul $a_n = 0$,

$$\{X_1 = a_1, ..., X_n = a_n\} = A_{n,2k}.$$

Dar în cazul $a_n = 1$ avem $A_{n,2k+1} = [\frac{2k+1}{2^n}, \frac{2k+2}{2^n}) = [a, a + \frac{1}{2^n})$, pentru că $\frac{2k+1}{2^n} = \frac{k}{2^{n-1}} + \frac{1}{2^n} = a' + \frac{1}{2^n} = \sum_{i=1}^{n-1} \frac{a_i}{2^i} + \frac{1}{2^n} = a$. La fel, în cazul $a_n = 0$ avem $A_{n,2k} = [a, a + \frac{1}{2^n})$. Cu aceasta am terminat stabilirea egalităţii (*).

Pentru a stabili independenţa variabilelor noastre este suficient să scriem

$$\mu\left(X_1 = a_1, ..., X_n = a_n\right) = \mu\left([a, a + \frac{1}{2^n})\right) = \frac{1}{2^n} =$$
$$= \mu\left(X_1 = a_1\right) ... \mu\left(X_n = a_n\right).$$

Cu aceasta am încheiat demonstraţia. □

Merită să notăm că de fapt modelul lui Lebesgue nu diferă în mod esenţial de un model de tip produs infinit. Pentru a preciza această afirmaţie, să notăm cu $D = \left\{x = \frac{k}{2^n} / n \in \mathbf{N}^*, k = 0, ...2^n - 1\right\}$, mulţimea numerelor diadice din intervalul $[0, 1)$. Scrierea diadică a unui număr $x \in [0, 1) \backslash D$, sub forma $x = \sum_{i=1}^{\infty} \frac{a_i}{2^n}$, cu $a_1, a_2, ... \in \{0, 1\}$ este unică. (Pentru un număr diadic există două scrieri de acest fel, amândouă cu un şir de numărători constanţi de la un rang încolo.) Cum mulţimea numerelor diadice este numărabilă, ea este neglijabilă pentru măsura Lebesgue. Rezultă că $([0, 1) \backslash D, \mathcal{B}([0, 1) \backslash D), \mu)$ este tot un spaţiu probabilizat, iar şirul de variabile construite mai sus (prin restricţie la $[0, 1) \backslash D$) rămân independente şi cu aceeaşi repartiţie. Definind $\Omega = \{0, 1\}^{\mathbf{N}^*}$ şi aplicaţia $\Phi : [0, 1) \backslash D \to \Omega$ prin $\Phi(x) = (a_1, a_2, ...)$, observăm că imaginea lui Φ, mulţimea $\Phi([0, 1) \backslash D)$ diferă de Ω doar prin mulţimea şirurilor constante de la un rang încolo:

$$\Omega \backslash \Phi([0, 1) \backslash D) = \{(a_1, a_2, ...) \in \Omega / \exists i \in \mathbf{N}^*, a_i = a_{i+l}, \forall l \in \mathbf{N}\},$$

care este o mulţime numărabilă. Aplicaţia Φ este biunivocă. Notăm cu $P = \mu \circ \Phi^{-1}$ repartiţia lui Φ, care este o măsură pe $(\Omega, \mathcal{F})$, unde $\mathcal{F}$ este σ -algebra produs pe Ω. Dacă notăm cu $Y_n : \Omega \to \{0, 1\}$ aplicaţiile de proiecţie pe componente, $Y_n(a_1, a_2, ...) = a_n$, putem scrie $\mathcal{F} = \sigma(Y_n / n \in \mathbf{N}^*)$. Se poate verifica că are loc egalitatea $Y_n \circ \Phi = X_n$, deci variabilele $Y_1, Y_2, ...$ sunt independente şi cu aceeaşi repartiţie ca variabilele $X_1, X_2, ...$ Măsura P se dovedeşte a fi tocmai măsura produs $\prod_{i \in \mathbf{N}^*} \lambda_i$, unde pe fiecare componentă a produsului se ia aceeaşi măsură: $\lambda_i = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$.

Construcţia lui Lebesgue poate fi generalizată pentru a produce o teoremă de existenţă pentru lanţuri Markov (vezi teorema 8.1, pag. 112

în cartea lui Billingsley citată la bibliografie). În particular, se obţin în acest fel un şir de variabile independente, toate repartizate Bernoulli cu un parametru $p \in (0, 1)$ arbitrar.

4. Medie şi dispersie

Presupunem că $(\Omega, \mathcal{F}, P)$ este un spaţiu probabilizat. În această secţiune vom discuta despre variabile aleatoare reale de tipul $X : \Omega \to E$, cu valori într-o mulţime cel mult numărabilă $E \subset \mathbf{R}$. O astfel de variabilă se scrie şi sub forma

$$X = \sum_{x \in E} x 1_{X^{-1}(x)} = \sum_{x \in E'} x 1_{X^{-1}(x)},$$

unde $E' = X(\Omega)$.

Pentru variabilele cu valori reale se pot face o serie de operaţii în plus faţă de variabilele generale. De exemplu se pot efectua operaţiile aritmetice obişnuite. Dacă avem două variabile $X : \Omega \to E, Y : \Omega \to F$, unde E şi F sunt mulţimi cel mult numărabile de numere reale, iar $a, b \in \mathbf{R}$, atunci putem construi aplicaţia sumă

$$Z = aX + bY : \Omega \to G,$$

definită prin $Z(\omega) = aX(\omega) + bY(\omega)$, cu valori în mulţimea cel mult numărabilă de numere reale $G = \{ax + by / x \in E, y \in F\}$. În mod similar se definesc produsul XY şi, în cazul în care $0 \notin F$, câtul X/Y, care sunt tot cu mulţimea valorilor numărabilă.

O operaţie importantă, specifică cazului real, este calcularea mediei.

Definiţia 4.3. *Fie $X : \Omega \to E$, o variabilă cu valori într-o mulţime cel mult numărabilă astfel ca $X(\Omega) = E \subset \mathbf{R}$. Vom spune că X admite medie, dacă familia de numere $(xP(X^{-1}(x)))_{x \in E}$ este absolut sumabilă. În caz că este îndeplinită această condiţie, se spune că suma*

$$EX = \sum_{x \in E} xP\left(X^{-1}(x)\right)$$

reprezintă media *variabilei X.*

Media unei variabile aleatoare mai este numită şi *valoarea medie* sau, de asemenea, acelei variabile, datorită semnificaţiei pe care o are în cazul diverselor modele probabiliste concrete. Termenul în limba engleză este „expectation", iar în limba franceză „espérance". Semnul E este preluat din literatura internaţională. În literatura română este utilizat de asemenea semnul M, ce reprezintă iniţiala cuvântului medie. Noţiunea aceasta este de fapt integrala din teoria măsurii. O variabilă aleatoare cu valori reale este o funcţie reală măsurabilă şi media nu este altceva decât integrala acestei funcţii. Notaţia utilizată în teoria măsurii este variată:

$$EX = \int X dP = \int_\Omega X(\omega) P(d\omega) = \int_\Omega X(\omega) dP(\omega).$$

Faptul că o variabilă admite medie este tot una cu faptul că ea este *integrabilă.*

Dacă variabila X are repartiţia descrisă de un tablou

$$P_X = \begin{pmatrix} a_1 & a_2 & ... & a_n \\ p_1 & p_2 & & p_n \end{pmatrix},$$

unde $a_1, ..., a_n \in \mathbf{R}$, sunt toate valorile pe care le ia, atunci media apare scrisă ca produsul scalar al celor doi vectori linie $EX = a_1p_1 + ... + a_np_n$.

Dacă variabila se reduce la o funcţie caracteristică, $1_A, A \in \mathcal{F}$, atunci avem $E1_A = P(A)$. Observăm că orice variabilă mărginită admite medie. Dacă $|X| \leq K$, atunci $EX \leq K$.

Exemplul 1.

Presupunem că se aruncă o monedă de 3 ori. Notăm cu X numărul de steme obţinute în urma experimentului. Aceasta este o variabilă aleatoare, iar media sa ne dă numărul mediu de steme care aşteptăm să apară. Prin repetarea experimentului de un număr mare de ori se poate observa că aceasta este media numerelor de steme obţinute la fiecare experiment. (Vezi şi legea numerelor mari de la sfârşitul secţiunii.)

Modelăm experimentul pe spaţiul $\Omega = \{0, 1\}^3$ cu probabilitatea şanselor egale $P(\{(x_1, x_2, x_3)\}) = \frac{1}{8}$, pentru orice punct $(x_1, x_2, x_3) \in \Omega$. Dacă stema este consemnată prin 0, iar 1 corespunde la faţa cu cifra, pentru $(x_1, x_2, x_3) \in \Omega$, variabila X se exprimă prin $X = 3 - x_1 - x_2 - x_3$. Avem $P(X = 0) = \frac{1}{8}, P(X = 1) = \frac{3}{8}, P(X = 2) = \frac{3}{8}, P(X = 3) = \frac{1}{8}$. Rezultă că

$$EX = \frac{3}{8} + 2\frac{3}{8} + 3\frac{1}{8} = \frac{3}{2}. \square$$

Exemplul 2.

Un model simplist privind fiabilitatea unui produs este următorul. Presupunem că un eşantion de 1000 de becurile produse de o fabrică se testează fiind puse să funcţioneze neîntrerupt în condiţii de uşoară suprasarcină, obţinându-se următoarele rezultate: 10 din becuri se defectează în prima săptămână de funcţionare, 21 cedează în cursul celei de a doua săptămâni, 52 în cursul celei de a treia, 90 în cursul celei de a patra săptămâni, 153 în cea de a cincea, 207 în cea de a şasea, 252 în a şaptea, 147 în a opta, 49 în a noua, 19 în a zecea.

Modelul probabilist al fiabilităţii becurilor este deci constituit în acest caz de spaţiul $\Omega = \{1, ..., 1000\}$, ce reprezintă becurile supuse experimentului, presupuse numerotate. Probabilitatea care se introduce pe acest spaţiu este, bineînţeles, $P(i) = \frac{1}{1000}$. Se defineşte o variabilă aleatoare $T : \Omega \to \mathbf{N}$, punând $T(i)$ să fie numărul de săptămâni întregi de funcţionare a becului cu numărul i. Astfel mulţimea $A_i = T^{-1}(i)$ este mulţimea becurilor ce s-au defectat în cursul săptămânii $i + 1$. Variabila T reprezintă timpul de funcţionare al unui bec, măsurat în săptămâni şi rotunjit la partea întreagă. Avem următoarele probabilităţi

$$P(A_0) = \frac{10}{1000}, P(A_1) = \frac{21}{1000}, P(A_2) = \frac{52}{1000}, P(A_3) = \frac{90}{1000},$$
$$P(A_4) = \frac{153}{1000}, P(A_5) = \frac{207}{1000}, P(A_6) = \frac{252}{1000},$$
$$P(A_7) = \frac{147}{1000}, P(A_8) = \frac{49}{1000}, P(A_9) = \frac{19}{1000}.$$

Speranţa de viaţă sau durata medie de funcţionare a unui bec aste exprimată de media variabilei T :

$$ET = 0\tfrac{10}{1000} + \tfrac{21}{1000} + 2\tfrac{52}{1000} + 3\tfrac{90}{1000} + 4\tfrac{153}{1000} + 5\tfrac{207}{1000} + 6\tfrac{252}{1000} + \\ +7\tfrac{147}{1000} + 8\tfrac{49}{1000} + 9\tfrac{19}{1000} = 5,146.$$

Rezultă că ne aşteptăm ca un bec arbitrar să funcţioneze 5 săptămâni, în condiţii de uşoară suprasolicitare. □

Exemplul 3.

Un jucător la curse de călărie pariază pe caii A,B,C, respectiv sumele $30, 10, 15$ euro. El apreciază că şansele de a ieşi învingători sunt următoarele: 70% pentru A, 10% pentru B şi 20% pentru C. Înainte de începerea cursei, jucătorul ştia că suma pariată pe A va produce un câştig cu coeficientul 2 în cazul că acesta va câştiga, suma pariată pe B va produce un câştig cu coeficientul $7,4$ în cazul că va ieşi învingător, iar suma pariată pe C va aduce un câştig cu coeficientul $4,5$ în cazul că acest cal va câştiga cursa.

Modelul probabilist pentru acest jucător este $\Omega = \{A, B, C\}$, corespunzător celor trei cai ce sunt consideraţi ca posibili câştigători. Probabilitatea este definită prin $P(A) = 0,7; P(B) = 0,1; P(C) = 0,2$. Notăm cu X variabila aleatoare ce reprezintă câştigurile posibile pentru jucător: $X(A) = 30 \times 2 = 60, X(B) = 10 \times 7,4 = 74, X(C) = 15 \times 4,5 = 67,5$. Speranţa de câştig, sau câştigul mediu aşteptat de jucător este exprimată de media variabilei:

$$EX = 60 \times 0,7 + 74 \times 0,1 + 67,5 \times 0,2 = 62,9.$$

Deci jucătorul joacă suma de 55 euro şi, pe baza estimărilor sale, se aşteaptă să câştige $62,9$. □

Vom examina în continuare câteva din principalele proprietăţi ale valorii medii. Pentru aceasta este convenabil să schimbăm un pic notaţia în exprimarea variabilelor presupunând că ele au forma

$$X = \sum_{i \in I} a_i 1_{A_i}, \quad (*)$$

unde $(A_i)_{i \in I}$ este o partiţie cel mult numărabilă a lui Ω cu elemente din $\mathcal{F}$, iar $(a_i)_{i \in I}$ este o familie de numere reale. Desigur că scrierea anterioară

$$X = \sum_{x \in E} x 1_{X^{-1}(x)},$$

unde $E = X(\Omega)$, este de acelaşi tip, pentru că $\mathcal{A} = \left(X^{-1}(e)\right)_{e\in E}$ este o partiţie. Dar în exprimarea (*) nu cerem ca $A_i = X^{-1}(a_i)$, asfel că pot exista valori $x \in \mathbf{R}$ pentru care mulţimea $\{i \in I/a_i = x\}$ să conţină mai mulţi indici diferiţi. Deci partiţia $(A_i)_{i\in I}$ este mai fină decât $\mathcal{A}$. Scrierea unei variabile sub o astfel de formă nu este unică, în general. Pe de altă parte, ţinând cont şi de convenţia făcută cum că atomii partiţiilor trebuie să fie nevizi, rezultă că mulţimea valorilor are descrierea

$$E = X(\Omega) = \{a_i/i \in I\}.$$

În acest caz media variabilei se calculează după formula dată de lema care urmează.

Lema 4.9. *Presupunem că variabila X se exprimă sub forma (*) de mai sus. Atunci X admite medie dacă şi numai dacă familia $(a_i P(A_i))_{i\in I}$ este absolut sumabilă şi are loc formula*

$$EX = \sum_{i\in I} a_i P(A_i).$$

Demonstraţie. Pentru fiecare valoare a variabilei $x \in E$ notăm $I_x = \{i \in I/a = x\}$, astfel că avem $X^{-1}(x) = \bigcup_{i\in I_x} A_i$ şi obţinem o partiţie a mulţimii de indici: $I = \bigcup_{x\in E} I_x$. Rezultă că putem face calculul următor fără probleme în cazul în care I este finită:

$$\begin{aligned} EX &= \sum_{x\in E} xP\left(X^{-1}(x)\right) = \sum_{x\in E} x \sum_{i\in I_x} P(A_i) = \sum_{x\in E} a_i \sum_{i\in I_x} P(A_i) = \\ &= \sum_{i\in I} a_i P(A_i). \end{aligned}$$

Afirmaţia privind absolut sumabilitatea şi justificarea formulei în cazul general se deduce recurgând la *lema* 4.3. □

Înşirăm acum principalele proprietăţi ale mediei.

1. Dacă X este o variabilă ce admite medie şi $X \geq 0$, atunci $EX \geq 0$.

2. Variabila $X : \Omega \to E \subset \mathbf{R}$ admite medie, dacă şi numai dacă $|X|$ admite medie. În plus are loc inegalitatea $EX \leq E|X|$.

3. Dacă variabila X admite medie, atunci, pentru orice $a \in \mathbf{R}$, variabila aX admite de asemenea medie şi are loc relaţia $E(aX) = aEX$.

4. Dacă X, Y sunt două variabile care admit medie, atunci $X + Y$ este o variabilă care admite de asemenea medie şi are loc formula $E(X + Y) = EX + EY$.

Pentru verificarea acestei proprietăţi pornim cu cele două variabile exprimate cu partiţii sub forma $X = \sum_{i\in I} a_i 1_{A_i}, Y = \sum_{j\in J} b_j 1_{B_j}$. Punând $C_{ij} = A_i \cap B_j$, şi

$$L = \{(i,j) \in I \times J/C_{ij} \neq \emptyset\},$$

am obţinut o nouă partiţie, $(C_{ij})_{(i,j)\in L}$, în raport cu care putem scrie

$$A_{i_0} = \bigcup_{j,(i_0,j)\in L} C_{i_0 j}, \;\; B_{j_0} = \bigcup_{i,(i,j_0)\in L} C_{i j_0},$$

pentru orice $i_0 \in I$ şi orice $j_0 \in J$. (Este vorba de fapt de partiţia asociată $\sigma-$algebrei generate de partiţiile $(A_i)_{i\in I}$ şi $(B_j)_{j\in J}$.) Aşa că putem scrie

$$X = \sum_{(i,j)\in L} a_i 1_{Cij}, \;\; Y = \sum_{(i,j)\in L} b_j 1_{Cij},$$
$$X + Y = \sum_{(i,j)\in L} (a_i + b_j)\, 1_{Cij}.$$

Prin aplicarea lemei anterioare se obţine imediat

$$E(X+Y) = \sum_{(i,j)\in L} (a_i + b_j)\, P(C_{ij}) =$$
$$= \sum_{(i,j)\in L} a_i P(C_{ij}) + \sum_{(i,j)\in L} b_j P(C_{ij}) = EX + EY. \; \square$$

5. Rezultă prin inducţie că dacă $X_1, ..., X_n$ sunt variabile aleatoare ce admit medie, atunci orice sumă de tipul $a_1X_1 + ... + a_nX_n$, cu $a_1, ..., a_n \in \mathbf{R}$, admite de asemenea medie şi are loc formula

$$E(a_1X_1 + ... + a_nX_n) = a_1EX_1 + ... + a_nEX_n.$$

6. Dacă X, Z sunt două variabile aleatoare astfel că Z este integrabilă şi are loc inegalitatea $|X| \le Z$, atunci şi X este integrabilă şi $EX \le E|X| \le EZ$.

Pentru verificare, presupunem mai întâi că $X \ge 0$. Atunci pornim cu exprimarea lui X cu o partiţie sub forma

$$X = \sum_{i\in I} a_i 1_{A_i}.$$

Notăm $Y = Z - X$, care este o variabilă nenegativă, conform ipotezei. Ea se poate exprima cu o partiţie sub forma $Y = \sum_{j\in J} b_j 1_{B_j}$. Vom avea $X + Y = Z$. Reluăm notaţia de la punctul 4. de mai sus pentru partiţia $(C_{ij})_{(i,j)\in L}$, în raport cu care scriem

$$X = \sum_{(i,j)\in L} a_i 1_{Cij}, \;\; Y = \sum_{(i,j)\in L} b_j 1_{Cij}, \;\; Z = \sum_{(i,j)\in L} (a_i + b_j)\, 1_{Cij}.$$

Numerele a_i şi b_j sunt toate nenegative, deoarece variabilele respective sunt nenegative. Prin urmare, avem $a_i P(C_{ij}) \le (a_i + b_j)\, P(C_{ij})$, pentru orice $(i,j) \in L$. Rezultă că X admite medie şi $EX \le EZ$.

Pentru a trata cazul general, observăm că se poate utiliza cazul anterior pentru $|X|$, ceea ce conduce la concluzia că $|X|$ admite medie şi $E|X| \le EZ$. Ţinând cont de punctul 2. de mai sus se obţine concluzia dorită. $\square$

Alte proprietăţi legate de medie le prezentăm mai jos sub forma unor enunţuri de leme.

Fie X o variabilă cu proprietatea că X^2 are medie. Atunci media $E\left(X^2\right)$ este numită *media pătratică* sau *momentul de ordinul doi* al variabilei.

Lema 4.10. *(Inegalitatea lui Schwartz) Dacă X, Y sunt variabile care admit momente de ordinul doi, atunci produsul XY admite medie şi are loc inegalitatea*

$$E\left(XY\right) \leq \sqrt{E\left(X^2\right)}\sqrt{E\left(Y^2\right)}$$

Demonstraţie. Vom porni cu observaţia că variabila XY admite medie, datorită inegalităţii $|XY| \leq \frac{1}{2}\left(X^2 + Y^2\right)$ şi a punctului 6. de mai sus. Rezultă că $X + \lambda Y$ admite moment de ordinul doi pentru orice număr real $\lambda \in \mathbf{R}$. Rezultă

$$E\left(X + \lambda Y\right)^2 \geq 0,$$

inegalitate care se scrie

$$EX^2 + 2\lambda E\left(XY\right) + \lambda^2 EY^2 \geq 0.$$

Privim acum expresia din stânga ca pe un polinom de ordinul doi în λ cu coeficienţi $EY^2, EX^2, 2E\left(XY\right)$. Punând condiţia ca discriminantul să fie negativ -condiţia care asigură semnul constant al polinomului -avem

$$\left[E\left(XY\right)\right]^2 - \left(EX^2\right)\left(EY\right)^2 \leq 0,$$

care este tocmai inegalitatea dorită. □

Inegalitatea aceasta poate fi privită ca o generalizare a inegalităţii *Cauchy -Buniakovschi*:

$$\sum_{i\in I} a_i b_i \leq \sqrt{\sum_{i\in I} a_i^2}\sqrt{\sum_{i\in I} b_i^2},$$

ce am întâlnit-o pentru familii finite $\left(a_i\right)_{i\in I}, \left(b_i\right)_{i\in I}$ de numere reale.

Luând $Y = 1$ în lemă se deduce că dacă o variabilă admite medie pătratică, atunci admite medie şi are loc inegalitatea $EX \leq \sqrt{E\left(X^2\right)}$.

Lema 4.11. *Dacă X şi Y sunt două variabile independente, astfel că atât X cât şi Y admit medie, atunci produsul XY admite medie şi ea se calculează după formula $E\left(XY\right) = \left(EX\right)\left(EY\right)$.*

Demonstraţie. Pentru demonstraţie pornim cu notaţiile

$$X = \sum_{i\in I} a_i 1_{A_i}, Y = \sum_{j\in J} b_j 1_{B_j}$$

şi cu partiţia $\left(C_{ij}\right)_{(i,j)\in L}$ introduse la punctul 4., însă vom alege partiţiile să fie astfel ca $A_i = X^{-1}\left(a_i\right)$, pentru orice $i \in I$ şi $B_j = Y^{-1}\left(b_j\right)$, pentru orice $j \in J$. Atunci produsul variabilelor se exprimă sub forma

$$XY = \sum_{(i,j)\in L} a_i b_j 1_{C_{ij}}.$$

Să verificăm absolut integrabilitatea familiei de numere $(a_i b_j P(C_{ij}))_{(i,j)\in L}$. Fie o mulţime finită $M \subset L$. Putem găsi două mulţimi finite $M_1 \subset I$ şi $M_2 \subset J$ astfel încât $M \subset M_1 \times M_2$. Atunci putem scrie

$$\sum_{(i,j)\in M} |a_i b_j| P(C_{ij}) \leq \sum_{(i,j)\in M_1\times M_2} |a_i b_j| P(C_{ij}) =$$

$$= \sum_{i\in M_1}\sum_{j\in M_2} |a_i b_j| P(C_{ij}) = \sum_{i\in M_1}\sum_{j\in M_2} |a_i| |b_j| P(A_i \cap B_j) =$$

$$= \sum_{i\in M_1}\sum_{j\in M_2} |a_i| |b_j| P(A_i) P(B_j) =$$

$$= \left(\sum_{i\in M_1} |a_i| P(A_i)\right)\left(\sum_{j\in M_2} |b_j| P(B_j)\right) \leq (E|X|)(E|Y|).$$

Am utilizat mai sus relaţia $P(C_{ij}) = P(A_i \cap B_j) = P(A_i) P(B_j)$, care exprimă independenţa variabilelor. Rezultă că XY admite medie şi media sa este calculată prin:

$$E(XY) = \sum_{(i,j)\in L} a_i b_j P(C_{ij}) = \sum_{(i,j)\in I\times J} a_i b_j P(C_{ij}) =$$

$$= \left(\sum_{i\in I} a_i P(A_i)\right)\left(\sum_{j\in J} b_j P(B_j)\right) = (EX)(EY).$$

Acest ultim calcul este evident în cazul în care mulţimile valorilor sunt finite, adică I, J finite. În cazul general trebuie făcută o trecere la limită, pe care o lăsăm cititorului. □

Integrala discretă ca medie

Fie E o mulţime cel mult numărabilă pe care avem o măsură de probabilitate $\mu = \sum_{e\in E} \mu_e \delta_e$. Tripletul $(E, \mathcal{P}(E), \mu)$ îl privim ca pe un spaţiu probabilizat. Orice funcţie reală $f : E \to \mathbf{R}$ poate fi considerată drept variabilă aleatoare. Notând imaginea sa prin $F = f(E)$ putem exprima media sa ca o integrală

$$\begin{aligned} Ef &= \sum_{a\in F} a\mu\left(f^{-1}(a)\right) = \sum_{a\in F} a \sum_{e\in f^{-1}(a)} \mu_e = \sum_{a\in F} f(e) \sum_{e\in f^{-1}(a)} \mu_e = \\ &= \sum_{e\in E} f(e)\mu_e = \int_E f d\mu. \end{aligned}$$

Putem aplica rezultatele anterioare în acest cadru. Spre exemplu, dacă avem două funcţii, f, g, astfel că f^2 şi g^2 sunt integrabile, atunci deducem că produsul este integrabil şi putem scrie inegalitatea lui Schwartz,

$$\int fgd\mu \leq \sqrt{\int f^2 d\mu}\sqrt{\int g^2 d\mu}.$$

În cazul în care μ este repartiţia unei variabile aleatoare putem exprima media variabilei în termeni de integrală în raport cu μ. Mai general, este adevărat următorul rezultat.

Lema 4.12. *Fie $X : \Omega \to E$ o variabilă aleatoare cu valori în mulţimea cel mult numărabilă E, iar $f : E \to R$ o funcţie arbitrară. Atunci variabila cu valori reale $Y = f(X)$ admite medie dacă şi numai dacă f este integrabilă în raport cu repartiţia P_X. În plus, dacă este îndeplinită una din aceste condiţii, atunci are loc relaţia*

$$EY = \int f dP_X.$$

DEMONSTRAŢIE. Vom presupune că $E = X(\Omega)$ şi notăm $F = f(E)$, astfel că vom avea $Y(\Omega) = F$. Funcţia f poate fi scrisă sub forma $f = \sum_{e\in E} f(e) 1_{\{e\}}$, astfel că variabila Y se scrie

$$Y = f \circ X = \sum_{e\in E} f(e) 1_{\{e\}} \circ X = \sum_{e\in E} f(e) 1_{X^{-1}(e)},$$

unde am utilizat faptul că $1_{\{e\}} \circ X = 1_{X^{-1}(e)}$. Suntem în măsură să aplicăm *lema* 4.9 variabilei Y exprimate prin partiţia asociată lui X. Integrabilitatea lui Y este acelaşi lucru cu sumabilitatea absolută a seriei $\sum_{e\in E} f(e) P(X^{-1}(e))$ şi

$$EY = \sum_{e\in E} f(e) P\left(X^{-1}(e)\right).$$

Dar itegrabilitatea lui f revine la acelaşi lucru:

$$\int f dP_X = \sum_{e\in E} f(e) P_X(\{e\}) = \sum_{e\in E} f(e) P\left(X^{-1}(e)\right),$$

ceea ce încheie demonstraţia. □

Media şi media pătratică ale unei variabilei aleatoare se pot exprima în funcţie de repartiţia sa astfel

$$EX = \int x P_X(dx),\ E\left(X^2\right) = \int x^2 P_X(dx).$$

Aceste relaţii rezultă prin aplicarea lemei anterioare funcţiilor reale $f(x) = x$ şi $f(x) = x^2$.

Dispersia

Pentru a distinge între ele variabilele aleatoare, se definesc diverşi parametri care exprimă sau măsoară diverse proprietăţi ale acestora. Un astfel de parametru este media unei variabile aleatoare. O altă mărime importantă asociată unei varibile este dispersia sa. Dispersia este definită pentru variabile aleatoare (reale) care admit momente de ordinul doi. Numărul $D^2X := E(X^2) - (EX)^2$ este numit *dispersia* variabilei.

Ţinând cont de relaţiile de mai sus rezultă că dispersia se exprimă prin

$$D^2X = \int x^2 P_X(dx) - \left(\int x P_X(dx)\right)^2 .$$

În particular, dacă două variabile au aceeaşi repartiţie, atunci ele au aceeaşi medie, acelaşi moment de ordin doi şi aceeaşi dispersie.

Dispersia se exprimă şi sub forma $D^2X = E(X - EX)^2$. Într-adevăr, dacă notăm media cu $m = EX$, atunci avem

$$E(X-m)^2 = E(X^2 - 2mX + m^2) = \\ = E(X^2) - 2mEX + m^2 = E(X^2) - m^2.$$

De fapt această a doua expresie a dispersiei pune în evidenţă interesul noţiunii. Dispersia este media pătratului diferenţei $X - EX$. Este clar că pentru o variabilă constantă $X \equiv a$, avem $D^2X = 0$.

Numărul $DX := \sqrt{D^2X}$ se numeşte *deviaţia standard* a variabilei sau *abaterea medie pătratică.* Dispersia ca şi deviaţia standard măsoară cât de împrăştiate sunt valorile variabilei faţă de media EX. Deviaţia standard este mai intuitivă din acest punct de vedere pentru că este de acelaşi ordin de mărime cu variabila X. Adică, dacă înmulţim pe X cu o constantă pozitivă, a, vom avea $D^2(aX) = a^2D^2X$, pe când $D(aX) = aDX$.

Desigur că o altă mărime care măsoară aceeaşi împrăştiere este şi $E|X - EX|$. Aceasta însă este mai puţin convenabilă pentru manipulare numerică. De exemplu, se comportă mai greoi în raport cu adunarea variabilelor aleatoare. Media pătratică este legată de norma pătratică care face din spaţiul variabilelor de pătrat integrabile, L^2, un spaţiu Hilbert şi acesta are proprietăţi speciale. Media $E|X|$ este legată de spaţiul L^1, care este mai dificil de abordat în privinţa calculelor. În lema de mai jos apare unul din avantajele ce le are calculul în L^2.

Lema 4.13. *Dacă $X_1, ..., X_n$ sunt variabile aleatoare independente, de pătrat integrabile, atunci dispersia sumei are expresia*

$$D^2(X_1 + ... + X_n) = D^2X_1 + ... + D^2X_n .$$

DEMONSTRAŢIE. Pentru verificarea formulei vom nota $m_i = EX_i, i = 1, ..., n$. Avem

$$E(X_1 + ... + X_n)^2 = EX_1^2 + ... + EX_n^2 + 2\sum_{i<j} m_i m_j ,$$

$$(EX_1 + ... + EX_n)^2 = m_1^2 + ... + m_n^2 + 2\sum_{i<j} m_i m_j .$$

Prin scăderea acestor relaţii se obţine expresia dispersiei sumei. □

Se obişnuieşte utilizarea următoarei notaţii prescurtate

$$P(X \geq a, Y \geq b) := P(\{\omega \in \Omega / X(\omega) \geq a, Y(\omega) \geq b\})$$

pentru mulţimi determinate de variabile aleatoare cu valori reale. Şi există o mare varietate de notaţii de acelaşi tip. În enunţul de mai jos utilizăm această notaţie. Următoarea inegalitate deşi elementară are numeroase aplicaţii.

Lema 4.14. *(Cebâşev -Bienaymé) Dacă variabila X admite moment de ordinul doi, atunci are loc inegalitatea*

$$P(|X - m| \geq \alpha) \leq \frac{1}{\alpha^2} D^2 X,$$

unde $\alpha > 0$ este o constantă arbitrară şi $m = EX$.

Demonstraţie. Se porneşte de la observaţia că are loc egalitatea de mulţimi

$$\{|X - m| \geq \alpha\} = \left\{(X - m)^2 \geq \alpha^2\right\},$$

care conduce la următoarea inegalitate între variabile aleatoare

$$\alpha^2 1_{\{|X - m| \geq \alpha\}} \leq (X - m)^2 .$$

Atunci mediile celor două variabile vor satisface inegalitatea

$$\alpha^2 E 1_{\{|X - m| \geq \alpha\}} \leq E(X - m)^2 .$$

Ţinând cont că $E1_{\{|X - m| \geq \alpha\}} = P(|X - m| \geq \alpha)$ se deduce inegalitate $\alpha^2 P(|X - m| \geq \alpha) \leq E(X - m)^2$, care este tocmai inegalitatea din enunţ. □

5. Legea numerelor mari

În continuare ne vom referi la un spaţiu probabilizat pe care există un şir infinit de variabile aleatoare independente cu anumite caracteristici date. Precizăm că independenţa unui şir de variabile înseamnă că fiecare număr finit dintre elementele şirului sunt independente între ele. La fel pentru un şir de evenimente, independenţa şirului are sensul de independenţă a fiecărui număr finit dintre elementele şirului.

Numim *schemă a lui Bernoulli* un spaţiu probabilizat $(\Omega, \mathcal{F}, P)$ pe care există un şir $A_1, A_2, ...$ de evenimente independente cu aceeaşi probabilitate $P(A_k) = p \in (0,1)$, pentru orice $k = 1, 2, ...$

[Putem de exemplu să gândim că se modelează o serie de aruncări cu zarul în care A_i reprezintă evenimentul de a ieşi valoarea 6 la cea de a $i-$a aruncare. Într-un astfel de caz este normal ca să presupunem aceste evenimente independente, iar $P(A_i) = \frac{1}{6}$. Dacă fiecare aruncare s-ar executa cu două zaruri, iar A_i ar corespunde realizării evenimentului că zarurile au ambele 6, atunci ar trebui să considerăm $p = P(A_i) = \frac{1}{36}$.]

Se numeşte *repartiţie Bernoulli* măsura $\mu = p\delta_1 + (1-p)\delta_0$, suportată de mulţimea cu două puncte $\{0, 1\}$, unde $p \in (0,1)$ este un parametru. Dacă un eveniment A, pe un spaţiu probabilizat $(\Omega, \mathcal{F}, P)$ are probabilitatea $P(A) = p \in (0,1)$, atunci variabila aleatoare 1_A va avea repartiţia lui Bernoulli cu parametrul p. Să observăm că pentru o mulţime dată, A, σ-algebra generată de ea coincide cu σ-algebra generată de variabila canonic asociată, 1_A, adică

$$\sigma(A) = \sigma(1_A) = \{A, A^c, \Omega, \emptyset\}.$$

Rezultă că mai multe evenimente, $A_1, ..., A_n$, sunt independente dacă şi numai dacă variabilele asociate, $1_{A_1}, ..., 1_{A_n}$, sunt independente. Atunci deducem că schema lui Bernoulli este tot una cu un spaţiu probabilizat pe care avem un şir de variabile aleatoare independente $X_1, X_2, ...$ toate având repartiţia Bernoulli cu acelaşi parametru.

Reamintim că în *propoziţia* 4.4 am construit un număr finit de variabile aleatoare independente cu repartiţii date, pe un spaţiu produs. Ca un caz particular, fiind dat $p \in (0,1)$, considerăm pe $E = \{0,1\}$ repartiţia lui Bernoulli $\mu = p\delta_1 + (1-p)\delta_0$ şi apoi facem produsul $\Omega = E^n$, pe care definim măsura produs $P = \mu \bigotimes ... \bigotimes \mu$. Atunci variabilele de proiecţie $X_i : \Omega \to E, i = 1, ..., n$, sunt independente, identic repartizate, cu repartiţia μ, şi mulţimile

$$A_i = X_i^{-1}(1) = E \times ... \times \{1\} \times ... \times E, \ i = 1, ..., n,$$

sunt independente şi au aceeaşi probabilitate $P(A_i) = p.$ Pentru a obţine un şir ar trebui făcut un produs infinit, dar produsul infinit de măsuri nu va fi abordat în acest curs. În cazul $p = \frac{1}{2}$ am prezentat construcţia lui Lebesgue, care asigură existenţa schemei lui Bernoulli de parametru $p = \frac{1}{2}$.

Să observăm că pentru a studia probabilistic perioadele de constanţă ale seriilor de aruncări cu moneda avem nevoie de schema lui Bernoulli, infinită. Pe un model finit, de exemplu, avem o restricţie apriori pentru numărul de rezultate identice consecutive. Chiar dacă probabilitatea de a se obţine secvenţe lungi de rezultate identice consecutive este mică, ea nu este zero şi avem nevoie de un model infinit pe care să apară posibilitatea unor astfel de secvenţe de orice lungime. Vezi secţiunile despre *Repartiţia geometrică* şi despre *Aruncarea repetată cu banul.*

Afirmaţia (iii) din propoziţia care urmează se numeşte *legea slabă a numerelor mari.* Ea dă un sens ideii că media unei variabile este tocmai valoarea pe care o obţinem prin repetarea experimentului probabilist care este modelat de variabilă, în condiţii identice şi fără nici o condiţionare. În acelaşi sens se interpretează şi relaţia de la punctul (i). Însă din punct de vedere practic, inegalitatea (ii) este cea mai utilă pentru că ne estimează o probabilitate. Această inegalitate este adesea numită tot *inegalitatea Cebâşev* (Pafnuti Lvovici, matematician rus, 1821 - 1894). De obicei, în exemple concrete, interesul este de a arăta că probabilitatea de la (ii) este mică.

Propoziţia 4.8. *Fie $X_1, X_2, ...$ un şir de variabile aleatoare independente cu aceeaşi medie şi deviaţie standard: $m = EX_i, \sigma = DX_i$. Notând $S_n = X_1 + ... + X_n$, atunci au loc următoarele relaţii:*
(i) $\lim_{n\to\infty} E\left(\frac{1}{n}S_n - m\right)^2 = 0$,
(ii) $P\left(\left|\frac{1}{n}S_n - m\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{n\varepsilon^2}$, *pentru orice* $\varepsilon > 0$,
(iii) $\lim_{n\to\infty} P\left(\left|\frac{1}{n}S_n - m\right| \geq \varepsilon\right) = 0$, *pentru orice* $\varepsilon > 0$.

Demonstraţie. (i) Media variabilei S_n este $ES_n = nm$. Ţinem cont de independenţa variabilelor şi calculăm

$$E\left(\frac{1}{n}S_n - m\right)^2 = \frac{1}{n^2}E\left(S_n - nm\right)^2 = \frac{1}{n^2}D^2\left(S_n\right) = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n},$$

care evident tinde la zero.

(ii) Variabila $\frac{1}{n}S_n$ are media m şi dispersia, calculată mai sus, $\frac{\sigma^2}{n}$. Inegalitatea de demonstrat se obţine direct din aplicarea inegalităţii Cebâşev -Bienaymé pentru această variabilă.

(iii) Relaţia rezultă direct din punctul (ii). Propoziţia este demonstrată. □

Relaţia de la punctul (iii) exprimă „*convergenţa în probabilitate*" a şirului de variabile aleatoare $\left(\frac{1}{n}S_n\right)_{n\in\mathbf{N}^*}$ către m. Un caz particular este formulat în următorul corolar, cunoscut sub numele de *legea slabă a numerelor mari,* rezultat stabilit de Jacob Bernoulli (1654 -1705). La vremea respectivă încă nu exista teoria măsurii, iar demonstraţia originală este foarte lungă. Variabila $\frac{1}{n}S_n$ reprezintă frecvenţa de apariţie a evenimentului caracteristic unui experiment într-o serie de n realizări independente ale experimentului. Formulat într-o frază, se poate spune: frecvenţa de apariţie a unui eveniment într-un şir de experimente identice şi independente converge în probabilitate la probabilitatea evenimentului.

Corolarul 4.1. *Fie $(A_n)_{n\in\mathbf{N}^*}$ un şir de evenimente independente cu aceeaşi probabilitate $P(A_n) = p$. Atunci variabila $S_n = 1_{A_1} + ... + 1_{A_n}$ satisface relaţia*

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}S_n - p\right| \geq \varepsilon\right) = 0,$$

pentru orice $\epsilon > 0$.

Un rezultat mult mai puternic este enunţat în teorema de mai jos, cunoscută sub numele de *legea tare a numerelor mari.* Demonstraţia este dificilă şi nu o vom prezenta.

Teorema 4.2. *Fie $(\Omega, \mathcal{F}, P)$ un spaţiu probabilizat pe care sunt date variabilele $X_1, X_2, ...$ independente, identic repartizate, cu media $EX_1 = m$. Atunci are loc relaţia*

$$\lim_{n\to\infty} \frac{1}{n}\left(X_1 + ... + X_n\right) = m, \ a.s.$$

Relaţia din enunţul teoremei are loc aproape sigur, adică cu excepţia unei mulţimi de probabilitate zero.

6. Exerciţii

Exerciţiul 4.1. *Fie* $f : \Lambda \to E$ *o aplicaţie de la o mulţime arbitrară* Λ *cu valori într-o mulţime cel mult numărabilă* E. *Notăm cu* $\sigma(f)$ *cea mai mică* $\sigma-$*algebră pe* Λ *care face din* f *o aplicaţie măsurabilă, când pe* E *se consideră* $\sigma-$*algebra* $\mathcal{P}(E)$, *a părţilor lui* E. *Să se arate că o funcţie* $g : \Lambda \to \mathbf{R}$ *este măsurabilă faţă de* $\sigma(f)$ *dacă şi numai dacă există o funcţie* $h : E \to \mathbf{R}$, *astfel încât* $g = h \circ f$.

Exerciţiul 4.2. ** Dacă* E *este o mulţime cu cel puţin două elemente, atunci mulţimea* $\Omega = E^{\mathbf{N}}$ *este nenumărabilă.*

Exerciţiul 4.3. *Se consideră trei urne cu bile albe şi negre după cum urmează: prima urnă are* 3 *bile albe şi* 3 *negre, a doua are* 2 *bile albe şi* 3 *negre, iar a treia* 4 *albe şi* 2 *negre. Se extrage câte o bilă din fiecare urnă, bila extrasă fiind imediat pusă la loc. Dacă se efectuează operaţia de trei ori la rând, care este probabilitatea de a se obţine de două ori câte o bilă albă şi două negre, şi o dată două bile albe şi una neagră?*

Exerciţiul 4.4. ** Fie* Ω *o mulţime şi* $\mathcal{U} = \{A_1, ..., A_n\}$ *o familie finită de părţi. Dacă* $a_1, ..., a_n \in \mathbf{R}$, *notăm* $\varphi = a_1 1_{A_1} + ... + a_n 1_{A_n}$, *care este o funcţie cu valori reale. (i) Să se arate că* φ *este constantă pe fiecare atom al partiţiei* $\left(A^\lambda\right)_{\lambda \in \Lambda}$ *asociate familiei* $\mathcal{U}$ *(vezi* lema *3.6). (ii) Dacă* $\mathcal{G}$ *este o algebră de părţi astfel că toate aplicaţiile de tipul lui* φ *sunt măsurabile faţă de ea, atunci* $a(\mathcal{U}) \subset \mathcal{G}$.

Pentru următoarele patru exerciţii presupunem că E este o mulţime cel mult numărabilă iar μ este o măsură de probabilitate pe E.

Exerciţiul 4.5. *Dacă* $G \subset E$ *este o submulţime arbitrară, să se arate că, pentru fiecare* $\varepsilon > 0$, *există o mulţime finită* $F \subset G$ *cu proprietatea că* $\mu(G \backslash F) < \varepsilon$.

Exerciţiul 4.6. *Presupunem că* $f \geq 0$ *este o funcţie integrabilă şi* g *este o altă funcţie astfel că* $|g| \leq f$. *Să se arate că atunci şi* g *este integrabilă.*

Exerciţiul 4.7. ** Să se arate că dacă* $(f_n)_{n \in \mathbf{N}}$ *este un şir crescător de funcţii pozitive, integrabile şi* $\sup_{n \in \mathbf{N}} \int f_n d\mu < \infty$, *atunci funcţia* $f = \lim_{n \in \mathbf{N}} f_n$ *este integrabilă şi* $\int f d\mu = \lim \int f_n d\mu$.

Exerciţiul 4.8. ** Fie* $(f_n)_{n \in \mathbf{N}}$ *un şir de funcţii care are limita punctuală* f. *Presupunem că există o funcţie integrabilă* $g \geq 0$ *astfel ca* $|f_n| \leq g$, *pentru orice* $n \in \mathbf{N}$. *Să se arate că atunci* f *este integrabilă şi că are loc relaţia* $\int f d\mu = \lim \int f_n d\mu$.

Exerciţiul 4.9. ** Fie E o mulţime cel mult numărabilă şi $(\mu_e)_{e\in E} \subset \mathbf{R}$ o familie absolut sumabilă de numere reale. Să se arate că există o unică aplicaţie $\mu : \mathcal{P}(E) \to \mathbf{R}$ astfel încât pentru orice şir descrescător $(A_n)_{n\in\mathbf{N}} \subset \mathcal{P}(E)$ cu intersecţia vidă ($A_{n+1} \subset A_n, \forall n \in \mathbf{N}$, $\bigcap_{n\in\mathbf{N}} A_n = \emptyset$) să aibă loc relaţia $\lim_{n\to\infty} \mu(A_n) = 0$ şi astfel încât $\mu(\{e\}) = \mu_e$, pentru orice $e \in E$. Spunem că μ este o măsură finită pe E. Să se calculeze numerele*

$$\sup\{\mu(A) \,/\, A \subset E\},\ \inf\{\mu(A) \,/\, A \subset E\},$$
$$|\mu| = \sup\left\{\int f d\mu / \ f : E \to R,\ |f| \le 1\right\}.$$

Exerciţiul 4.10. *Fie E_1, E_2, E_3 trei mulţimi cel mult numărabile şi μ_1, μ_2, μ_3 măsuri de probabilitate definite pe ele. Identificând spaţiile produs $(E_1 \times E_2) \times E_3 \cong E_1 \times (E_2 \times E_3) \cong E_1 \times E_2 \times E_3$ să se arate că are loc egalitatea $(\mu_1 \bigotimes \mu_2) \bigotimes \mu_3 = \mu_1 \bigotimes (\mu_2 \bigotimes \mu_3) = \mu_1 \bigotimes \mu_2 \bigotimes \mu_3$.*

Exerciţiul 4.11. ** Fie $(E_i)_{i\in I}$ o familie finită de mulţimi cel mult numărabile. Presupunem că mulţimea I este partiţionată sub forma $I = \cup_{j\in J} I_j$, unde mulţimile din partiţie sunt indexate astfel încât $I_j \cap I_l = \emptyset$, dacă $j \neq l$. Notăm $E^j = \prod_{i\in I_j} E_i$ şi $\mu^j = \bigotimes_{i\in I_j} \mu_i$. Să se arate că $\prod_{i\in I} E_i$ este izomorf cu $\prod_{j\in J} E^j$ şi sub acest izomorfism avem $\bigotimes_{i\in I} \mu_i = \bigotimes_{j\in J} \mu^j$.*

Exerciţiul 4.12. *Fiind date numerele $p_1, ..., p_n \in (0,1)$, să se construiască un spaţiu probabilizat pe care există evenimentele $A_1, ..., A_n$, independente şi cu probabilităţile $P(A_1) = p_1, ..., P(A_n) = p_n$.*

Exerciţiul 4.13. *Fie $n \geq 2$. Presupunem că dintr-o urnă cu n bile se fac două extrageri fără revenire. Fie X_1 variabila aleatoare care arată rezultatul primei extrageri şi X_2 variabila care arată rezultatul celei de a doua extrageri. (i) Să se arate că cele două variabile au aceeaşi repartiţie. (ii) Ele nu sunt independente.*

Exerciţiul 4.14. ** Fie $(\Omega, \mathcal{F}, P)$ un spaţiu probabilizat şi $\mathcal{A}_1 = (A_{1,i})_{i\in I_1}$, $\mathcal{A}_2 = (A_{1,i})_{i\in I_2}$ două partiţii. Definim variabilele aleatoare $X_k : \Omega \to I_k$ prin $X_k(\omega) = i$, dacă $\omega \in A_i$, unde $k = 1, 2$. Să se verifice că partiţiile $\mathcal{A}_1$ şi $\mathcal{A}_2$ sunt independente dacă şi numai dacă $P_{X_1} \bigotimes P_{X_2} = P_Z$, unde $Z : \Omega \to I_1 \times I_2$ este aplicaţia definită prin $Z(\omega) = (X_1(\omega), X_2(\omega))$. Să se deducă o nouă demonstraţie pentru afirmaţia că independenţa partiţiilor este echivalentă cu independenţa $\sigma-$algebrelor $\sigma(\mathcal{A}_1)$ şi $\sigma(\mathcal{A}_2)$. Generalizare la cazul mai multor partiţii.*

Exerciţiul 4.15. *Un zar măsluit are următorul comportament la aruncare: feţele cu numere pare au aceeaşi probabilitate de apariţie; feţele cu numere impare au, de asemenea, probabilităţi egale între ele; o faţă cu număr par are şanse duble de a ieşi faţă de o faţă cu număr* *impar. 1) Să se determine probabilitatea cu care iese fiecare faţă separat. 2) Se aruncă de trei ori zarul. Să se determine probabilitatea: a) să iasă exact de două ori o cifră impară, b) să iasă trei cifre consecutive, în orice ordine.*

Exerciţiul 4.16. *Trei sportivi trag la ţintă pe rând. Se ştie că primul are probabilitatea de a nimeri ţinta* $p_1 = 0,4$, *al doilea are probabilitatea* $p_2 = 0,5$, *iar al treilea* $p_3 = 0,7$. *Să se determine probabilitatea ca exact doi sportivi să fi nimerit ţinta.*

Exerciţiul 4.17. *La o roată de ruletă există* 38 *de sectoare în care se poate opri indicatorul, cu şanse egale pentru fiecare sector. Un jucător care pariază simplu pe un sector plăteşte suma* x *la intrarea în joc. În caz că, după rotirea ruletei, iese numărul ales, jucătorul câştigă de* 36 *de ori valoarea pe care a pariat. Care este media câştigului pe care îl obţine cazinoul în acest caz?*

Exerciţiul 4.18. *Doi vânători merg la vânătoare de vulpi într -o pădure unde există* 20 *de exemplare.* V_1 *nimereşte cu probabilitatea* $p_1 = 0,15$, *iar* V_2 *cu probabilitatea* $p_2 = 0,20$. *Cei doi discută posibilităţile de organizare a vânătorii în trei variante: (i) Îşi împart locul de vânătoare în două zone, fiecare având* 10 *vulpi şi fiecare vânează separat. (ii) Merg împreună la vânătoare şi trag amândoi asupra fiecărei vulpi. O vulpe găsită cu două gloanţe este împărţită în două. O vulpe găsită cu un singur glonţ este împărţită proporţional cu probabilităţile condiţionate corespunzătoare fiecărui vânător. (iii) Locul vânătorii este împărţit în două zone, fiecare având câte* 10 *vulpi şi în timpul dimineţii fiecare vânează pe câte o zonă, trăgând asupra tuturor vulpilor de acolo. După -amiază îşi schimbă între ei zonele şi trag în vulpile rămase.*

Se presupune că fiecare trage doar un foc asupra fiecărei vulpi şi că este suficient de abil ca să descopere toate vulpile, în orice variantă. Să se determine câştigul mediu aşteptat de fiecare vânător pentru fiecare variantă, ştiind că valoarea unei vulpi este de 200 *lei.*

Exerciţiul 4.19. *La o loterie* 6 *din* 49 *se notează cu* X *cel mai mare din numerele extrase şi* Y *cel mai mic din numerele extrase. 1) Să se determine repartiţiile lui* X *şi* Y. *2) Să se determine repartiţia vectorului* (X, Y). *3) Fie* Z *numărul bilelor extrase ce poartă numere mai mari decât toate numerele neextrase. Să se determine repartiţia lui* Z.

Exerciţiul 4.20. *Se extrag cu întoarcere două numere din mulţimea finită* $\{1, 2, ..., n\}$. *Fie* X *diferenţa în valoare absolută dintre aceste numere. Să se determine repartiţia şi media lui* X.

Definiţia 4.4. *Fie* μ *o repartiţie pe* $\mathbf{R}$. *Numim numere aleatoare cu repartiţia* μ *un şir* $x_1, x_2, ...$ *de numere reale ce reprezintă o realizare* $x_i = X_i(\omega), i = 1, 2, ...$ *pentru un şir de variabile aleatoare* $X_1, X_2, ...$, *independente şi identic repartizate, cu repartiţia* μ.

Din punct de vedere practic, această definiţie face cu greu distincţie concretă între un şir de numere aleatoare şi unul care nu este. De fapt, în practică avem de a face doar cu un număr finit de termeni dintr-un şir. Se vorbeşte concret doar de calitatea unui şir de numere aleatoare, iar această calitate este măsurată prin proprietăţile sale statistice: verificarea legii numerelor mari, teoremei limită centrală şi altele. În general evaluarea calităţii numerelor aleatoare este o problemă dificilă.

Totuşi lucrul cu numere aleatoare, calculele concrete, trece prin manipularea acestei definiţii. Prin această definiţie se fac transformările curente care conduc la probleme de teoria măsurii şi apoi la probleme de analiză privind estimările numerice.

Calculatoarele au programe de generare a numerelor pseudo -aleatoare. Pentru şiruri foarte lungi acestea devin de fapt periodice, ceea ce înseamnă că nu pot fi considerate aleatoare. Dar pentru multe probleme de simulare aceste numere pseudo -aleatoare se dovedesc a da rezultate foarte bune, ele aproximând bine numerele aleatoare. De obicei ele aproximează numerele aleatoare uniform repartizate pe mulţimea numerelor din $[0,1)$ ce se exprimă cu un număr dat de zecimale exacte.

Exerciţiul 4.21. *Fie* $M=\{0,1,2,3,4,5,6,7,8,9\}$ *mulţimea cifrelor de exprimare a numerelor în baza* 10 *şi fie* $E=\left\{x=\sum_{i=1}^{5}\frac{x_i}{10^i}/\ x_i\in M\right\}$ *mulţimea numerelor din intervalul* $[0,1)$ *ce se exprimă cu* 5 *zecimale exacte. Măsura uniformă pe* E *încarcă fiecare punct cu ponderea* 10^{-5}. *a) Definim aplicaţia* $g:E\times E\to F$ *prin* $g(x,y)=x+10^{-5}y$, *unde* $F=\left\{x=\sum_{i=1}^{10}\frac{x_i}{10^i}/\ x_i\in M\right\}$ *este mulţimea numerelor din* $[0,1)$ *ce se exprimă cu* 10 *zecimale exacte. Arătaţi că dacă* $x_1,x_2,...$ *este un şir de numere aleatoare uniform repartizate pe* E, *atunci* $g(x_{2n-1},x_{2n}), n=1,2,...$ *este un şir de numere aleatoare uniform repartizate pe* F. *b) Fie* μ *măsura uniformă pe* E *şi* $X_i:E\to M$ *aplicaţia care reprezintă a i -a cifră zecimală a unui număr din* E, *pentru* $i=1,2,3,4,5$. *Să se arate că, sub măsura* μ, *variabilele acestea sunt independente, identic repartizate, cu repartiţia uniformă pe* M. *c) Arătaţi că fiind dat un şir de numere aleatoare uniform repartizate pe* E *şi o permutare* π *a mulţimii* $\{1,2,3,4,5\}$, *prin permutarea cifrelor zecimale cu* π *se obţine tot un şir de numere aleatoare cu aceeaşi repartiţie.*

CAPITOLUL 5

Câteva repartiţii pe N

În acest capitol vom discuta despre câteva exemple importante de repartiţii pe mulţimea numerelor naturale $\mathbf{N}$. Deoarece orice măsură μ pe $(\mathbf{N},\mathcal{P}(\mathbf{N}))$ se poate identifica cu şirul valorilor pe puncte $\mu_n = \mu(\{n\}), n \in \mathbf{N}$ ea se exprimă cu ajutorul măsurilor Dirac sub forma: $\mu = \sum_{n\in\mathbf{N}} \mu_n \delta_n$.

Reamintim că *repartiţia Bernoulli* este o măsură suportată de punctele 0 şi 1 şi este dată de formula $q\delta_0 + p\delta_1$, cu parametrul $p \in (0,1)$ şi $q = 1 - p$.

Fiind dat un eveniment $A \in \mathcal{F}$ pe un spaţiu probabilizat $(\Omega, \mathcal{F}, P)$ astfel încât $P(A) \in (0,1)$, se poate defini variabila aleatoare $X = 1_A$, care are repartiţia $P_X = q\delta_0 + p\delta_1$. Este deci o repartiţie Bernoulli cu parametrul $p = P(A)$. Media variabilei este egală cu momentul de ordinul doi $EX = EX^2 = p$, iar dispersia sa este $D^2X = p - p^2 = pq$.

1. Repartiţia geometrică

Aceasta este o repartiţie pe $\mathbf{N}$, dată de expresia $\sum\limits_{k=0}^{\infty} pq^k\delta_k$, care are masele ataşate punctelor în progresie geometrică cu primul termen $p \in (0,1)$ şi cu raţia $q = 1 - p$. Se verifică imediat că suma totală a acestor mase este $\sum\limits_{k=0}^{\infty} pq^k = \frac{p}{1-q} = 1$.

O variabilă aleatoare tipică, având repartiţia geometrică, este construită în felul următor. Presupunem că pe spaţiul probabilizat $(\Omega, \mathcal{F}, P)$ avem un şir de evenimente independente $A_1, A_2, ...$ şi astfel că $P(A_i) = p \in (0,1)$ este o valoare fixă pentru $i = 1, 2, ...$ Cum am mai spus, aceasta este o schemă Bernoulli. Dacă evenimentul A_i are semnificaţia de „succes la încercarea a $i-$a”, putem să definim o variabilă $V : \Omega \to \mathbf{N}$ care ne spune câte eşecuri s-au petrecut înainte de primul succes. Punctual ea este definită prin $V(\omega) = i$ pentru $\omega \in D_i$, unde evenimentele $D_0, D_1, D_2, ...$ sunt definite astfel: $D_0 = A_1$ şi pentru $i \geq 1, D_i = A_1^c \cap ... \cap A_i^c \cap A_{i+1}$, care este evenimentul ce constă din nerealizarea evenimentelor $A_1, ..., A_i$ şi realizarea lui A_{i+1}.

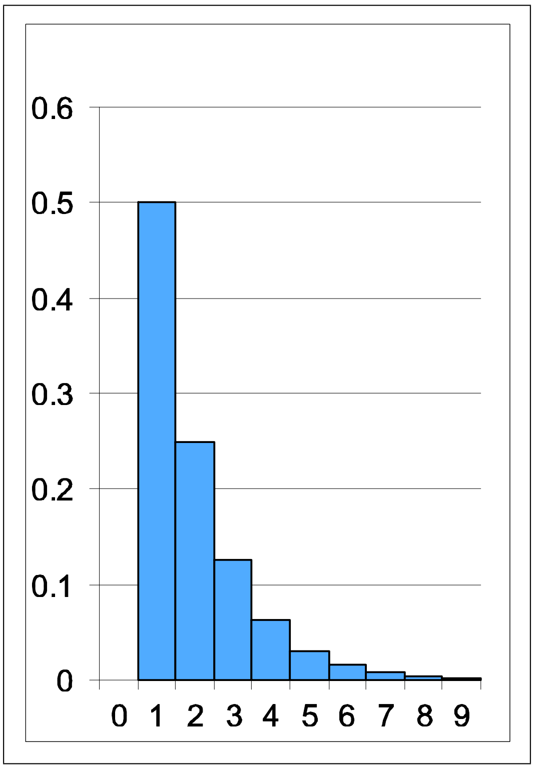


FIGURA 1. Repartiţia geometrică cu $p = 0,5$.

Este clar că evenimentele $D_0, D_1, ..., D_n$ sunt disjuncte şi

$$\bigcup_{i=1}^{n+1} A_i = \bigcup_{i=0}^{n} D_i.$$

Avem $P(D_i) = q^i p$ şi atunci $P\left(\bigcup_{i=1}^{n+1} A_i\right) = P\left(\bigcup_{i=0}^{n} D_i\right) = p + pq + ... + pq^n = p\frac{1-q^{n+1}}{1-q} = 1 - q^{n+1}$. Rezultă că $P\left(\bigcup_{i=1}^{\infty} A_i\right) = 1$ şi, prin urmare, $P\left(\Omega \backslash \bigcup_{i=0}^{\infty} A_i\right) = 0$. Dacă această ultimă mulţime este nevidă punem $D_\infty = \Omega \backslash \bigcup_{i=0}^{\infty} A_i$ şi $\overline{\mathbf{N}} = \mathbf{N} \cup \{\infty\}$, şi atunci familia de evenimente $(D_i)_{i \in \overline{\mathbf{N}}}$ formează o partiţie. Are loc formula

$$P(V = k) = pq^k,$$

pentru $k = 0, 1, ...$ şi $P(D_\infty) = 0$. Dacă punem $V(\omega) = \infty$ pentru $\omega \in D_\infty$, obţinem o variabilă cu valori în $\overline{\mathbf{N}}$. Pentru a avea o variabilă care ia numai valori reale, putem de exemplu să definim $V = 0$ şi pe mulţimea D_∞. Oricum V este repartizată geometric, valoarea ∞ nefiind importantă.

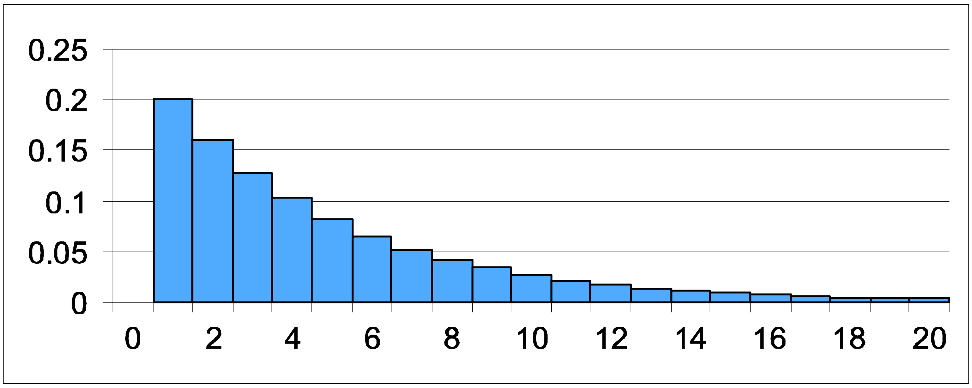


FIGURA 2. Repartiţia geometrică cu $p = 0,8$.

Dacă se modelează un joc de noroc care se desfăşoară în serii de partide, iar A_i este evenimentul „la a $i-$a partidă jucătorul a câştigat", atunci V reprezintă numărul de partide pierdute înainte de a câştiga prima oară.

Media unei variabile repartizate geometric este

$$EV = \sum_{k=0}^{\infty} kP(V=k) = \sum_{k=1}^{\infty} kpq^k = pq \sum_{k=1}^{\infty} \frac{d}{dq}\left(q^k\right) = pq\frac{d}{dq}\left(\sum_{k=1}^{\infty} q^k\right) =$$

$$= pq\frac{d}{dq}\left(\frac{q}{1-q}\right) = pq\frac{1}{(1-q)^2} = \frac{q}{p}.$$

Pentru a calcula dispersia, calculăm mai întâi media lui V^2 :

$$EV^2 = \sum_{k=0}^{\infty} k^2pq^k = p\sum_{k=1}^{\infty} k(k-1)q^k + p\sum_{k=1}^{\infty} kq^k =$$

$$= pq^2\frac{d^2}{dq^2}\sum_{k=2}^{\infty} q^k + \frac{q}{p} = pq^2\frac{d^2}{dq^2}\left(\frac{q^2}{1-q}\right) + \frac{q}{p} = pq^2\frac{2}{(1-q)^3} + \frac{q}{p} = \frac{2q^2+qp}{p^2}.$$

Deci dispersia este $D^2V = \frac{2q^2+qp}{p^2} - \frac{q^2}{p^2} = \frac{q}{p^2}$.

2. Repartiţia binomială

Repartiţia binomială este o repartiţie suportată de punctele $0, 1, ..., n$ şi are expresia

$$\sum_{k=0}^{n} C_n^k p^k q^{n-k} \delta_k,$$

unde $n \in \mathbf{N}^*$ este numit ordinul repartiţiei, $p \in (0,1)$ este un alt parametru şi $q = 1 - p$. Pentru a verifica că masa totală a acestei măsuri este 1 se trece prin formula binomului lui Newton:

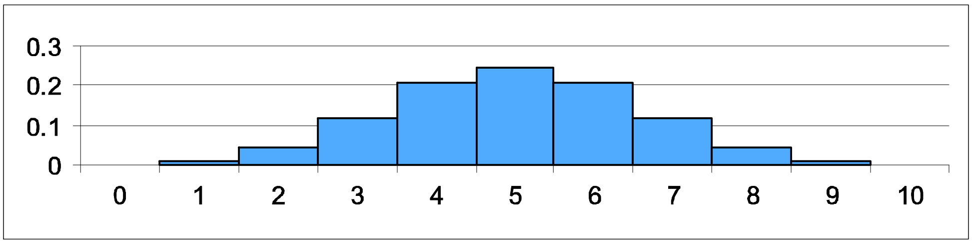


FIGURA 3. Repartiţia binomială cu $n = 10$ şi $p = 0,5$.

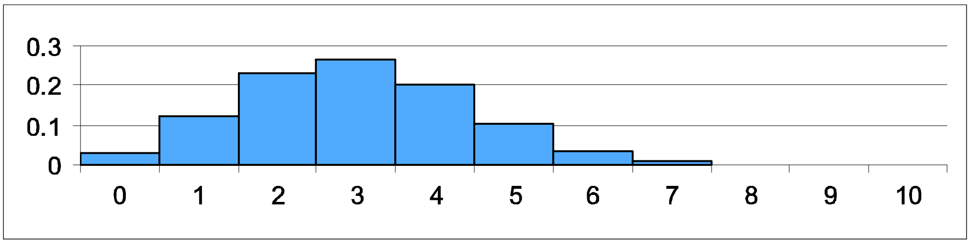


FIGURA 4. Repartiţia binomială cu $n = 10$ şi $p = 0,3$.

$$1 = (p+q)^n = \sum_{k=0}^{n} C_n^k p^k q^{n-k}.$$

De aici vine denumirea de "binomială" pentru repartiţie. Vom calcula acum media şi dispersia unei variabile aleatoare, X, definite pe un spaţiu probabilizat $(\Omega, \mathcal{F}, P)$, care este repartizată binomial cu ordinul n şi parametrul p :

$$EX = \sum_{k=0}^{n} kP(X=k) = \sum_{k=0}^{n} k \frac{n!}{k!\,(n-k)!} p^k q^{n-k} =$$

$$= np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!\,(n-k)!} p^{k-1} q^{n-k} = np \sum_{l=0}^{n-1} \frac{(n-1)!}{l!\,(n-1-l)!} p^l q^{n-1-l} = np,$$

$$EX^2 = \sum_{k=0}^{n} k^2 P(X=k) = \sum_{k=0}^{n} k^2 \frac{n!}{k!\,(n-k)!} p^k q^{n-k} =$$

$$= np \sum_{k=1}^{n} k \frac{(n-1)!}{(k-1)!\,(n-k)!} p^{k-1} q^{n-k} = np \sum_{l=0}^{n-1} (l+1) \frac{(n-1)!}{l!\,(n-1-l)!} p^l q^{n-1-l} =$$

$$= np\left[(n-1)p+1\right] = (np)^2 + npq.$$

În concluzie avem următoarea dispersie

$$D^2X = EX^2 - (EX)^2 = npq.$$

Construcţia unei variabile aleatoare tipice cu repartiţia binomială este obţinută pe un spaţiu probabilizat $(\Omega, \mathcal{F}, \mathcal{P})$ pe

care sunt date n evenimente independente $A_1, ..., A_n$ care au toate aceeaşi probabilitate $P(A_i) = p, i = 1, ..., n,$ cu $p \in (0,1)$. În acest cadru are loc următorul rezultat.

Propoziţia 5.1. *Variabila aleatoare* $S = \sum\limits_{i=1}^{n} 1_{A_i}$ *este repartizată binomial cu parametrii* n, p.

Demonstraţie. Vom proceda prin inducţie după parametrul n. Notăm, pentru $k = 1, ..., n$,

$$S_k = \sum_{i=1}^{k} 1_{A_i}.$$

Presupunem că am demonstrat formula

$$P(S_k = j) = C_k^j p^j q^{k-j},$$

pentru orice $j = 0, ..., k$ şi o vom demonstra cu $k+1$ în locul lui k. Avem $S_{k+1} = S_k + 1_{A_{k+1}}$, iar S_k şi $1_{A_{k+1}}$ sunt independente. Pentru a vedea acest lucru, mai întâi vom examina consecinţele independenţei evenimentelor $A_i, i = 1, ..., n$. Notăm $a_i = a(A_i) = \{A_i, A_i^c, \Omega, \emptyset\}$, algebra generată de evenimentul A_i, care este chiar o $\sigma-$algebră. Acestea sunt independente, ca urmare a ipotezei şi a *lemei* 2.1, iar teorema de asociativitate a independenţei, *teorema* 4.1, ne spune că $\sigma(a_1, ..., a_k)$ şi a_{k+1} sunt independente. Pe de altă parte, putem scrie $S_k = f(X_1, ..., X_k)$, unde $f : \mathbf{R}^k \to \mathbf{R}$ este funcţia sumă, definită prin $f(x_1, ..., x_k) = x_1 + ... + x_k$, iar $X_i = 1_{A_i}$. Atunci *proprietatea* 4. din paragraful despre $\sigma-$algebra generată de o variabilă ne asigură că S_k este măsurabilă în raport cu $\sigma(a_1, ..., a_k)$. Rezultă deci că S_k şi $1_{A_{k+1}}$ sunt independente.

Mai departe introducem notaţia

$$B_{k,j} = \{S_k = j\}.$$

şi observăm că mulţimile $B_{k,0}, B_{k,1}, ..., B_{k,k}$ formează o partiţie a lui Ω. Se verifică direct relaţiile

$$\begin{gathered}
B_{k+1,0} = \{S_{k+1} = 0\} = \left\{S_k = 0, 1_{A_{k+1}} = 0\right\} = B_{k,0} \cap A_{k+1}^c, \\
B_{k+1,k+1} = \{S_{k+1} = k+1\} = \left\{S_k = k, 1_{A_{k+1}} = 1\right\} = B_{k,k} \cap A_{k+1}, \\
B_{k+1,j} = \{S_{k+1} = j\} = \left\{S_k = j, 1_{A_{k+1}} = 0\right\} \cup \left\{S_k = j-1, 1_{A_{k+1}} = 1\right\} \\
= (B_{k,j} \cap A_{k+1}^c) \cup (B_{k,j-1} \cap A_k),
\end{gathered}$$

pentru $j = 1, ..., k$. Pe de altă parte, ţinând cont că mulţimile $B_{k,j}$ sunt fiecare independente de A_{k+1}, putem scrie

$$P(B_{k+1,0}) = P(B_{k,0})\, q,$$
$$P(B_{k+1,k+1}) = P(B_{k,k})\, p,$$
$$P(B_{k+1,j}) = P(B_{k,j})\, q + P(B_{k,j-1})\, p,$$

pentru $j = 1, ..., k$. Ţinând cont de ipoteza de inducţie aceste relaţii devin

$$P(B_{k+1,0}) = C_k^0 q^{k+1} = C_{k+1}^0 q^{k+1},$$
$$P(B_{k+1,k+1}) = C_k^k p^{k+1} = C_{k+1}^{k+1} p^{k+1},$$
$$P(B_{k+1,j}) = C_k^j p^j q^{k-j} q + C_k^{j-1} p^{j-1} q^{k-j+1} p =$$
$$= \left(C_k^j + C_k^{j-1}\right) p^j q^{k+1-j} = C_{k+1}^j p^j q^{k+1-j},$$

ceea ce încheie demonstraţia. □

Este instructiv de văzut că se poate demonstra că S este repartizată binomial şi printr-un calcul brut direct. Facem acest lucru mai jos. Prin comparaţie, putem spune că demonstraţia anterioară, de fapt, evită calculul prin utilizarea proprietăţii de asociere a independenţei.

Altă demonstraţie. În acest scop introducem notaţia $A_{i,1} = A_i$ şi $A_{i,-1} = A_i^c$, pentru $i = 1, ..., n$,

$$A^\lambda = A_{1,\lambda_1} \cap ... \cap A_{n,\lambda_n},$$

pentru orice sistem $\lambda = (\lambda_1, ..., \lambda_n) \in \Gamma := \{-1, 1\}^n$ şi mai notăm $s(\lambda) = \frac{n+\lambda_1+...+\lambda_n}{2}$, o mărime care indică numărul de componente ale lui λ care sunt egale cu 1. Se verifică direct că are loc relaţia următoare

$$\{S = i\} = \bigcup_{s(\lambda)=i} A^\lambda,$$

unde reuniunea se ia după toate sistemele $\lambda \in \Gamma$ astfel încât $s(\lambda) = i$, iar $i = 1, ..., n$. Mulţimile care participă la această reuniune formează o partiţie; mai precis, dacă $\lambda \neq \tau$, atunci $A^\lambda \cap A^\tau = \emptyset$. Rezultă

$$P(S = i) = \sum_{s(\lambda)=i} P\left(A^\lambda\right).$$

Dar $P\left(A^\lambda\right) = P(A_{1,\lambda_1}) \cdot ... \cdot P(A_{n,\lambda_n})$, datorită independenţei şi relaţiilor $P(A_{j,1}) = p, P(A_{j,-1}) = q$, ceea ce conduce la $P\left(A^\lambda\right) = p^i q^{n-i}$, pentru un sistem $\lambda \in \Gamma$ astfel ca $s(\lambda) = i$. Atunci avem $P(S = i) = p^i q^{n-i} card\{\lambda \in \Gamma \mid s(\lambda) = i\}$. Dar mulţimea $\{\lambda \in \Gamma \mid s(\lambda) = i\}$ este în corespondenţă bijectivă cu mulţimea părţilor lui $\{1, 2, ..., n\}$ care au cardinalul i. Acestea sunt în număr de C_n^i, ceea ce permite a trage concluzia că

$$P(S = i) = p^i q^{n-i} C_n^i,$$

terminând astfel demonstraţia. □

Prima metodă de demonstraţie are şi avantajul de a se extinde producând, cu o demonstraţie identică, următorul rezultat.

Propoziţia 5.2. *Fie* $X_1, ..., X_k$, *variabile aleatoare independente, repartizate binomial având ordinele, respectiv* $n_1, ..., n_k$ *şi cu al doilea parametru identic* $p \in (0,1)$. *Atunci suma* $X_1 + ... + X_k$ *este repartizată binomial cu ordinul* $n_1 + ... + n_k$ *şi acelaşi parametru* p.

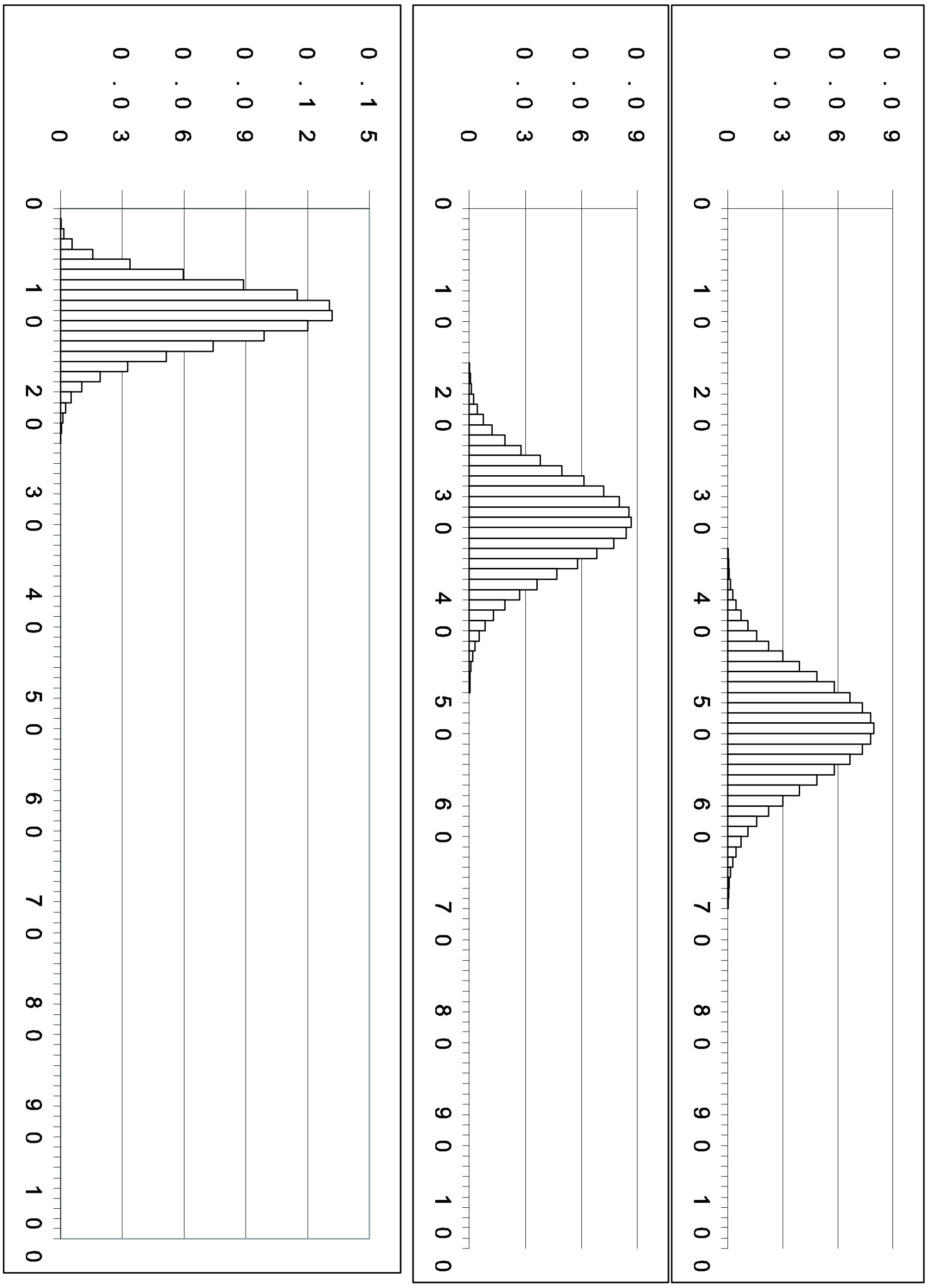


FIGURA 5. Repartiţii binomiale cu $n = 100$ şi $p = 0,5; p = 0,3; p = 0,1$.

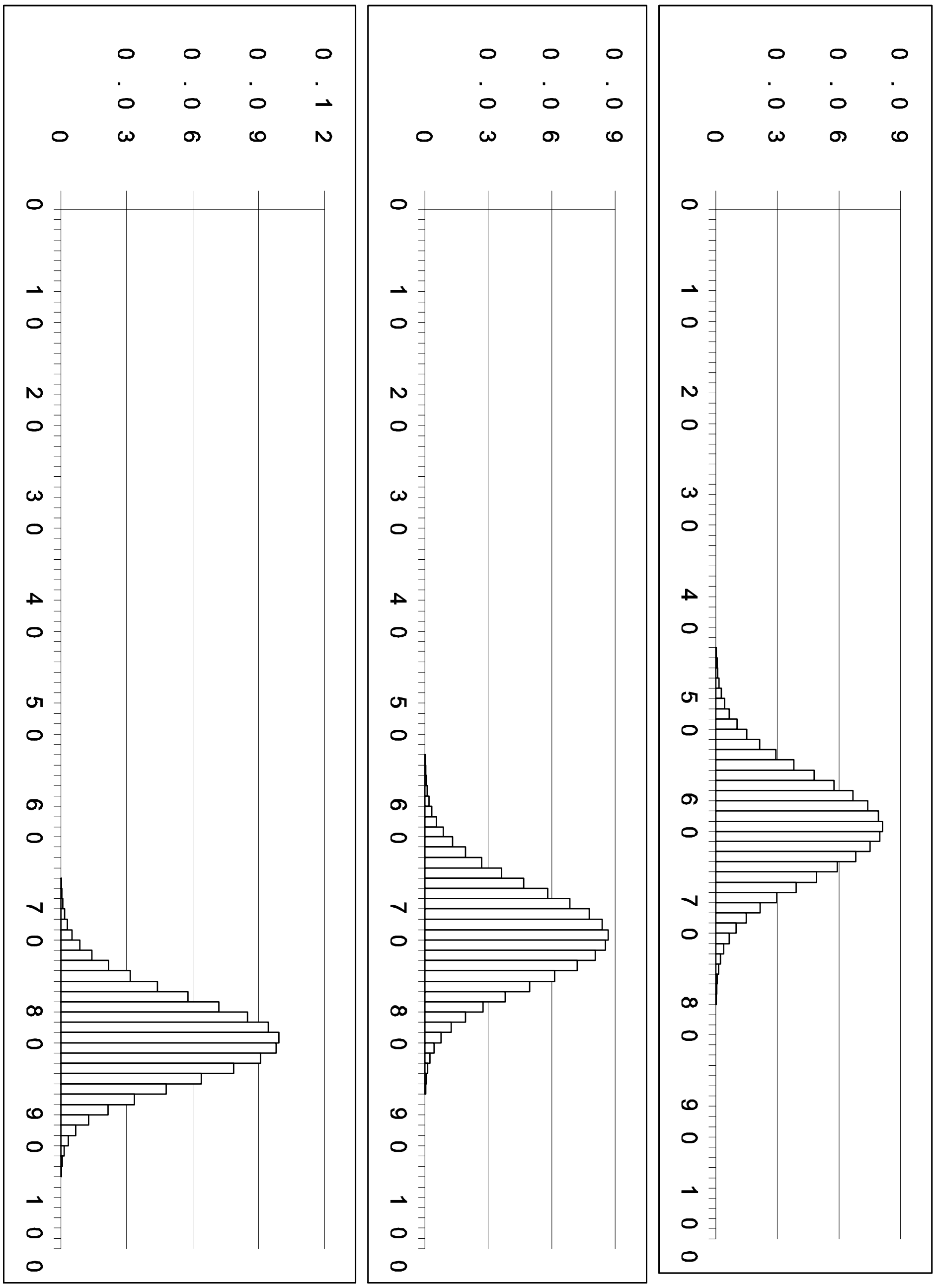


FIGURA 6. Repartiţii binomiale cu $n = 100$ şi $p = 0,6; p = 0,7; p = 0,8$.

3. Histograme

O *histogramă* este o reprezentare grafică pentru o distribuţie discretă. Ele pot fi de mai multe tipuri. Cele mai obişnuite histograme sunt cele în care pentru fiecare punct din suportul distribuţiei se desenează un dreptunghi de înălţime proporţională cu ponderea repartizată punctului. Astfel de histograme sunt curent utilizate în economie. Programul Excel are facilităţi speciale pentru executarea histogramelor.

Am redat în figurile din acest capitol câteva histograme ale unor repartiţii pe care le-am întâlnit. Este de notat că scara pentru axa verticală este dilatată faţă de cea orizontală, pentru a putea da un aspect grafic vizibil repartiţiilor. De exemplu, valorile repartiţiilor binomiale fiind foarte mici, în cazul în care n este mare, la o reprezentare cu aceeaşi scară pe ambele axe nu ar mai fi posibil să vedem nimic într-un grafic dintr-o pagină de carte. Chiar şi aşa, marea majoritate a valorilor apar în grafic confundate cu zero. În figurile 5 şi 6 se observă numai valorile din jurul mediei np.

Repartiţiile de acelaşi fel sunt redate în figurile noastre păstrând cele două scări, pentru a putea fi comparate între ele. Dar tipurile diferite de repartiţii au scări diferite.

4. Repartiţia hipergeometrică*

Fie $M = \{1, 2, ..., N\}$ o mulţime ce reprezintă bilele dintr-o urnă. Presupunem că $R \subset M$ reprezintă o submulţime de bile marcate, de exemplu vopsite în roşu. Notăm cu r cardinalul mulţimii R. Ne interesează experimentul extragerii a n bile deodată şi rezultatul acestuia cuantificat prin numărul de bile marcate ce apar la extragere. Am văzut cu ocazia exemplului privind controlul calităţii că se poate face modelarea acestui experiment pe spaţiul

$$\Gamma = \{A \in \mathcal{P}(M) \, / \, card A = n\}.$$

Bineînţeles că presupunem $r, n \leq N$. Probabilitatea pentru un eveniment elementar $A \in \Gamma$ este dată de $Q(\{A\}) = \frac{1}{C_N^n}$.

Numărul de bile roşii extrase în acest experiment este specificat de variabila aleatoare $Y : \Gamma \to \mathbf{N}$ definită prin $Y(A) = card(A \cap R)$. De fapt Y ia valori în mulţimea $\{0 \vee (n - N + r), ..., n \wedge r\}$. Conform discuţiei făcute în legătură cu identitatea Van der Monde, avem

$$Q(Y = k) = \frac{C_r^k C_{N-r}^{n-k}}{C_N^n} = \frac{r!}{k!\,(r-k)!} \frac{(N-r)!}{(n-k)!\,(N-r-n+k)!} \frac{n!\,(N-n)!}{N!},$$

pentru orice număr natural k care satisface condiţiile $n - N + r \leq k \leq n \wedge r$. Repartiţia variabilei Y este numită *repartiţia hipergeometrică* cu parametrii N şi r şi rangul n. Explicitarea acestei repartiţii în sumă de măsuri Dirac ponderate este deci

$$\sum_{k=0 \vee (n-N+r)}^{n \wedge r} \frac{C_r^k C_{N-r}^{n-k}}{C_N^n} \delta_k.$$

Pentru valori mari ale lui N şi valori relativ mici ale lui n se poate face aproximarea numerică a experimentului extragerilor a n bile deodată cu experimentul seriilor de n extrageri cu revenire, care sunt mai facile din punct de vedere calculatoric. Vom explica acum acest fapt, estimând şi eroarea care apare.

Pentru aceasta vom nota cu $\Omega = M^n$ spaţiul pe care se modelează seriile de n extrageri cu revenire şi cu

$$\Omega' = \{(x_1, ..., x_n) \in \Omega / \ x_i \neq x_j \ \forall i \neq j\}$$

vom nota submulţimea seriilor în care rezultatele tuturor extragerilor sunt diferite. Vom lua mulţimea $\Omega' \subset \Omega$ ca spaţiu de modelare a experimentului seriilor de extrageri fără revenire. Dacă P este probabilitatea definită pentru punctele din Ω prin $P(\{(x_1, ..., x_n)\}) = \frac{1}{N^n}$, pe Ω' avem probabilitatea P' definită prin $P'(\{(x_1, ..., x_n)\}) = \frac{1}{N(N-1)...(N-n+1)}$.

Fie $X_i : \Omega \to \{0, 1\}$ aplicaţia definită prin $X_i(x_1, ..., x_n) = 1_R(x_i)$. Această aplicaţie arată culoarea bilei extrase la cea de a i-a extragere din seria $(x_1, ..., x_n)$, prin faptul că $X_i(x_1, ..., x_n) = 1$, dacă $x_i \in R$ şi $X_i(x_1, ..., x_n) = 0$, dacă $x_i \notin R$. Variabilele $X_i, i = 1, ..., n$, sunt independente şi identic repartizate sub P, pentru că $P(X_i = 1) = \frac{r}{N}$. Variabila $S = X_1 + ... + X_n$ este repartizată binomial cu parametrul $p = \frac{r}{N}$ şi ordinul n. Ea indică de câte ori a fost extrasă o bilă roşie într-o serie de n extrageri cu revenire. Restricţia acestei variabile la Ω', variabila $S_{|\Omega'}$ indică numărul de bile roşii ce au fost obţinute în seria de extrageri fără revenire $(x_1, ..., x_n)$, sau numărul de bile roşii ce se află în extragerea „fără ordine" a grupului de bile $\{x_1, ..., x_n\}$. A se vedea paragraful despre extrageri fără ordine din capitolul 2. Variabila $S_{|\Omega'}$ este repartizată hipergeometric sub P', cu parametrii N, r şi rangul n, adică are aceeaşi repartiţie ca variabila Y de mai sus.

Pentru orice mulţime $A \subset \Omega'$ avem

$$P'(A) = \frac{N^n}{N(N-1)...(N-n+1)} P(A).$$

Se observă că $P'(A) \geq P(A)$. Diferenţa dintre cele două probabilităţi se poate estima evaluând factorii din dreapta astfel că obţinem următoarea lemă.

Lema 5.1. *Dacă* $\varepsilon = \frac{n^2}{N} < 1$, *atunci* $P'(A) - P(A) \leq \frac{\varepsilon}{2} P'(A)$.

DEMONSTRAŢIE. Înlocuind pe $P(A)$ în funcţie de $P'(A)$ în inegalitatea din enunţ, rezultă că ea este echivalentă cu:

$$1 - \frac{N(N-1)...(N-n+1)}{N^n} \leq \frac{\varepsilon}{2}. \qquad (*)$$

Pentru aceasta avem de evaluat expresia

$$\frac{N(N-1)...(N-n+1)}{N^n} = \left(1 - \frac{1}{N}\right) ... \left(1 - \frac{n-1}{N}\right).$$

Aceasta este minorată ţinând cont de inegalitatea $e^x e^{-\frac{x^2}{1+x}} \leq 1+x$, valabilă pentru $x > -1$:

$$\geq \exp\left(-\sum_{i=1}^{n-1}\frac{i}{N}\right)\exp\left(-\sum_{i=1}^{n-1}\frac{i^2}{N^2}\frac{1}{1-\frac{i}{N}}\right) \geq$$

$$\geq \exp\left(-\frac{n(n-1)}{2N} - \frac{n(n-1)(2n-1)}{6N(N-n)}\right).$$

Apoi ţinem cont că $\frac{n-1}{(N-n)} \leq \frac{n}{N}$, şi minorăm ultima expresie prin

$$\geq \exp\left(-\frac{n(n-1)}{2N} - \frac{n^3}{3N^2}\right) \geq \exp -\frac{\varepsilon}{2}.$$

După aceea ţinem cont de inegalitatea $e^{-x} \geq 1-x$, care permite a deduce estimarea (*). $\square$

Această estimare a diferenţei dintre probabilităţile P' şi P conduce, în particular, la o estimare a diferenţei dintre repartiţia hipergeometrică şi repartiţia binomială care este de ordinul $\frac{n^2}{N}$. Dar de fapt sunt valabile chiar estimări de ordinul $\frac{n}{N}$. Acestea se obţin prin metode mai elaborate, pe care nu este locul să le dezvoltăm aici.

5. Repartiţia negativ -binomială*

Repartiţia numită *negativ -binomială* este legată de repartiţia binomială, dar şi mai mult de repartiţia geometrică, pe care o generalizează. Expresia repartiţiei negativ -binomiale cu parametrii $n \in \mathbf{N}^*, p \in (0,1)$ este

$$\sum_{k=0}^{\infty} C_{k+n-1}^{k} p^n q^k \delta_k, \quad (*)$$

unde $q = 1-p$. Se observă că pentru $n = 1$ se obţine repartiţia geometrică. Verificarea directă, prin calcul analitic, a faptului că masa unei repartiţii negativ- binomiale este 1 se poate face trecând prin relaţia binomului lui Newton în forma generală:

$$(1+x)^r = \sum_{k=0}^{\infty}\frac{(r)_k}{k!}x^k,$$

serie convergentă pentru $|x| < 1, r \in \mathbf{R}$ şi în care utilizăm notaţia $(r)_k = r(r-1)...(r-k+1)$, cu convenţia $(r)_0 = 1$. În cazul nostru această relaţie este utilă sub forma

$$\frac{1}{p^n} = (1-q)^{-n} = \sum_{k=0}^{\infty}\frac{(-n)_k}{k!}(-q)^k.$$

Numele repartiţiei vine de la acest binom la putere negativă, care prin dezvoltare produce coeficienţii repartiţiei. Mai jos vom face o a doua demonstraţie, probabilistă, pentru faptul că masa măsurii din expresia (*) este 1.

Situaţia tipică în care apare această repartiţie, la fel ca în cazul repartiţiei geometrice, este în contextul unui şir de evenimente independente $(A_i)_{i\in\mathbf{N}^*}$ pe un spaţiu probabilizat $(\Omega,\mathcal{F},P)$, cu probabilitatea constantă $P(A_i)=p, i\in\mathbf{N}^*$. Notăm $X_i=1_{A_i}$ şi $S_j=\sum_{i=1}^{j}X_i, j\in\mathbf{N}^*$. Definim apoi variabilele $T_n:\Omega\to\mathbf{N}^*\cup\{\infty\}, n\in\mathbf{N}^*$, prin

$$T_n(\omega)=\inf\{j\in\mathbf{N}^*/S_j=n\},$$

cu convenţia că $T_n(\omega)=\infty$, dacă mulţimea din dreapta relaţiei de definiţie este vidă. Să zicem că şirul $(A_i)_{i\in\mathbf{N}^*}$ reprezintă realizarea succesivă a unui experiment probabilistic cu probabilitatea p de succes. Atunci T_n ne spune la a câta încercare apare al n -lea succes, iar $V_n=T_n-n:\Omega\to\mathbf{N}\cup\{\infty\}$ reprezintă numărul de eşecuri până la obţinerea celui de al n -lea succes. O astfel de situaţie apare adesea în aplicaţiile legate de modelele din biologie. De exemplu un tratament este repetat de mai multe ori, până ce rezultatele cumulate ating un nivel care, să zicem, asigură dispariţia paraziţilor.

Lema care urmează arată că V_n are repartiţia negativ -binomială, în particular, că măsura din formula (*) este o măsură de probabilitate.

Lema 5.2. *Au loc egalităţile*

$$P(V_n=k)=C_{k+n-1}^{k}p^nq^k,\ k\in\mathbf{N},$$
$$P(V_n=\infty)=0.$$

Demonstraţie. Avem

$$\{V_n=k\}=\{T_n=n+k\}=\{S_{n+k-1}=n-1,S_{n+k}=n\}=$$
$$=\{S_{n+k-1}=n-1,X_{n+k}=1\}.$$

Ţinând cont că S_{n+k-1} şi X_{n+k} sunt independente, obţinem prima relaţie:

$$P(T_n-n=k)=P(S_{n+k-1}=n-1)P(X_{n+k}=1)=$$
$$=C_{n+k-1}^{n-1}p^{n-1}q^kp=C_{n+k-1}^{k}p^nq^k.$$

Pentru verificarea celei de a doua relaţii din enunţ vom proceda prin inducţie. Pentru $n=1$ avem $\{T_1<\infty\}=\cup_iA_i$ şi, prin urmare,

$$P(T_1=\infty)=P(\cap_iA_i^c)=\lim_{k\to\infty}P\left(\cap_{i=1}^{k}A_i^c\right)=\lim_{k\to\infty}q^k=0.$$

Mai departe presupunem că ştim că $P(T_n=\infty)=0$ şi vom arăta că aceeaşi relaţie este valabilă şi pentru $n+1$. Mai precis, vom arăta că, pentru fiecare $k\geq 0$, are loc relaţia

$$P(T_n=n+k,T_{n+1}=\infty)=0.$$

Pentru aceasta notăm că are loc incluziunea

$$\{T_n=n+k\}\cap\left(\cup_{i=n+k+1}^{\infty}A_i\right)\subset$$
$$\subset\{S_{n+k}=n\}\cap\left\{\sup_{i\geq n+k+1}S_i\geq n+1\right\}\subset\{T_{n+1}<\infty\}.$$

Aceasta implică

$$\{T_n = n + k, T_{n+1} = \infty\} \subset \cap_{i=n+k+1}^{\infty} A_i^c.$$

Deoarece $P\left(\cap_{i=n+k+1}^{\infty} A_i^c\right) = 0$, rezultă concluzia dorită. □

Relaţia dintre repartiţia negativ -binomială şi repartiţia geometrică este pusă în evidenţă şi din alt punct de vedere în lema care urmează. Facem convenţia că $T_0 = 0$ şi notăm $\tau_n = T_n - T_{n-1}$, pentru $n \in \mathbf{N}^*$.

Lema 5.3. *Variabilele $\tau_1 - 1, \tau_2 - 1, ...$ sunt independente şi identic repartizate, cu repartiţia geometrică de parametru p.*

Demonstraţie. Este clar că $\tau_1 - 1 = T_1 - 1$ şi ştim că acesta are repartiţia geometrică de parametru p. În continuare pornim cu $k_1, ..., k_n \in \mathbf{N}$ şi scriem egalitatea

$$\{\tau_1 = 1 + k_1, ..., \tau_n = 1 + k_n\} =$$
$$= \{X_1 = 0, ..., X_{k_1} = 0, X_{k_1+1} = 1, X_{k_1+2} = 0, ...X_{k_1+1+k_2} = 0,$$
$$X_{k_1+1+k_2+1} = 1, ..., X_{k_1+1+...+k_{n-1}+1+k_n} = 0, X_{k_1+1+...+k_n+1} = 1\}$$

Aceasta permite să calculăm, ţinând cont şi de independenţă,

$$P(\tau_1 = 1 + k_1, ..., \tau_n = 1 + k_n) = q^{k_1} p q^{k_2} p ... q^{k_n} p = p^n q^{k_1 + ... + k_n}. \quad (**)$$

De aici se deduce, luând $n = 2$,

$$P(\tau_2 = 1 + k) = \sum_{l=0}^{\infty} P(\tau_1 = 1 + l, \tau_2 = 1 + k) = \sum_{l=0}^{\infty} p^2 q^{l+k} = pq^k,$$

adică $\tau_2 - 1$ are repartiţia geometrică. Acest calcul se generalizează şi, în mod similar, se deduce că repartiţia lui $\tau_n - 1$ este tot geometrică. Odată stabilit faptul că toate variabilele $\tau_n - 1, n \in \mathbf{N}^*$, sunt repartizate identic, observăm că relaţia (**) devine

$$P(\tau_1 = 1 + k_1, ..., \tau_n = 1 + k_n) = P(\tau_1 = 1 + k_1) ... P(\tau_n = 1 + k_n),$$

ceea ce arată independenţa lor. Cu aceasta se încheie demonstraţia. □

Dar este evident că $V_n = \sum_{i=1}^{n} (\tau_i - 1)$. De aceea putem spune că repartiţia negativ -binomială de parametru p şi ordin n este repartiţia sumei a n variabile independente, identic repartizate, cu repartiţie geometrică de parametru p.

6. Aruncarea repetată cu banul

Această secţiune are drept scop de a da o idee despre validarea practică a teoriei probabilităţilor. Pentru aceasta vom discuta următorul experiment: un număr de 26 de studenţi au aruncat câte o monedă, fiecare de câte 100 de ori. Rezultatele au fost notate la fiecare aruncare, punându-se un 0 dacă a ieşit stema şi un 1 dacă a ieşit cifra. Pentru fiecare serie de 100 de aruncări au fost numărate 0−urile şi redăm mai jos rezultatele obţinute în cele 26 de serii:

$$53, 58, 52, 47, 48, 54, 49, 43, 46, 43, 47, 51, 44, 52, 61, 50,$$

$$49, 50, 52, 61, 50, 48, 47, 49, 50.$$

S-a constatat că în fiecare serie de 100 de aruncări numărul de 0−uri a fost cuprins între 43 şi 61, iar majoritatea seriilor aveau numărul 0−urilor cuprins între 46 şi 54.

Apoi au fost puse toate seriile în prelungire şi s-a format o serie lungă, corespunzând celor 2600 de aruncări. Pentru această serie s-au numărat secvenţele maximale constante, pe care le vom numi, prescurtat, în continuare *secvenţe constante*. Ca să lămurim termenul de secvenţă constantă vom presupune că am notat cu $x_1, x_2, ..., x_{2600}$ semnele consemnate în timpul experimentului. Facem convenţia că $x_0 = 1 - x_1$ şi definim o *secvenţă constantă* ca fiind un grup de semne consecutive de tipul

$$x_{k-1} \neq x_k = x_{k+1} = ... = x_{k+l-1} \neq x_{k+l}$$

astfel încât $k \geq 1, k + l \leq 2600$. Un astfel de grup va fi numit secvenţă constantă de lungime l. După cum se vede, partea propriu-zis constantă a secvenţei este formată din l elemente consecutive, iar în capete apare câte un alt element care marchează schimbarea de semn. Ultimul element, x_{2600}, nu poate face parte din partea constantă a unei secvenţe constante. De fapt, grupul ultimelor semne care sunt la rând egale cu x_{2600} nu constituie o secvenţă constantă, în accepţiunea pe care o dăm noi. Alegerea făcută pentru x_0 face ca primele semne egale cu x_1 să reprezinte partea constantă a unei secvenţe constante. Vom explica mai târziu de ce am exclus ultima parte de la numărătoare.

Au fost numărate în total 1318 secvenţe constante. În tabelul următor redăm pe rândul de jos numărul secvenţelor găsite, de diverse lungimi:

lung.	1	2	3	4	5	6	7	≥ 8
nr.sv.	674	332	153	77	37	29	6	10

Se observă că numerele acestea scad aproximativ ca o progresie geometrică cu raţia $\frac{1}{2}$. Singur, numărul secvenţelor de constanţă de lungime 6, care este 29, nu este destul de apropiat de jumătatea celui anterior în tabel. Mai observăm că numărul 1318 este apropiat de jumătatea numărului total de aruncări.

Pentru a explica teoretic aceste constatări începem prin a calcula valorile repartiţiei binomiale cu $p = \frac{1}{2}$ şi cu ordinul $n = 100$. Acestea sunt puse în tabelele de mai jos, unde am notat $p_k = C_{100}^k \left(\frac{1}{2}\right)^{100}$.

p_{50}	p_{51}	p_{52}	p_{53}	p_{54}	p_{55}	p_{56}	p_{57}	p_{58}	p_{59}
$0,079$	$0,078$	$0,073$	$0,066$	$0,057$	$0,048$	$0,038$	$0,030$	$0,022$	$0,015$

p_{60}	p_{61}	p_{62}	p_{63}	p_{64}	p_{65}
$0,010$	$0,007$	$0,004$	$0,003$	$0,002$	$0,001$

Repartiţia este simetrică faţă de $k = 50$, astfel că $p_{50+l} = p_{50-l}$. Calculăm $\sum_{l=40}^{60} p_l \approx 0,9648$ şi $\sum_{l=35}^{65} p_l \approx 0,9982$. Aceste valori arată că probabilitatea de a se obţine un număr de $0-$uri cuprins între 40 şi 60, din 100 de aruncări, este covârşitoare, iar probabilitatea de a obţine între 35 şi 65 este practic 1. Cu aceasta am explicat de ce în fiecare serie de 100 de aruncări numărul de $0-$uri şi numărul de $1-$uri sunt echilibrate.

Mai departe ne ocupăm de modelarea secvenţelor constante. Pe un spaţiu probabilizat $(\Omega, \mathcal{F}, P)$ considerăm $A_1, A_2, ...$, un şir de evenimente independente cu aceeaşi probabilitate $P(A_i) = \frac{1}{2}$. Evenimentul A_i are semnificaţia „la cea de a $i-$a aruncare cu banul a fost notat un 1". Pentru analiza noastră teoretică, aşa cum se va vedea, este absolut necesar să considerăm un şir infinit de evenimente. Notăm $X_i = 1_{A_i}, i = 1, 2...$, obţinând un şir de variabile independente repartizate Bernoulli cu parametrul $p = \frac{1}{2}$. Notăm $T_0 = 1$,

$$T_n(\omega) = \inf\left\{j > T_{n-1}(\omega) / X_j(\omega) \neq X_{T_{n-1}}(\omega)\right\}$$

şi $\tau_n = T_n - T_{n-1}, n \in \mathbf{N}^*$. Pentru fiecare $\omega \in \Omega$, secvenţele constante ale şirului $X_1(\omega), X_2(\omega), ...$ sunt descrise sub forma

$$X_{T_{n-1}-1}(\omega) \neq X_{T_{n-1}}(\omega) = ... = X_{T_n-1}(\omega) \neq X_{T_n}(\omega),$$

cu $n \in \mathbf{N}^*$ şi cu convenţia $X_0 = 1 - X_1$. Lungimea acestei secvenţe este $\tau_n(\omega)$.

Lema 5.4. *Variabilele $\tau_1 - 1, \tau_2 - 1, ...$ sunt independente, identic repartizate cu repartiţia geometrică de parametru $p = \frac{1}{2}$.*

Demonstraţie. Adaptăm raţionamentul făcut în demonstraţia *lemei* 5.3. Cu $k_1, ..., k_n \in \mathbf{N}$, are loc egalitatea

$$\{\tau_1 = 1 + k_1, ..., \tau_n = 1 + k_n\} =$$

$$= \{X_1 = X_2 = ... = X_{1+k_1} \neq X_{1+k_1+1} = X_{1+k_1+2}... =$$

$$= X_{1+k_1+1+k_2} \neq X_{1+k_1+1+k_2+1} = ... = X_{1+k_1+...+1+k_n} \neq X_{1+k_1+...+1+k_n+1}\}.$$

Pentru a calcula probabilitatea acestei mulţimi va trebui să despărţim analiza după valoarea primei variabile. Avem

$$\{\tau_1 = 1 + k_1, ..., \tau_n = 1 + k_n, X_1 = 0\} =$$

$$\{X_1 = 0, X_2 = 0, ..., X_{1+k_1} = 0, X_{1+k_1+1} = 1, X_{1+k_1+2} = 1, ..., X_{1+k_1+1+k_2} = 1,$$

$$X_{1+k_1+1+k_2+1} = 0, ..., X_{1+k_1+...+1+k_n} = \epsilon, X_{1+k_1+...+1+k_n+1} = 1 - \epsilon\},$$

unde am pus ϵ pentru a desemna 0, dacă n este impar şi 1, dacă n este par. Ţinând cont de independenţa variabilelor $X_1, X_2, ...$ rezultă

$$P(\tau_1 = 1 + k_1, ..., \tau_n = 1 + k_n, X_1 = 0) = \left(\frac{1}{2}\right)^{k_1+...+k_n+n+1}.$$

În mod similar se deduce că

$$P(\tau_1 = 1 + k_1, ..., \tau_n = 1 + k_n, X_1 = 1) = \left(\frac{1}{2}\right)^{k_1+...+k_n+n+1},$$

ceea ce implică

$$P(\tau_1 = 1 + k_1, ..., \tau_n = 1 + k_n,) = \left(\frac{1}{2}\right)^{k_1+...+k_n+n}. \quad (*)$$

Făcând $n = 1$ în această relaţie se deduce că $P(\tau_1 = 1 + k) = \left(\frac{1}{2}\right)^{k+1}, k \in \mathbf{N}$. Apoi scriem aceeaşi relaţie cu $n = 2$ pentru a obţine

$$P(\tau_2 = 1 + k) = \sum_{j=0}^{\infty} P(\tau_1 = 1 + j, \tau_2 = 1 + k) = \sum_{j=0}^{\infty} \left(\frac{1}{2}\right)^{j+k+2} = \left(\frac{1}{2}\right)^{k+1}.$$

În acelaşi mod se deduce că fiecare din variabilele $\tau_n - 1, n \in \mathbf{N}^*$, este repartizată geometric cu parametrul $\frac{1}{2}$. Atunci relaţia (*) devine

$$P(\tau_1 - 1 = k_1, ..., \tau_n - 1 = k_n) = P(\tau_1 - 1 = k_1) ... P(\tau_n - 1 = k_n),$$

ceea ce încheie demonstraţia. □

Corolarul 5.1. *Variabila aleatoare $T_n - n - 1$ este repartizată negativ- binomial cu parametrii n şi $p = \frac{1}{2}$.*

Demonstraţie. Se ţine cont de *lema* 5.3 şi de faptul că $T_n = T_0 + \tau_1 + ... + \tau_n$.□

Mai departe, vom defini, pentru fiecare $l \in \mathbf{N}^*$, variabila aleatoare $Y_l(\omega) = \sup\{n/T_n \leq l\}$, care reprezintă numărul de secvenţe de constanţă care au fost terminate până la momentul l. Nu se iau în considerare secvenţele de constanţă care se încheie, prin schimbarea valorii, după l.

Lema 5.5. *Variabila aleatoare Y_l este repartizată binomial cu parametrul $p = \frac{1}{2}$ şi ordinul $l - 1$.*

Demonstraţie. Pentru $k = 0, 1, ..., l$ are loc egalitatea

$$\{Y_l = k\} = \{T_k \leq l, T_{k+1} > l\} = \{T_k \leq l, \tau_{k+1} > l - T_k\} =$$
$$= \bigcup_{j=0}^{l-k-1} \{T_k = k + 1 + j, \tau_{k+1} > l - j - 1\}.$$

În această ultimă exprimare am luat toate valorile posibile pentru T_k : cea mai mică este $k + 1$. De aici se deduce că

$$P(Y_l = k) = \sum_{j=0}^{l-k-1} P(T_k = k + 1 + j) P(\tau_{k+1} > l - j - 1) =$$

$$= \sum_{j=0}^{l-k-1} C_{j+k-1}^{k-1} \left(\frac{1}{2}\right)^j \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{l-j-1} = \left(\frac{1}{2}\right)^{l+k-1} \sum_{j=0}^{l-k-1} C_{j+k-1}^{k-1} =$$
$$= \left(\frac{1}{2}\right)^{l+k-1} C_{l-1}^{k}.$$

Cu aceasta am încheiat demonstraţia. □

Faptul că Y_l este repartizată binomial este neaşteptat. O repartiţie binomială este de obicei produsă de numărarea succeselor într-o schemă Bernoulli finită. Dar aici nu este vorba despre schema iniţială, $A_1, A_2, ..., A_l$, cea care corespunde aruncărilor succesive cu banul. Se poate demonstra că este vorba de evenimentele $B_i = A_i \Delta A_{i+1}, i = 1, ..., l-1$, care sunt şi ele independente.

Pentru experimentul nostru, ne interesează variabila Y_{2600}. Găsim că repartiţia binomială cu parametrul $p = \frac{1}{2}$ şi ordinul 2599 are media $1299,5$ şi determinăm $P(1250 \leq Y_{2599} \leq 1350) \approx 0,95$, prin utilizarea calculatorului. Se observă că valoarea de 1318 secvenţe constante, pe care am consemnat-o, este în intervalul $(1250, 1350)$.

Pentru a ne explica de ce datele din tabelul cu secvenţele constante sunt în progresie geometrică vom raţiona astfel. Cele 1318 secvenţe constante, prin lungimile lor, ne furnizează realizări ale variabilelor aleatoare independente $\tau_1, \tau_2, ..., \tau_{1318}$. Variabila τ_i este repartizată astfel

$$P(\tau_i = 1) = \frac{1}{2}, P(\tau_i = 2) = \frac{1}{2^2}, P(\tau_i = 3) = \frac{1}{2^3}, ...$$

Observăm că probabilităţile acestea sunt în progresie geometrică cu raţia $\frac{1}{2}$ şi intuiţia ne spune că atunci este normal ca proporţia în care s-au realizat valorile lui τ_i să fie la fel. O măsurare mai precisă a concordanţei realizărilor concrete cu repartiţia teoretică se face cu testul χ^2, dar acest subiect depăşeşte nivelul prezentului curs.

7. Repartiţia Poisson

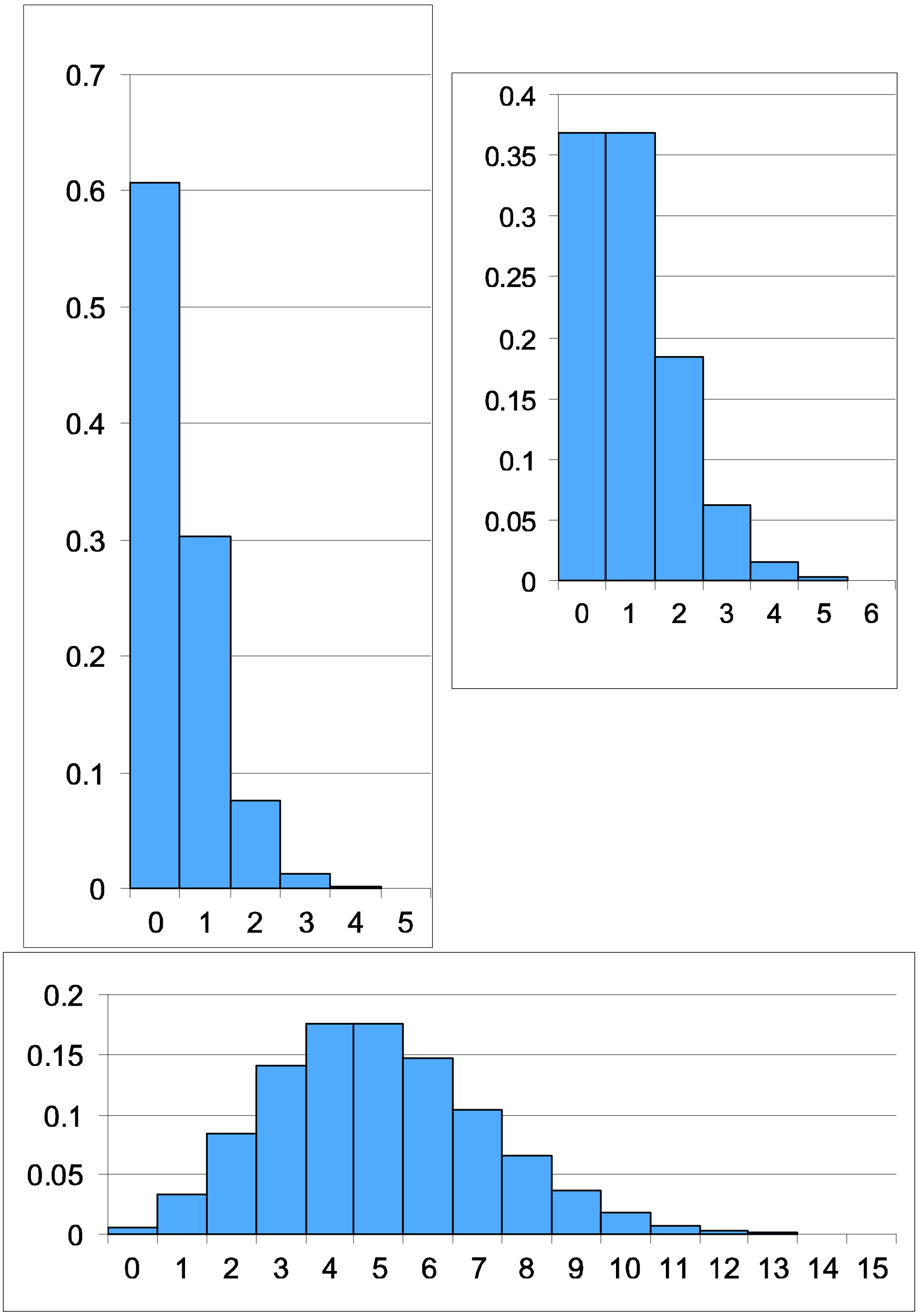
Repartiţia Poisson poartă numele matematicianului francez Denis Poisson (1781-1840). *Repartiţia Poisson* de parametru $\lambda > 0$ este o repartiţie pe $\mathbf{N}$, definită prin relaţia

$$\mu = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} \delta_k.$$

O variabilă alearoare X (definită pe un spaţiu probabilizat $(\Omega, \mathcal{F}, P)$) este repartizată Poisson cu parametru $\lambda > 0$, dacă

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}.$$

Vom stabili mai jos câteva din proprietăţile mai importante ale variabilelor repartizate Poisson. Un calcul direct arată imediat că $EX = \lambda$ şi $D^2 X = \lambda$:



Repartiţii Poisson cu parametrii $\lambda = 0,5$ (stânga sus) $\lambda = 1$ (dreapta sus) şi $\lambda = 5$ (jos)

$$EX = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} k = e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} = e^{-\lambda} \lambda \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \lambda,$$

$$EX^2 = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} k^2 = e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} k = e^{-\lambda} \lambda \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} (k+1) =$$

$$= e^{-\lambda} \lambda \left(\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} + \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} k \right) = \lambda + \lambda^2,$$

$$D^2 X = EX^2 - (EX)^2 = \lambda.$$

Propoziţia 5.3. *Fie $X_1, ..., X_n$ variabile aleatoare independente repartizate Poisson cu parametrii respectivi $\lambda_1, ..., \lambda_n > 0$. Atunci suma lor $X_1 + ... + X_n$ este repartizată tot Poisson, cu parametrul $\lambda = \lambda_1 + ... + \lambda_n$.*

Demonstraţie. Este suficient să tratăm cazul a două variabile, pentru că apoi rezultă şi cazul general printr-o inducţie simplă. Pentru două variabile putem scrie

$$\{X_1 + X_2 = k\} = \bigcup_{i=0}^{k} \{X_1 = i\} \cap \{X_2 = k - i\},$$

$$P(X_1 + X_2 = k) = \sum_{i=0}^{k} P(X_1 = i) P(X_2 = k - i) =$$

$$= \sum_{i=0}^{k} e^{-\lambda_1} \frac{\lambda_1^i}{i!} e^{-\lambda_2} \frac{\lambda_2^{k-i}}{(k-i)!} = e^{-\lambda_1} e^{-\lambda_2} \frac{1}{k!} \sum_{i=0}^{k} C_k^i \lambda_1^i \lambda_2^{k-i} =$$

$$= e^{-(\lambda_1 + \lambda_2)} \frac{1}{k!} (\lambda_1 + \lambda_2)^k.$$

Pentru a trata cazul general se face o inducţie după n. Presupunem demonstrată propoziţia pentru sume de $n-1$ variabile. Fiind date n variabile, scriem $X_1 + ... + X_n = (X_1 + ... + X_{n-1}) + X_n$ şi ţinem cont că variabilele $X_1 + ... + X_{n-1}$ şi X_n, sunt independente datorită *teoremei* 4.1. □

Repartiţia Poisson este legată de repartiţia binomială printr-o teoremă de aproximare asemănătoare teoremei de Moivre-Laplace ce va fi demonstrată în capitolul următor. Fie $(X_i)_{i \in \mathbf{N}}$ un şir de variabile aleatoare repartizate binomial cu parametrii p_i, n_i, care corespund fiecărui indice $i \in \mathbf{N}$. Notăm $q_i = 1 - p_i$ şi presupunem că sunt îndeplinite următoarele condiţii

$$\lim_{i \to \infty} n_i = \infty,$$

$$\lim_{i \to \infty} n_i p_i = \lambda,$$

unde $\lambda > 0$ este un număr fixat.

Teorema 5.1. *În condiţiile de mai sus repartiţiile P_{X_i} converg la repartiţia Poisson de parametru λ, în sensul următor*

$$\lim_{i\to\infty} P\left(X_i = k\right) = e^{-\lambda}\frac{\lambda^k}{k!}\ , \quad k \in \mathbf{N}.$$

Demonstraţie. Ştim că

$$P\left(X_i = k\right) = C_{n_i}^k p_i^k q_i^{n_i - k} = \frac{n_i\left(n_i - 1\right)...\left(n_i - k + 1\right)}{k!} p_i^k \left(1 - p_i\right)^{n_i - k} =$$

$$= \frac{1}{k!}\left(n_i p_i\right)^k \left(1 - \frac{1}{n_i}\right)...\left(1 - \frac{k-1}{n_i}\right)\left(1 - p_i\right)^{\frac{1}{p_i}\left(p_i n_i - p_i k\right)}.$$

Vom trece acum la limită în acest produs. Pentru k fixat, produsul cu $k-1$ factori are limita

$$\lim_{i\to\infty}\left(1 - \frac{1}{n_i}\right)...\left(1 - \frac{k-1}{n_i}\right) = 1.$$

Apoi, pentru ultimul factor observăm că expresia din paranteza de la exponent are limita:

$$\lim_{i\to\infty}\left(p_i n_i - p_i k\right) = \lambda.$$

Deci limita ultimului factor este

$$\lim_{i\to\infty}\left(1 - p_i\right)^{\frac{1}{p_i}\left(p_i n_i - p_i k\right)} = e^{-\lambda}.$$

Toate aceste observaţii conduc la rezultatul anunţat. □

În baza unor estimări numerice s-a ajuns la concluzia că în practică se poate utiliza aproximarea repartiţiei binomiale prin repartiţia Poisson dacă $n \geq 20$ şi $p \leq 0,05$. Aproximarea este excelentă dacă $n \geq 100$ şi $np \leq 10$.

Exemplul 1.

În procesul de fabricaţie al unui obiect apar produse defecte cu frecvenţa de $\frac{1}{100}$. Ne punem problema de a determina care este probabilitatea de a găsi cinci sau mai multe produse defecte la testarea a 200 de obiecte.

Soluţie. Vom nota cu X o variabilă aleatoare repartizată binomial cu parametrii $p = 0,01, n = 200$, care modelează numărul de produse defecte ce apar la 200 de teste. Repartiţia lui X poate fi aproximată cu o repartiţie Poisson de parametru $\lambda = 2$. Prin calcul direct avem $P\left(X \leq 4\right) \approx e^{-2}\sum_{k=0}^{4}\frac{2^k}{k!} = 0,94735$ şi prin urmare răspunsul la problema pusă este $P\left(X \geq 5\right) \approx 0,053$. □

Exemplul 2.

Repartiţia Poisson este cunoscută ca fiind cea care reprezintă evenimentele rare. Pentru a explica acest aspect vom reproduce mai jos una din primele consemnări în date reale a fenomenului. Este vorba despre următorul tabel, alcătuit de Ladislaus von Bortkiewicz şi care reprezintă datele despre accidente mortale petrecute la 10 corpuri de cavalerie prusacă în decursul a 20 de ani.

nr de accidente	*nr de aparitii*	*număr teoretic*
0	109	$108,7$
1	65	$66,3$
2	22	$20,2$
3	3	$4,1$
4	1	$0,6$
	$---$	$---$
	200	$199,9$

Pe prima coloană sunt înşirate cinci tipuri de rezultate ce au fost consemnate pentru fiecare din cele 10 corpuri în fiecare din cei 20 de ani: 0 accidente, 1 accident, 2 accidente, 3 accidente, 4 accidente. Pe coloana a doua se menţionează de câte ori a fost consemnat acel rezultat. Numărul total de rezultate consemnate este de $10 \times 20 = 200$. Numărul total de accidente este, conform tabelului, $65 + 22 \times 2 + 3 \times 3 + 4 = 122$. Numărul mediu de accidente pe an şi corp este $\frac{122}{200} = 0,61$. Bortkiewicz a calculat valorile repartiţiei Poisson cu media $\lambda = 0,61$ ($p_k = e^{-0,61}\frac{(0,61)^k}{k!}, k = 0,1,2,3,4$) şi le-a multiplicat cu 200, obţinând valorile din ultima coloană. Aceste valori se observă a fi foarte apropiate de valorile din coloana a doua. Pentru a explica mai precis această apropire vom observa că probabilitatea ca la o zi de antrenament cu calul să se producă un accident este un număr foarte mic p. Pentru a ne face o idee despre valoarea sa să presupunem că fiecare corp avea un număr de 300 de călăreţi şi că se efectuau cam 250 de antrenamente pe an. Atunci numărul total de antrenamente care au avut loc în 20 de ani la cele 10 corpuri a fost în jur de $300 \times 250 \times 20 \times 10 = 15.000.000$. Cele 122 de accidente conduc la $p = \frac{122}{15.000.000} \approx 8,1 \times 10^{-6}$. Să privim acum ce se petrece la un singur corp, într-un an: repetarea de 300×250 ori a antrenamentului conduce la modelarea evenimentului accident cu o variabilă aleatoare repartizată binomial cu $n = 75.000$ şi $p = 8,1 \times 10^{-6}$. Dar această repartiţie este bine aproximată de repartiţia Poisson cu media

$$\lambda = np = 300 \times 250 \times \frac{122}{300 \times 250 \times 20 \times 10} = \frac{122}{200} = 0,61.$$

Deci evenimentul, în fond rar, al accidentului, prin repetarea de un număr mare de ori a experimentului de antrenare, face ca la un an de zile numărul de accidente să apară modelat de o variabilă repartizată Poisson cu media $\lambda = 0,61$. Cele 200 de rezultate consemnate apar ca 200 de realizări ale unor variabile aleatoare independente, repartizate Poisson, cu parametrul $0,61$.

Dar este clar că în mod excepţional exemplul acesta are datele foarte apropiate de valorile teoretice. Putem presupune că descoperitorul său a ales din mai multe pe cel mai apropiat de idealul teoretic. □

Exemplul 3.

Un magazin de piese auto se aprovizionează pentru următorul trimestru cu un număr de n caroserii ale unui anumit model de autoturism. Vânzarea unei caroserii aduce 40 euro profit net magazinului. În caz că o caroserie nu este vândută în cursul unui trimestru, cheltuielile de depozitare conduc la o pierdere de 5 euro. Vânzarea caroseriilor corespunde evenimentelor rare, care sunt accidentele. Ea este modelată de o variabilă aleatoare repartizată Poisson şi, pentru trimestrul ce urmeză, se estimează că va avea media 10. Să se determine stocul optim, ce maximizează profitul mediu.

Soluţie. Presupunem că pe spaţiul probabilizat $(\Omega, \mathcal{F}, P)$ avem variabila aleatoare X repartizată Poisson cu media 10, care reprezintă cererea. Atunci profitul, depinzând de X şi de stocul n, va fi tot o variabilă aleatoare, care se explicitează astfel:

$$Y_n = \begin{cases} 40X - 5(n - X) & pe\ \{X < n\} \\ 40n & pe\ \{X \geq n\} \end{cases} .$$

Aceasta poate fi scrisă şi sub forma

$$Y_n = 45X \cdot 1_{\{X<n\}} - 45n \cdot 1_{\{X<n\}} + 40n.$$

Pentru a estima cum variază profitul în raport cu stocul vom calcula diferenţa

$$Y_{n+1} - Y_n = 45X \cdot 1_{\{X=n\}} - 45 \cdot 1_{\{X<n\}} - 45(n+1) \cdot 1_{\{X=n\}} + 40 =$$

$$= -45 \cdot 1_{\{X \leq n\}} + 40 .$$

Notăm cu $f(n) = EY_n$ media profitului şi atunci avem

$$f(n+1) - f(n) = EY_{n+1} - EY_n = -45P(X \leq n) + 40 =$$

$$= 45\left(\frac{40}{45} - P(X \leq n)\right) = 45(0,8888 - P(X \leq n)).$$

Este clar că expresia $P(X \leq n)$ creşte cu n. Pentru a putea determina semnul diferenţei $f(n+1) - f(n)$ se află cu calculatorul valorile

$$P(X \leq 13) = 0,8644 ,$$
$$P(X \leq 14) = 0,9165 .$$

Rezultă că avem $f(n+1) > f(n)$, pentru $n \leq 13$ şi $f(n+1) < f(n)$, pentru $n \geq 14$. Deci valoarea maximă este obţinută pentru $n_0 = 14$, care este stocul optim. Putem determina această valoare maximă, ba chiar vom face calculul general pentru profitul mediu:

$$f(n) = EY_n = 45E\left[X \cdot 1_{\{X<n\}}\right] - 45nP(X < n) + 40n.$$

Pe de altă parte putem explicita

$$\begin{aligned} E\left[X \cdot 1_{\{X<n\}}\right] &= \sum_{k=0}^{n-1} kP(X=k) = e^{-\lambda} \sum_{k=0}^{n-1} k \frac{\lambda^k}{k!} = \\ &= \lambda e^{-\lambda} \sum_{k=0}^{n-2} \frac{\lambda^k}{k!} = \lambda P(X \leq n-2), \end{aligned}$$

ceea ce ne permite să scriem

$$f(n) = 450P(X \leq n-2) - 45nP(X \leq n-1) + 40n.$$

Cu $n = 14$ rezultă $f(14) = 371,60$. Pentru comparaţie, calculăm profitul mediu în cazul că se face aprovizionarea doar cu $n = 10$ caroserii, care reprezintă numărul mediu estimat al cererii. Avem $f(10) = 344,60$.

(Am utilizat următoarele valori scoase din calculator: $P(X \leq 12) = 0,7915$; $P(X \leq 9) = 0,4276$; $P(X \leq 8) = 0,3025$) $\square$

Exemplu 4.

O companie de asigurări încheie contracte de asigurare pentru accidente de circulaţie cu coliziune frontală cu despăgubire de 2000 euro. Se ştie că pentru o perioadă de 4 ani media aşteptată a unui astfel de accident este de $0,05$. Cheltuielile companiei pentru un contract sunt de 50 euro în 4 ani. a) Ştiind că are 100 de contracte, iar fiecare client plăteşte 50 de euro anual, să se determine probabilitatea ca respectiva companie să fie pusă în situaţia de a plăti mai mult decât sumele încasate în 4 ani. b) Aceeaşi problemă în cazul în care compania ar avea 1000 de contracte şi clientul ar plăti 45 de euro anual.

Soluţie. Accidentele pe care le poate avea un autoturism în decursul a 4 ani sunt modelate de o variabilă aleatoare repartizată Poisson cu media $\lambda = 0,05$.

a) Pentru 100 de contracte accidentele posibile sunt modelate de variabila $Y = \sum_{i=1}^{100} X_i$, unde X_i ar reprezenta accidentele ce le poate avea autoturismul cu contractul i. Variabilele $X_i, i = 1, ..., 100$, trebuiesc presupuse independente. Rezultă că Y este repartizată Poisson cu media $EY = 100 \times 0,05 = 5$. Situaţia financiară este următoarea: se încasează de la clienţi un total de $50 \times 4 \times 100 = 20.000$ euro. Cheltuielile cu toate dosarele pentru întreaga perioadă sunt $50 \times 100 = 5.000$ euro. Rămâne o sumă disponibilă de $20.000 - 5.000 = 15.000$ euro. Cu această sumă pot fi despăgubiţi 7 clienţi. Cel de al 8 -lea client care suferă accident pune compania în pierdere. Ne interesează deci să calculăm probabilitatea $P(Y > 7) = 0,1333$, după cum se găseşte uşor într-un program de calculator.

b) În mod similar, variabila care modelează accidentele posibile este $Z = \sum_{i=1}^{1000} X_i$ şi are media $EZ = 1000 \times 0,05 = 50$. Situaţia financiară este următoarea: se încasează $45 \times 4 \times 1000 = 180.000$ euro. Se cheltuiesc $50 \times 1000 = 50.000$ euro şi rămân $180.000 - 50.000 =$

130.000 euro. Cu această sumă pot fi despăgubiţi 65 de clienţi. Cel de al 66 -lea client ce are accident aduce compania în pierdere. Probabilitatea ce ne interesează este $P(Z > 65) = 0,0173$, după cum se află cu un program de calculator

După cum se vede, situaţia este total diferită în cazul b). Dacă numărul clienţilor se înmulţeşte cu 10 nu se face pur şi simplu o înmulţire cu 10 a tuturor datelor. Explicaţia stă în proprietăţile repartiţiei Poisson. Deviaţia standard a unei repartiţii Poisson cu media λ este $\sqrt{\lambda}$, iar deviaţia standard măsoară cât de împrăştiate sunt valorile variabilei aleatoare faţă de media sa.

Repartiţia Poisson cu media 5 are deviaţia standard $DY = \sqrt{D^2Y} = \sqrt{5} = 2,24$ în timp ce repartiţia Poisson cu media 50 are deviaţia standard $DZ = \sqrt{D^2Z} = \sqrt{50} = 7,07$. În cazul lui Y deviaţia standard este mare relativ la medie, este aproape jumătate din medie. În cazul lui Z deviaţia standard este mică relativ la media sa. Fenomenul aleator joacă un rol mai mic în problema respectivă. $\square$

8. Repartiţia multinomială*

Repartiţia multinomială este o generalizare a repartiţiei binomiale. Pentru a pune în evidenţă acest lucru vom arunca mai întâi o nouă privire asupra cazului binomial. Pe spaţiul probabilizat $(\Omega, \mathcal{F}, P)$ avem n variabile aleatoare independente, identic repartizate, $X_1, ..., X_n$, care iau valori într-o mulţime cu două elemente $\{0, 1\}$. Aceste variabile iau valoarea 1 cu probabilitatea p şi valoarea 0 cu probabilitatea $q = 1 - p$. Variabila sumă $X_1 + ... + X_n$ ne spune de câte ori apare 1 în şirul dat de variabile. Ea este repartizată binomial.

Dar putem privi lucrurile şi altfel. Considerând simetric cele două rezultate posibile pe care le indică variabilele, vom nota $p_1 = p$ şi $p_2 = q$, astfel că repartiţia comună a variabilelor X_i este $p_1\delta_1 + p_2\delta_0$. Vom mai schimba notaţia şi vom pune $Y_1 = X_1 + ... + X_n$ şi $Y_2 = n - Y_1$. Variabila Y_2 ne spune de câte ori apare valoarea 0 în şirul nostru. Dacă variabila Y_1 are repartiţia binomială cu ordinul n şi parametrul p_1, variabila Y_2 este complet similară, fiind repartizată binomial cu ordinul n şi parametrul p_2. Dacă $k_1, k_2 \in \mathbf{N}$ satisfac condiţia $k_1 + k_2 = n$, atunci are loc egalitatea $\{Y_1 = k_1\} = \{Y_2 = k_2\}$ şi putem spune că repartiţia binomială caracterizează simultan cele două variabile prin relaţia

$$P(Y_1 = k_1, Y_2 = k_2) = \frac{n!}{k_1!k_2!}.$$

De aici urmează generalizarea. Fie $p_1, ..., p_m \in [0, 1]$ astfel ca $p_1 + ... + p_m = 1$. Să presupunem acum că variabilele $X_1, ..., X_n$ sunt independente, identic repartizate, definite pe un spaţiu probabilizat $(\Omega, \mathcal{F}, P)$, cu valori într-o mulţime $E = \{e_1, ..., e_m\}$, astfel că repartiţia comună acestor variabile este măsura $p_1\delta_{e_1} + ... + p_m\delta_{e_m}$. Cu alte cuvinte, variabila X_l ia veloarea e_i cu probabilitatea p_i şi mai explicit, avem

$P(X_l = e_i) = p_i$, oricare ar fi $l = 1, ..., n$ şi $i = 1, ..., m$. Atunci variabilele

$$Y_i = \sum_{l=1}^{n} 1_{X_l^{-1}(e_i)}, i = 1, ..., m,$$

ne spun de câte ori apare rezultatul e_i în şirul $X_1, ..., X_n$. Are loc relaţia $Y_1 + ... + Y_m = n$, astfel că pentru orice numere $k_1, ..., k_m \in \mathbf{N}$ ce satisfac relaţia $k_1 + ... + k_m = n$ avem

$$\{Y_1 = k_1, ..., Y_{m-1} = k_{m-1}\} = \{Y_1 = k_1, ...Y_{m-1} = k_{m-1}, Y_m = k_m\}$$

Vectorul $Y = (Y_1, ..., Y_n)$ este cel care are repartiţia ce ne interesează şi pe care o descriem în lema ce urmează.

Lema 5.6. *Variabila* $(Y_1, ..., Y_m)$ *are repartiţia dată de relaţia*

$$P(Y_1 = k_1, ..., Y_m = k_m) = \frac{n!}{k_1!...k_m!} p_1^{k_1}...p_m^{k_m},$$

valabilă pentru orice sistem de numere $k_1, ..., k_m \in \mathbf{N}$ *astfel încât* $k_1 + ... + k_m = n$.

Demonstraţie. Fie $\omega \in \Omega$ fixat. Dacă $Y_i(\omega) = k_i$, înseamnă că mulţimea $M_i = \{l / X_l(\omega) = e_i\}$ are k_i elmente. Pentru $i = 1, ..., m$ avem m mulţimi care partiţionează mulţimea de indici ale variabilelor X_l,

$$\{1, ..., n\} = \bigcup_{i=1}^{m} M_i.$$

Lăsând acum ω liber, putem scrie

$$\{Y_1 = k_1, ..., Y_m = k_m\} = \bigcup_{(M_1, ..., M_m)} \{X_l = e_i, \forall l \in M_i, i = 1, ..., m\},$$

unde reuniunea se face după toate partiţiile numerotate $(M_1, ..., M_m)$ ale mulţimii $\{1, ..., n\}$ astfel încât $card(M_i) = k_i, i = 1, ..., m$. Atunci putem calcula

$$P(Y_1 = k_1, ..., Y_m = k_m) = \sum_{(M_1, ..., M_m)} P(X_l = e_i, \forall l \in M_i, i = 1, ..., m) =$$

$$= \sum_{(M_1, ..., M_m)} p_1^{k_1}...p_m^{k_m} = \frac{n!}{k_1!...k_m!} p_1^{k_1}...p_m^{k_m},$$

ţinând cont de *propoziţia* 1.2. Aceasta demonstrează lema. □

Repartiţia vectorului Y este numită repartiţia multinomială. Este o repartiţie definită de parametrii $n, m \in \mathbf{N}, n \geq 1, m \geq 2$, şi $p_1, ..., p_m \in (0, 1)$ care trebuie să satisfacă relaţia $p_1 + ... + p_m = 1$ şi este suportată de mulţimea

$$K = \{(k_1, ..., k_m) \in \mathbf{N}^m / k_1 + ... + k_m = n\} \subset \mathbf{N}^m.$$

Expresia repartiţiei multinomiale este dată de

$$\sum_{(k_1,...,k_m)} \frac{n!}{k_1!...k_m!} p_1^{k_1}...p^{k_m} \delta_{(k_1,...,k_m)} \ ,$$

unde sumarea se face după toate punctele din suportul K.

Exemplele de repartiţii multinomiale apar în două formulări tipice:

a) Într-o urnă se află bile colorate în culorile $e_1, ..., e_m$. Numărul bilelor de culoarea e_i este proporţional cu p_i. Se fac n extrageri cu întoarcere. Extragerile sunt modelate de variabilele $X_l : \Omega \to \{e_1, ..., e_m\}, l = 1, ..., n$, care indică ce culoare are bila extrasă la a $i-$a extragere. Atunci probabilitatea de a fi extrase k_i bile de culoarea $e_i, i = 1, ..., m$, este reprezentată de

$$P(Y_1 = k_1, ..., Y_m = k_m) = \frac{n!}{k_1!...k_m!} p_1^{k_1}...p_m^{k_m}.$$

b) Se aruncă în mod aleatoriu n bile în m cutii notate $e_1, ..., e_m$, aşa încât probabilitatea de a arunca o bilă în cutia i este p_i. Dacă aruncarea unei bile este modelată de o variabilă X_i, alocarea bilelor în cele m cutii este descrisă de repartiţia lui $Y = (Y_1, ..., Y_m)$, repartiţia multinomială.

9. Împrăştierea aleatoare*

O serie de fenomene fizice sunt descrise printr-o mulţime finită de puncte răspândite aleator într-un "spaţiu " E, spaţiu care la rândul său este o submulţime, de obicei continuă (în sensul de nediscretă şi de puterea continului), dintr-un spaţiu euclidian. Spaţiul euclidian respectiv poate avea orice dimensiune, inclusiv unu.

Exemple de astfel de fenomene pot fi următoarele: a) Punctele de impact ale picăturilor de ploaie care cad într-o secundă pe o suprafaţă determinată. b) Punctele de impact ale meteoriţilor căzuţi pe pământ. c) Microbii care sunt împrăştiaţi într-un volum de aer contaminat. d)Momentele de timp la care trec autovehiculele prin dreptul unui reper pe şosea. Aici spaţiul euclidian este unu-dimensional şi reprezintă timpul. e) Momentele de timp la care sosesc apelurile la o centrală telefonică. f) Momentele de timp la care sunt emise particule de către o cantitate dată de substanţă radioactivă. g) Momentele la care soseşte autobuzul în staţie, etc. Toate aceste fenomene sunt modelate de obiectul matematic numit *măsură aleatoare Poisson.*

Fie $(E, \mathcal{E})$ un spaţiu măsurabil şi $\lambda : \mathcal{E} \to \mathbf{R}_+$ o măsură finită pe $(E, \mathcal{E})$. Notăm cu Ω mulţimea părţilor finite din E,

$$\Omega = \{F \in \mathcal{P}(E) \mid card(F) < \infty\},$$

şi pentru fiecare mulţime $A \in \mathcal{E}$ notăm $X_A : \Omega \to \mathbf{R}_+$, aplicaţia de numărare

$$X_A(F) = card(A \cap F) \quad , \quad F \in \Omega.$$

Se observă că dacă $A, B \in \mathcal{E}$ şi $A \subset B$ atunci $X_A \leq X_B$, iar dacă $A \cap B = \emptyset$, atunci $X_{A \cup B} = X_A + X_B$.

Pe mulţimea Ω definim $\mathcal{F}$ drept cea mai mică σ- algebră care face toate aplicaţiile X_A măsurabile:

$$\mathcal{F} = \sigma\left(X_A \mid A \in \mathcal{E}\right).$$

Definiţia 5.1. *O măsură de probabilitate P pe $(\Omega, \mathcal{F})$ va fi numită* măsură aleatoare Poisson *cu intensitate λ, dacă sunt verificate următoarele proprietăţi:*

1^0 Pentru orice $A \in \mathcal{E}$, variabila X_A este repartizată Poisson cu parametrul $\lambda(A)$.

2^0 Dacă $A_1, ..., A_n \in \mathcal{E}$ sunt mulţimi disjuncte, atunci variabilele $X_{A_1}, ..., X_{A_n}$ sunt independente.

În această definiţie admitem implicit că o variabilă aleatoare nulă a.s. este repartizată Poisson degenerat, cu parametrul 0. Un eveniment elementar $F \in \Omega$ trebuie văzut ca o configuraţie finită de puncte împrăştiate în spaţiul E. Măsura aleatoare Poisson descrie probabilistic împrăştierea aleatoare.

Rezultă imediat din definiţie că $EX_A = \lambda(A)$ pentru orice mulţime $A \in \mathcal{E}$. Mai notăm că în cazul a două mulţimi $A, B \in \mathcal{F}$ astfel încât $A \subset B$ şi $\lambda(A) = \lambda(B)$, rezultă că $X_A = X_B$, aproape sigur.

Măsurile aleatoare Poisson sunt utile atât în tratarea aplicaţiilor cât şi în studiul teoretic al unor probleme legate de procesele Markov. Enunţăm mai jos teorema generală de construcţie a unei măsuri aleatoare Poisson cu intensitate dată. Demonstraţia se bazează pe unele idei legate de procesele Poisson şi nu îşi are locul cel mai bun aici.

Teorema 5.2. *Fiind dat un spaţiu măsurabil $(E, \mathcal{E})$ şi o măsură difuză şi finită, λ, pe el, există şi este unică o măsură de probabilitate P pe $(\Omega, \mathcal{F})$ care este o măsură aleatoare Poisson cu intensitatea λ.*

Noţiunea de măsură difuză este următoarea. O măsură λ pe $(E, \mathcal{E})$ se numeşte difuză dacă are proprietatea că pentru orice mulţime $A \in \mathcal{E}$ astfel ca $\lambda(A) > 0$ şi orice număr $0 < b < \lambda(A)$, există o submulţime $B \subset A, B \in \mathcal{E}$, astfel încât $\lambda(B) = b$.

În continuare vom prezenta alte fapte legate de măsurile aleatoare Poisson păstrând notaţiile introduse mai sus. Vom avea nevoie de următoarea lemă tehnică.

Lema 5.7. *Fie $(X_n)_{n \in \mathbf{N}}$ un şir crescător de variabile aleatoare repartizate Poisson (definite pe un spaţiu probabilizat $(\Omega, \mathcal{F}, P)$).*

(i) Dacă $\sup\limits_n EX_n < \infty$ atunci limita şirului $X = \lim\limits_n X_n$ este o variabilă aleatoare repartizată Poisson cu parametrul $\sup\limits_n EX_n$.

(ii) Dacă $\sup\limits_n EX_n = \infty$, atunci $\lim\limits_n X_n = \infty$, aproape sigur.

(iii) În cazul în care (X_n) este descrescător, limita este sau nulă sau tot o variabilă repartizată Poisson, cu parametrul $\inf_n EX_n$.

Demonstraţie. Vom nota $\lambda_n = EX_n$ şi observăm că şirul $(\lambda_n)_{n\in\mathbf{N}}$ este crescător cu limita ce o notăm $\lambda = \lim \lambda_n$.

(i) Pentru fiecare $k \in \mathbf{N}$, şirul de mulţimi $\{X_n \leq k\}, n \in \mathbf{N}$ este descrescător şi avem $\cap_n \{X_n \leq k\} = \{X \leq k\}$. De aceea putem calcula limita

$$P(X \leq k) = \lim_{n\to\infty} P(X_n \leq k) = \lim_{n\to\infty} \sum_{l=0}^{k} e^{-\lambda_n} \frac{\lambda_n^l}{l!} = \sum_{l=0}^{k} e^{-\lambda} \frac{\lambda^l}{l!}, \quad (*)$$

de unde rezultă că X este repartizată Poisson cu parametrul λ.

(ii) Relaţia (*) de mai înainte arată că în cazul în care $\lim_{n\to\infty} \lambda_n = \infty$, rezultă $P(X \leq k) = 0$, pentru orice k. Aceasta demonstrează afirmaţia din enunţ.

(iii) În cazul în care şirul de variabile este descrescător rezultă că şirul de mulţimi $\{X_n \leq k\}, n \in \mathbf{N}$, este crescător. Din nou este valabilă relaţia $(*)$ şi se raţionează similar cazurilor dinainte. Dacă $\lim_{n\to\infty} \lambda_n = 0$, se obţine $X = 0$, aproape sigur. $\square$

Următoarea teoremă ajută la modelarea unor exemple concrete. Anume, faptul că anumite fenomene sunt descrise de o măsură aleatoare Poisson este greu de acceptat pentru că impune apriori repartiţia Poisson. Teorema de mai jos arată că de fapt nu este necesar să presupunem că variabilele X_A sunt repartizate Poisson. Această condiţie este implicată de alte condiţii mai slabe şi mai uşor de acceptat. Deşi este enunţată în cazul unidimensional, teorema se poate generaliza şi la cazul unor spaţii E multidimensionale. Pentru cazul real, se poate arăta că noţiunea de măsură difuză revine la o măsură ce nu încarcă punctele. De exemplu, orice măsură ce se reprezintă cu o densitate faţă de măsura Lebesgue este difuză.

Teorema 5.3. *În notaţia anterioară, presupunem că $E \subset \mathbf{R}$, este un interval deschis, $\mathcal{E} = \mathcal{B}(E)$, iar $\lambda : \mathcal{B}(E) \to [0, \infty]$ este o măsură difuză finită. Presupunem că P este o probabilitate pe $(\Omega, \mathcal{F})$ cu următoarele proprietăţi:*

1^0 Pentru orice interval mărginit $I \subset \mathbf{R}$ are loc relaţia $EX_I = \lambda(I)$,

2^0 Pentru orice două intervale mărginite $I, J \subset E$, astfel încât $\lambda(I) = \lambda(J)$ rezultă $P(X_I = 0) = P(X_J = 0)$.

3^0 Dacă $I_1, ..., I_n \subset E$, sunt intervale disjuncte, atunci variabilele $X_{I_1}, ..., X_{I_n}$ sunt independente.

În aceste condiţii $(\Omega, \mathcal{F}, P)$ este măsura aleatoare Poisson cu intensitatea λ.

Demonstraţie. Începem prin a studia mulţimea E' care constă din toate punctele $x \in E$ care au câte o vecinătate V ce este interval deschis, $x \in V \subset E$ şi $\lambda(V) = 0$. Se observă din această definiţie că E' este o mulţime deschisă şi, prin urmare, ea se scrie ca o reuniune $E' = \bigcup_{l\in\Lambda} J_l$, unde $J_l, j \in \Lambda$, sunt intervale deschise disjuncte, iar Λ este cel mult numărabilă. Se observă că $\lambda(J_l)$, pentru orice $l \in \Lambda$ şi, cum măsura este difuză, şi închiderea acestor intervale este de măsură nulă: $\lambda\left(\overline{J_l}\right) = 0$, pentru orice $l \in \Lambda$. Notăm $E'' = \bigcup_{l\in\Lambda} \overline{J_l}$ şi vom avea $\lambda(E'') = 0$. O proprietate ce decurge imediat din această definiţie, şi pe care o vom utiliza mai jos, este următoarea: pentru orice două puncte $x, y \in E \setminus E''$, intervalul determinat de ele are măsura strict pozitivă, adică $\lambda((x, y)) > 0$.

Notăm acum cu $\Omega_0 = \{F \in \Omega / \ F \cap E'' = \emptyset\}$. Deoarece $E'' \in \mathcal{E}$ şi pentru că are loc scrierea

$$\Omega_0 = \{F \in \Omega / \ X_{E''}(F) = 0\}$$

rezultă că $\Omega_0 \in \mathcal{F}$. Pe de altă parte, deoarece $EX_{E''} = \lambda(E'') = 0$ şi din exprimarea complementarei sub forma

$$\Omega_0^c = \{F \in \Omega / \ X_{E''}(F) \geq 1\},$$

rezultă că $P(\Omega_0^c) = 0$, deci $P(\Omega_0) = 1$.

Următorul pas va fi de a arăta că pentru un interval de forma $I = [a, b)$ cu $a, b \in E, X_I$ este o variabilă aleatoare repartizată Poisson cu parametrul $\lambda(I)$. Dacă $\lambda(I) = 0$, atunci $EX_I = 0$ şi, prin urmare, avem $X_I = 0, a.s.$, ceea ce încheie demonstraţia. Pentru cazul $\lambda(I) > 0$ demonstraţia este mai lungă. Notăm $\alpha = \lambda(I)$ şi împărţim intervalul I în 2^n intervale consecutive, de măsură egală. În acest fel introducem următoarea notaţie: $x_0^n = a$ şi luăm punctul $x_1^n \in (a, b)$ astfel încât $\lambda([x_0^n, x_1^n)) = \frac{\alpha}{2^n}$. Prin recurenţă se construieşte sistemul de puncte $a = x_0^n < x_1^n < x_2^n < ... < x_{2^n}^n = b$, astfel încât intervalele $I_i^n = [x_{i-1}^n, x_i^n)$ să aibă fiecare măsura $\lambda(I_i^n) = \frac{\alpha}{2^n}$.

Notăm $Y_i^n = 1_{\left\{X_{I_i^n} \geq 1\right\}}, i = 1, ..., 2^n$ şi $Y^n = \sum\limits_{i=1}^{n} Y_i^n$.

a) Vom arăta acum că şirul (Y^n) este crescător. Mai întâi se observă că diviziunea

$$\left\{x_0^{n+1} < x_1^{n+1} < x_2^{n+1} < ... < x_{2^{n+1}}^{n+1}\right\}$$

rafinează diviziunea $\{x_0^n < x_1^n < x_2^n < ... < x_{2^n}^n\}$. Mai precis, avem $x_2^{n+1} = x_1^n$ şi $I_1^n = [x_0^n, x_1^n) = [x_0^{n+1}, x_1^{n+1}) \cup [x_1^{n+1}, x_2^{n+1}) = I_1^{n+1} \cup I_2^{n+1}$. În general, $x_{2i}^{n+1} = x_i^n$şi $I_i^n = [x_{i-1}^n, x_i^n) = [x_{2i-2}^{n+1}, x_{2i-1}^{n+1}) \cup [x_{2i-1}^{n+1}, x_{2i}^{n+1}) = I_{2i-1}^{n+1} \cup I_{2i}^{n+1}$, pentru orice indice $i = 1, ..., 2^n$. Aceasta implică

$$X_{I_i^n} = X_{I_{2i-1}^{n+1}} + X_{I_{2i}^{n+1}}$$

şi deci $Y_i^n \leq Y_{2i-1}^{n+1} + Y_{2i}^{n+1}$, care conduce la $Y^n \leq Y^{n+1}$.

b) În continuare vom verifica relaţia $\lim\limits_{n\to\infty} Y^n = X_I$ pe Ω_0. Vom arăta convergenţa $Y^n(F) \to X_i(F)$ pentru orice element $F \in \Omega_0$. Să presupunem deci că $F \subset E \setminus E''$ şi că intersecţia $F \cap I$ se compune din m puncte

$$F \cap I = \{y_1, ..., y_m\},$$

puncte ce le vom presupune ordonate $y_1 < y_2 < ... < y_m$. Pentru a raţiona în continuare, măsurăm distanţa dintre două puncte cu ajutorul măsurii λ, prin măsura lungimii intervalului determinat de cele două puncte. Astfel, numărul $d = \min\{\lambda([y_j, y_{j+1})) \mid j = 1, ..., m-1\}$, exprimă distanţa minimă dintre două puncte din $F \cap I$. El este strict pozitiv pentru că punctele sunt din $E \setminus E''$. Dacă n este astfel încât $\frac{\alpha}{2^n} < d$, atunci un interval de tipul I_i^n nu poate conţine mai mult de un punct din F. Rezultă că $X_{I_i^n}(F) \leq 1$ şi deci $X_{I_i^n}(F) = Y_i^n(F)$. Cum această egalitate are loc pentru orice $i = 1, ..., 2^n$, rezultă

$$X_I(F) = \sum_{i=1}^{2^n} X_{I_i^n}(F) = \sum_{i=1}^{2^n} Y_i^n(F) = Y^n(F).$$

În concluzie, şirul $Y^n(F), n \in \mathbf{N}^*$, este staţionar şi $\lim_{n\to\infty} Y^n(F) = X_I(F)$.

Vom utiliza acum faptele demonstrate la a) şi b) de mai sus pentru a aplica *teorema* 5.1 şirului (Y^n). Începem prin a observa că, pentru fiecare n fixat, mulţimile $\left\{X_{I_i^n} \geq 1\right\}$, $i = 1, ..., 2^n$ sunt independente, datorită ipotezei 3^0 şi numărul $p_n = P\left(X_{I_i^n} \geq 1\right)$ nu depinde de i, datorită ipotezei 2^0. Rezultă că variabilele $\{Y_i^n \mid i = 1, ..., 2^n\}$ sunt independente, repartizate Bernoulli cu parametrul p_n şi, prin urmare, Y^n este repartizată binomial cu parametrul p_n şi rangul 2^n. Avem $EY^n = 2^n p_n$ şi, datorită punctului a) de mai sus, şirul acestor numere este crescător. Cum $Y^n \leq X_I$, rezultă $2^n p_n \leq EX_I = \lambda(I)$. În orice caz există limita $\lim_{n\to\infty} 2^n p_n = \gamma \leq \lambda(I)$. *Teorema* 5.1 ne asigură că

$$\lim_{n\to\infty} P(Y^n = k) = e^{-\gamma}\frac{\gamma^k}{k!}.$$

Pe de altă parte, are loc relaţia

$$\{X_I \leq k\} \cap \Omega_0 = \bigcap_{n\in\mathbf{N}^*} \{Y^n \leq k\} \cap \Omega_0,$$

în care şirul de mulţimi din dreapta este descrescător, pentru fiecare $k \in \mathbf{N}$ fixat. Rezultă

$$P(X_I \leq k) = \lim_{n\to\infty} P(Y^n \leq k) = \lim_{n\to\infty} \sum_{l=0}^{k} P(Y^n = l) = \sum_{l=0}^{k} e^{-\gamma}\frac{\gamma^l}{l!}.$$

Se deduce că X_I este repartizată Poisson cu parametrul γ, şi atunci $\gamma = \lambda(I)$.

Deoarece λ nu încarcă punctele, rezultă că intervalele $I' = (a, b)$ şi $I'' = [a, b]$ dau variabile de numărare egale:

$$X_{I'} = X_I = X_{I''} \quad , \quad a.s.$$

Fie acum o mulţime deschisă $D \subset E$. Ea se scrie ca o reuniune cel mult numărabilă de intervale deschise disjuncte $D = \bigcup_n I_n$ şi are loc relaţia

$$X_D = \sum_n X_{I_n}.$$

Dacă $\lambda(D) < \infty$, rezultă că X_D este o variabilă repartizată Poisson cu parametrul $\lambda(D)$, prin aplicarea *propoziţiei* 5.3 şi a punctului (i) din *lema* 5.7. În plus, dacă $D_1, ..., D_l$ sunt mulţimi deschise disjuncte din E, atunci, utilizând ipoteza 3^0 şi descompunerile în intervale ale acestor mulţimi, se ajunge la concluzia că $X_{D_1}, ..., X_{D_l}$ sunt independente.

Să considerăm acum cazul unei mulţimi compacte $K \subset \mathbf{R}$. Notăm $D_n = \left\{x \in E | d(x, K) < \frac{1}{n}\right\}$ vecinătatea de rază $\frac{1}{n}$ a lui K, pentru orice $n \in \mathbf{N}^*$. Acestea sunt mulţimi deschise mărginite şi $K = \bigcap_n D_n$. Rezultă că

$$X_K = \lim_{n\to\infty} X_{D_n}.$$

Din lema anterioară, rezultă că X_K este repartizată Poisson şi parametrul său este $\lambda(K)$. În plus, dacă $K_1, ..., K_l$ sunt mulţimi compacte disjuncte, atunci vecinătăţile lor de rază $\frac{1}{n}$ sunt disjuncte când n este mare. Rezultă că $X_{K_1}, ..., X_{K_l}$ se aproximează cu variabile independente. La limită şi $X_{K_1}, ..., X_{K_l}$ se obţin independente.

Mai departe, trecem la cazul unei mulţimi boreliene $A \in \mathcal{B}(E)$. Pentru aceasta avem nevoie de un rezultat mai fin de teoria măsurii. Acesta spune că pentru orice mulţime boreliană se pot construi două şiruri $(K_n)_{n\in\mathbf{N}}$, $(D_n)_{n\in\mathbf{N}}$ astfel încât $K_n \subset K_{n+1} \subset A \subset D_{n+1} \subset D_n$, pentru orice $n \in \mathbf{N}$ şi

$$\lambda(A) = \lim_{n\to\infty} \lambda(K_n) = \lim_{n\to\infty} \lambda(D_n),$$

mulţimile $K_n, n \in \mathbf{N}$, fiind compacte, iar mulţimile $D_n, n \in \mathbf{N}$, fiind deschise. Notând $A' = \bigcup_n K_n$ şi $A'' = \bigcap_n D_n$, vom avea

$$X_{A'} \le X_A \le X_{A''}.$$

Deoarece $\lambda\left(A'\right) = \lambda\left(A''\right)$, rezultă că $X_{A'} = X_A = X_{A''}$, aproape sigur. Rezultă că X_A este repartizată Poisson cu parametrul $\lambda(A)$. Ca mai sus, prin aproximare, se deduce şi faptul că $X_{A_1}, ..., X_{A_l}$ sunt variabile independente dacă $A_1, ..., A_l$ sunt boreliene, disjuncte. □

Observaţia 5.1. *O ipoteză care implică punctele 1^0 şi 2^0 din enunţul de mai sus este următoarea:*

4^0 Dacă $\lambda(I) = \lambda(J)$, atunci variabilele X_I şi X_J sunt identic repartizate.

În multe aplicaţii este uşor de acceptat direct această ipoteză, care altfel este o concluzie a teoremei.

Picăturile de ploaie.

Atunci când se examinează modelarea picăturilor de ploaie, este, la o primă vedere, greu de acceptat că măsura aleatoare Poisson este obiectul matematic care descrie fenomenul. În primul rând obişnuinţa comună ne îndeamnă să privim cantitatea de picături care cade pe un metru pătrat ca fiind o constantă. Dar este clar că nu este aşa. Faptul că la măsurători făcute în două locuri diferite, în aceeaşi perioadă de timp, se obţin cantităţi de apă foarte apropiate, nu înseamnă că fenomenul aleator nu este prezent. O variabilă aleatoare repartizată Poisson cu parametrul $\lambda = 10.000$ (corespunzător la 10.000 de picături) are deviaţia standard 100. Deci este normal ca variaţiile să fie de acest ordin. Ele apar ca nesemnificative din punct de vedere practic, dar fenomenul este aleator şi poate corespunde unei măsuri aleatoare Poisson. În cazul a 1.000.000 de picături variaţia dată de model este şi mai mică, relativ; ea este de 1.000 de picături. Faptul că numărul de picături ce cade pe fiecare metru pătrat este aleator poate fi observat atunci când se consideră o perioadă scurtă de timp, sau în cazul unei ploi foarte scurte de vară. Independenţa numărului de picături care cad în locuri diferite pe suprafeţe egale este naturală. Teorema dinainte ne conduce la acceptarea modelului dat de o măsură aleatoare Poisson cu intensitatea măsura Lebesgue multiplicată de o constantă ce corespunde amploarei fenomenului. □

Emisia de particule radioactive.

În cazul emisiunii de particule, fenomen care este modelat ca o împrăştiere aleatoare de puncte pe axa timpului, ilustrăm buna modelare cu cifrele din următorul tabel care conţine datele consemnate de Rutherford şi Geiger.

k	*nr. observatii*	*nr. teoretic*
0	57	54
1	203	211
2	383	407
3	525	526
4	532	508
5	408	394
6	273	254
7	139	140
8	45	68
9	27	29
10	10	11
≥ 11	6	6

Cei doi fizicieni au măsurat numărul de particule α emise pe minut de o sursă de polonium. Anume au făcut măsurători pentru un număr de 2608 minute. În prima coloană sunt trecute cele 12 tipuri de rezultate consemnate pentru fiecare din cele 2608 minute: 0 particule emise, 1 particulă, 2 particule, ... În coloana a doua este trecut numărul de minute pentru care a fost consemnat tipul de rezultat din coloana întâi. Numărul total de particule emise a fost

$$203 + 2 \times 383 + 3 \times 525 + 4 \times 532 + 5 \times 408 + 6 \times 273 + 7 \times 139 + \\ + 8 \times 45 + 9 \times 27 + 10 \times 10 + 11 \times 6 = 10.092$$

Media numărului de particule emise pe minut rezultă $\lambda = \frac{10.092}{2608} = 3,86$. Valorile repartiţiei Poisson cu media $3,86$ multiplicate cu numărul de minute măsurate, 2608, sunt trecute pe coloana a treia. Se observă o foarte mare apropiere între rezultatele din coloanele doi şi trei. □

Rata de defectare.

O legătură interesantă există între noţiunea de rată de defectare care apare în fiabilitate, sau rata de deces din actuariat, şi măsura aleatoare Poisson. Fie $\beta : [0, \infty) \to \mathbf{R}_+$ o funcţie continuă. Notăm λ măsura ce are β drept densitate: $\lambda(dx) = \beta(x)\, dx$. Fie P măsura aleatoare Poisson pe $(\mathbf{R}_+, \mathcal{B}(\mathbf{R}_+))$ cu intensitatea λ. Definim apoi T ca poziţia primului punct (sau momentul primului punct, dacă privim $\mathbf{R}_+$ ca reprezentând timpul) :

$$T(F) = \inf F = \inf \left\{ t \geq 0 \mid X_{[0,t]}(F) > 0 \right\} \quad , \quad F \in \Omega,$$
$$T : \Omega \to [0, \infty],$$

cu convenţia $T(\emptyset) = \infty$. Atunci β este rata de defectare corespunzătoare lui T.

De fapt, în enunţul de mai sus este suficient să presupunem că β este o funcţie măsurabilă şi local integrabilă, adică $\int_0^x \beta(t)\, dt < \infty$, pentru orice $x > 0$. Pentru a evita discuţii delicate de analiză, presupunem β o funcţie continuă.

Fiind o variabilă aleatoare pozitivă, putem interpreta T ca un timp de funcţionare. Din relaţia

$$\{T > t\} = \left\{ X_{[0,t]} = 0 \right\},$$

rezultă $P(T > t) = e^{\lambda([0,t])}$. Prin derivare obţinem densitatea de repartiţie

$$\frac{d}{dt} P(T > t) = -\beta(t)\, e^{-\int_0^t \beta(s) ds}.$$

Deci β are interpretarea ratei de defectare asociate lui T.

În cazul particular în care β este o constantă, rezultă că repartiţia primului punct, sau repartiţia lui T este exponenţială de parametru β.□

10. Exerciţii

Exerciţiul 5.1. *Un motor de un anumit tip este utilizat atât pentru avioane bimotoare cât şi pentru cvadrimotoare. Presupunem că probabilitatea ca un astfel de motor să se defecteze după o anumită perioadă* *de funcţionare este d. Pe de altă parte, normele de siguranţă obligă un bimotor care are defect unul din motoare să întrerupă zborul. Pentru un cvadrimotor regula este ca dacă are mai puţin de 3 motoare în funcţiune să întrerupă călătoria. Să se determine probabilităţile: 1) ca un bimotor să funcţioneze respectiva perioadă fără probleme, 2) ca un cvadrimotor să funcţioneze aceeaşi perioadă cu succes. 3) Pentru ce valori ale lui d este mai avantajos un bimotor?*

Exerciţiul 5.2. *O clădire este prevăzută cu două lifturi. Probabilitatea ca un lift să se defecteze într-o zi este de* $0,01$. *Să se determine repartiţia numărului de zile pe an în care este defect primul lift. Să se determine repartiţia numărului de zile pe an în care sunt defecte ambele lifturi, presupunând că acestea se defectează independent unul de celălalt. Să se determine repartiţia numărului de zile pe an în care cel puţin unul din lifturi este defect.*

Exerciţiul 5.3. *Un automat telefonic poate să răspundă la* 90% *din apeluri. Se fac* 10 *apeluri la rând la acest automat. Să se calculeze cu o precizie de* 10^{-4} *probabilitatea de a fi reuşit contactarea de: 1) exact* 10 *ori, 2) exact* 5 *ori, 3) cel puţin o dată.*

Exerciţiul 5.4. *Se constată că o persoană care scrie în stare de oboseală poate să deformeze aproximativ* 20% *din litere. Experţii grafologi analizează un text în care, din* 140 *de litere,* 42 *sunt deformate. Pentru a examina ipoteza unei prefaceri intenţionate a scrisului se calculează probabilitatea ca, în ipoteza stării de oboseală, să se ajungă la un număr de litere deformate mai mare sau egal cu* 42. *Să se afle această probabilitate.*

Exerciţiul 5.5. *O familie îşi planifică copii pe care ar dori să-i aibă şi analizează următoarele scheme:*

(a) lasă să se nască orice copil până ce apare o fată şi atunci se opreşte.

(b) lasă să se nască orice copil până ce vor avea cel puţin doi copii de sexe diferite şi atunci se opreşte.

Presupunând că nu există naşteri de gemeni şi că nou -născutul are şanse egale de a fi băiat sau fată, să se determine media familiei în fiecare caz.

Exerciţiul 5.6. ** Matematicianul Ştefan Banach avea două cutii de chibrite, fiecare cu câte* 50 *de beţe, în acelaşi buzunar. De câte ori a avut nevoie, el a scos câte o cutie la întâmplare şi a utilizat un băţ de chibrit. La un moment dat cutia pe care o scoate din buzunar are un singur băţ şi după utilizare este terminată. Să se construiască un model probabilist cu o variabilă aleatoare care să modeleze numărul de chibrite ce se pot afla în cealălaltă cutie în momentul în care este golită prima cutie. Să se determine repartiţia acestei variabile. Care este numărul mediu de chibrite ce rămân în cea de a doua cutie?*

Exerciţiul 5.7. ** O persoană hotăreşte să arunce cu banul până ce obţine cel puţin două rezultate cu fiecare faţă a monedei. Care este numărul mediu de aruncări pe care ar trebui să le execute persoana respectivă?*

Exerciţiul 5.8. *Care este probabilitatea ca la o extragere loto „6 din 49", la care s-au vândut 2.000.000 bilete, să se obţină rezultatele: a) nu a ieşit nici un bilet câştigător, b) a ieşit un bilet câştigător, c) au ieşit două bilete câştigătoare, d) au ieşit trei bilete câştigătoare. Care este probabilitatea ca să nu iasă câştigător nici un bilet trei extrageri la rând, la care s-au vândut respectiv 2.000.000, 1.500.000 şi 2.500.000 bilete?*

Exerciţiul 5.9. *Un tipograf face în medie o greşeală la 5000 de semne puse în pagină. El realizează pagini de câte 43 de linii, fiecare conţinând câte 70 de semne. Calculaţi probabilitatea ca pe o pagină dată: 1) să fie exact două erori, 2) să fie cel mult două erori.*

Exerciţiul 5.10. *În procesul de fabricaţie al unui obiect apar produse defecte cu probabilitatea de 0,01. Care este probabilitatea de a găsi două sau mai multe produse defecte la testarea a 200 de obiecte?*

Exerciţiul 5.11. *Fie X, Y variabile aleatoare repartizate Poisson cu mediile λ, respectiv μ. Să se calculeze $P(X = k / X + Y = n)$ şi să se arate că X este repartizată binomial sub $P(\cdot / X + Y = n)$.*

Exerciţiul 5.12. *Se ştie că apelurile la o centrală telefonică sosesc aleator, aproximativ uniform în intervalul de 12 ore al unei zile. Într-o anumită zi au sosit 180 de apeluri. Care este probabilitatea ca fixând un interval de 4 ore să constatăm că în el s-au petrecut între 50 şi 70 de apeluri.*

[Se va constata că există două modele plauzibile: 1) Considerăm fiecare din cele 180 de apeluri modelate de câte o variabilă aleatoare repartizată Bernoulli cu $p = \frac{1}{3}\left(= \frac{4}{12}\right)$. 2) Considerăm modelul unei măsuri aleatoare Poisson pe intervalul $I = [0, 12]$, cu intensitatea $a\mathcal{L}$ unde am notat cu $\mathcal{L}$ măsura Lebesgue pe dreaptă. Notăm cu J intervalul de lungime 4 care reprezintă perioada în care suntem interesaţi şi punem $J' = I \backslash J$. Variabilele X_J şi $X_{J'}$ sunt independente şi repartizate Poisson cu parametrii $4a$, respectiv $8a$. Calculând repartiţia lui X_J, condiţionată de $X_J + X_{J'} = 180$ se ajunge la acelaşi rezultat ca prin metoda 1).]

Exerciţiul 5.13. *Un vas conţine apă în care se află bacterii. S-a estimat că numărul de bacterii este atât de mare încât, dacă ar fi uniform repartizate, la fiecare picătură de apă ar reveni 2 bacterii. Presupunând că se spală vasele cu această apă şi după spălare rămâne câte o picătură de apă pe pe fiecare farfurie, după câteva zile fiecare bacterie dă naştere la o colonie. Care este probabilitatea de a avea 3 sau mai multe colonii pe o farfurie?*

Exerciţiul 5.14. *Microbii sunt împrăştiaţi aleator pe o lamelă de laborator cu densitatea de* $5.000/cm^2$*. Câmpul vizibil prin microscop este de* 10^{-4} cm^2*. Care este probabilitatea ca în câmpul vizibil să se afle cel puţin un microb?*

CAPITOLUL 6

Variabile aleatoare şi repartiţii continue

1. Generalităţi

Până acum am avut de a face în special cu spaţii probabilizate discrete, în sensul că σ -algebrele au fost generate de o partiţie cel mult numărabilă. În acest moment suntem obligaţi să discutăm despre repartiţia normală, o repartiţie care îşi face apariţia implicit prin teorema de Moivre -Laplace, care va interveni explicit în teorema limită centrală şi care care ne scoate din cadrul discret studiat în capitolele anterioare. Scopul acestui capitol este de a prezenta un număr minim de fapte absolut necesare pentru a clarifica noţiunea de repartiţie în cazul nediscret. Evităm însă problemele mai complexe de teoria măsurii.

Vom presupune în continuare că este dat un spaţiu probabilizat $(\Omega, \mathcal{F}, P)$. Vom numi *variabilă aleatoare* cu valori în E orice aplicaţie $X : \Omega \to E$, cu valori într-un spaţiu măsurabil $(E, \mathcal{E})$, care are proprietatea că $X^{-1}(A) \in \mathcal{F}$, pentru orice $A \in \mathcal{E}$. Termenul din teoria măsurii, utilizat pentru aceeaşi noţiune, este de aplicaţie măsurabilă de la $\mathcal{F}$ la $\mathcal{E}$. Când se omite atributul "cu valori în E" se subînţelege că $E = \mathbf{R}$, iar $\mathcal{E} = \mathcal{B}(\mathbf{R})$ este σ -algebra mulţimlor boreliene pe $\mathbf{R}$.

Dacă $X : \Omega \to E$ este o aplicaţie arbitrară, se notează $\sigma(X) = \{X^{-1}(A) / A \in \mathcal{E}\}$ familia preimaginilor întoarse de aplicaţie. Se verifică fără probleme, ca şi în cazul discret, că această familie este o σ -algebră şi ea poartă denumirea de σ -algebra generată de X. Faptul de a fi variabilă aleatoare revine deci la verificarea incluziunii $\sigma(X) \subset \mathcal{F}$. Independenţa unor variabile aleatoare $X_1, ..., X_n$ este definită prin independenţa σ -algebrelor $\sigma(X_1), ..., \sigma(X_n)$.

Fiind dată o variabilă aleatoare $X : \Omega \to E$ cu valori în spaţiul măsurabil $(E, \mathcal{E})$, se notează cu P_X şi se numeşte *repartiţia* lui X, la fel ca în cazul discret, măsura de probabilitate definită pe $\mathcal{E}$ prin formula

$$P_X(A) = P\left(X^{-1}(A)\right), \quad A \in \mathcal{E}.$$

Deci repartiţia lui X este măsura obţinută transportând P prin aplicaţia X. Faptul că aceasta este o măsură de probabilitate se verifică direct. De exemplu, proprietatea de σ -aditivitate se verifică astfel: dacă $A_1, A_2, ...$ sunt mulţimi disjuncte din $\mathcal{E}$, atunci $X^{-1}(A_1), X^{-1}(A_2), ...$

sunt mulţimi disjuncte din $\mathcal{F}$ şi are loc relaţia

$$X^{-1}\left(\bigcup_{n=1}^{\infty} A_n\right) = \bigcup_{n=1}^{\infty} X^{-1}(A_n),$$

care implică

$$P_X\left(\bigcup_{n=1}^{\infty} A_n\right) = P\left(X^{-1}\left(\bigcup_{n=1}^{\infty} A_n\right)\right) = P\left(\bigcup_{n=1}^{\infty} X^{-1}(A_n)\right) =$$

$$\sum_{n=1}^{\infty} P\left(X^{-1}(A_n)\right) = \sum_{n=1}^{\infty} P_X(A_n). \square$$

Observăm că orice măsură de probabilitate pe un spaţiu măsurabil $(E, \mathcal{E})$ poate fi privită ca repartiţia unei variabile aleatoare. Întradevăr, dacă μ este o măsură de probabilitate pe $(E, \mathcal{E})$, atunci $(E, \mathcal{E}, \mu)$ este un spaţiu probabilizat. Aplicaţia identitate $X(x) = x$ de la $(E, \mathcal{E}, \mu)$ la $(E, \mathcal{E})$ poate fi considerată ca o variabilă aleatoare cu valori în E. Măsura transportată prin aplicaţia identitate este neschimbată, deci $\mu(A) = \mu(X^{-1}(A))$, pentru orice $A \in \mathcal{E}$, ceea ce înseamnă că repartiţia lui X este μ (adică $P_X = \mu$). Acesta este motivul pentru care orice măsură de probabilitate mai este numită repartiţie.

1.1. Repartiţii ce admit densitate*. În acest capitol vom examina o serie de repartiţii care nu mai sunt suportate de puncte, ci care pun măsura zero pe orice punct, ba chiar vor asocia o măsură nulă pentru orice mulţime cel mult numărabilă. Aceste repartiţii se numesc *continue*. De fapt, repartiţiile ce ne interesează vor fi obţinute din măsura Lebesgue pe dreaptă prin introducerea unei *densităţi*. Concret, să considerăm un interval $(a, b) \subset \mathbf{R}$,finit sau infinit, şi o funcţie $\rho : (a, b) \to \mathbf{R}_+$ continuă pe porţiuni şi cu integrala unu:

$$\int_a^b \rho(x)\, dx = 1.$$

(Faptul că funcţia ρ este continuă pe porţiuni înseamnă că există o partiţie $a = t_0 < t_1 < ... < t_n = b$, astfel încât ρ să fie continuă pe fiecare interval deschis $(t_{i-1}, t_i), i = 1, ..., n.$) Atunci putem defini o măsură μ punând, pentru orice interval $[x_1, x_2] \subset (a, b)$,

$$\mu([x_1, x_2]) = \int_{x_1}^{x_2} \rho(x)\, dx.$$

Este clar că o astfel de măsură nu va încărca punctele, iar pentru intervale deschise sau semideschise vom avea aceeaşi expresie

$$\mu((x_1, x_2)) = \mu((x_1, x_2]) = \mu([x_1, x_2)) = \mu([x_1, x_2]) = \int_{x_1}^{x_2} \rho(x)\, dx.$$

Dacă avem o mulţime A care se scrie ca o reuniune de intervale, sau mai general pentru orice $A \in \mathcal{B}((a,b))$, definim

$$\mu(A) = \int_A \rho(x)\, dx.$$

În acest fel obţinem aplicaţia $\mu : \mathcal{B}((a,b)) \to [0,1]$, care se dovedeşte a fi σ -aditivă, deci o măsură de probabilitate. Nu facem demonstraţia, care rezultă cu argumente standard de teoria măsurii. Se utilizează notaţia $\mu = \rho(x)\, dx$ pentru această măsură.

Un calcul de repartiţie.

Să presupunem că privim tripletul $((a,b), \mathcal{B}((a,b)), \mu)$ de mai sus ca spaţiu probabilizat. Să presupunem că avem o bijecţie $h : (a,b) \to (c,d)$, cu imaginea un alt interval şi astfel încât $h \in \mathcal{C}^1(a,b)$ şi $|h'| > 0$. Vom privi h ca variabilă aleatoare cu valori în $\mathbf{R}$ şi dorim să-i determinăm repartiţia. Mai precis, vrem să arătăm că repartiţia lui h, pe care o notăm cu μ', este dată de densitatea $r(y) = \frac{\rho \circ h^{-1}(y)}{|h'| \circ h^{-1}(y)}, y \in (a,b)$.

Pentru simplificare presupunem $h' > 0$ şi începem prin a scrie formula schimbării de variabilă

$$\int_{y_1}^{y_2} f(y)\, dy = \int_{x_1}^{x_2} f(h(x))\, h'(x)\, dx,$$

unde $a \le x_1 \le x_2 \le b$ şi $h(x_i) = y_i, i = 1, 2$. Deci avem $h^{-1}([y_1, y_2]) = [x_1, x_2]$ şi trebuie să calculăm

$$\mu\left(h^{-1}([y_1,y_2])\right) = \mu([x_1,x_2]) = \int_{x_1}^{x_2} \rho(x)\, dx = \int_{x_1}^{x_2} \rho(x) \frac{h'(x)}{h'(x)} dx =$$

$$= \int_{x_1}^{x_2} \left(\rho \circ h^{-1}\right)(h(x)) \frac{h'(x)}{\left(h' \circ h^{-1}\right)(h(x))} dx = \int_{y_1}^{y_2} \frac{\rho \circ h^{-1}(y)}{h' \circ h^{-1}(y)} dy =$$

$$= \int_{y_1}^{y_2} r(y)\, dy.$$

La penultimul semn egal am folosit formula schimbării de variabilă. Egalitatea stabilită pentru intervale se extinde prin aditivitate la reuniuni de intervale şi cu un argument de clasă monotonă se obţine, pentru orice $A \in \mathcal{B}((c,d))$,

$$\mu'(A) = \mu\left(h^{-1}(A)\right) = \int_A r(y)\, dy. \qquad ((*))$$

Aceasta este relaţia dorită.□

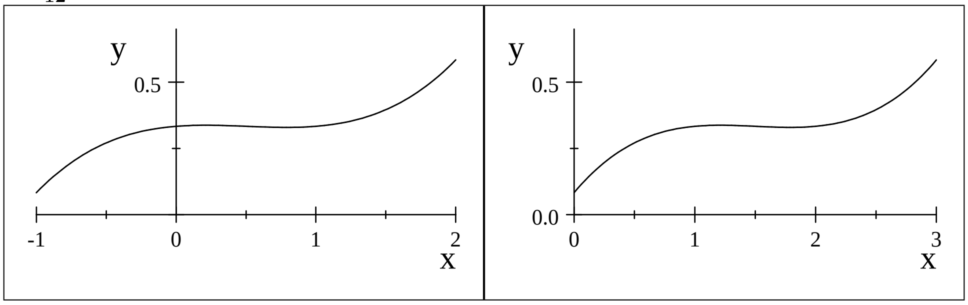
Să presupunem acum că repartiţia cu densitate $\mu = \rho(x)\, dx$ este repartiţia unei variabile X definite pe un alt spaţiu probabilizat $(\Omega, \mathcal{F}, P)$, adică $P_X = \mu = \rho(x)\, dx$. Atunci $Y = h(X)$ este o altă variabilă aleatoare şi repartiţia ei este μ', după cum se verifică uşor:

$$P\left(Y^{-1}(A)\right) = P\left(X^{-1}\left(h^{-1}(A)\right)\right) = \mu\left(h^{-1}(A)\right) = \mu'(A).$$

Translaţia.

Cazul în care h este o translaţie, adică $h(x) = \xi + x$, cu $\xi \in \mathbf{R}$ un număr fix, corespunde unei translaţii a măsurii μ. Intervalul (a, b) este translatat în $(a + \xi, b + \xi)$, iar densitatea ρ este şi ea translatată $r(x) = \rho(x - \xi)$. Graficul lui r este punctual translatat cu ξ.

Redăm mai jos graficul funcţiei $\rho(x) = \frac{1}{12}x^3 - \frac{1}{8}x^2 + \frac{1}{24}x + \frac{1}{3}$ şi graficul translatat cu $\xi = 1$, care corespunde funcţiei $\rho(x-1) = \frac{1}{12}x^3 - \frac{3}{8}x^2 + \frac{13}{24}x + \frac{1}{12}$



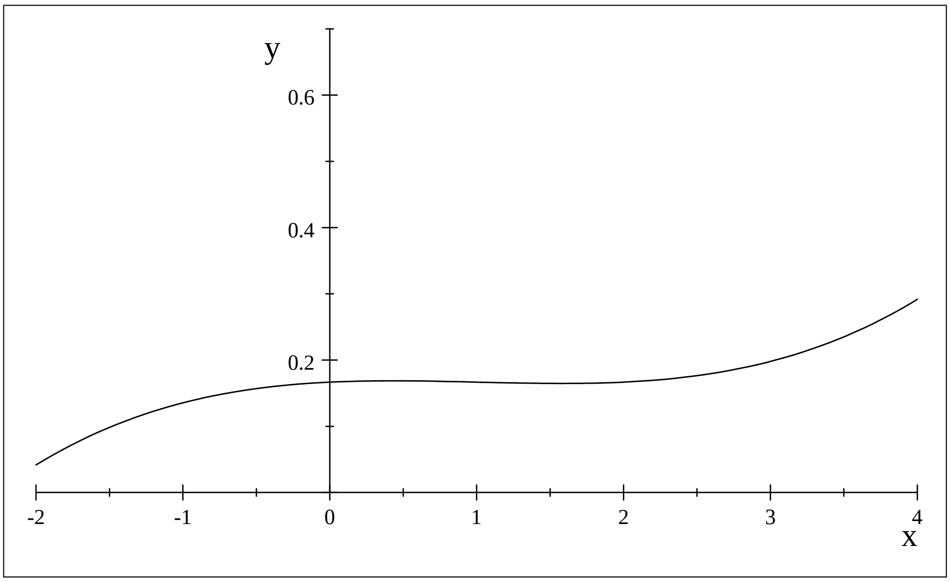
Cazul unei variabile aleatoare adunate cu o constantă. Dacă avem o variabilă aleatoare X, definită pe un spaţiu probabilizat $(\Omega, \mathcal{F}, P)$, a cărui repartiţie este $P_X = \mu = \rho(x)\,dx$, atunci $Y = h(X) = X + \xi$ este o variabilă cu repartiţia translatată.

Omotetia.

Să presupunem acum că h este o omotetie, adică are o expresie de tipul $h(x) = \theta x$, unde $\theta > 0$ este un număr fix. În această situaţie intervalul (a, b) este transportat peste intervalul $(a\theta, b\theta)$, iar densitatea devine $r(x) = \frac{1}{\theta}\rho\left(\frac{x}{\theta}\right)$. Când $\theta > 1$, această transformare revine la întinderea graficului în lung şi, în acelaşi timp, turtirea sa. Când $\theta < 1$, graficul se strânge în lungime şi se înalţă.

În desenul de mai jos avem transformarea graficului lui ρ după o omotetie cu $\theta = 2$. Mai concret este graficul funcţiei $\frac{1}{2}\rho\left(\frac{x}{2}\right) = \frac{1}{192}x^3 - \frac{1}{64}x^2 + \frac{1}{96}x + \frac{1}{6}$.

Cazul unei variabile aleatoare înmulţite cu o constantă. Dacă X este o variabilă aleatoare definită pe un spaţiu probabilizat $(\Omega, \mathcal{F}, P)$ care are repartiţia $P_X = \mu = \rho(x)\,dx$, atunci $Y = \theta X$ are drept repartiţie măsura transportată prin omotetie. Adică $P_Y = \frac{1}{\theta}\rho\left(\frac{x}{\theta}\right)dx$.



Produsul de măsuri cu densitate şi independenţa variabilelor. Să presupunem că pe fiecare din intervalele $(a_i, b_i), i = 1, ..., n,$ avem câte o măsură de probabilitate $\mu_i : \mathcal{B}((a_i, b_i)) \to [0, 1], i = 1, ..., n$. Se demonstrează în teoria măsurii că atunci există o singură măsură de probabilitate pe spaţiul produs $(a_1, b_1) \times ... \times (a_n, b_n)$, să o notăm $\mu : \mathcal{B}((a_1, b_1) \times ... \times (a_n, b_n)) \to [0, 1]$, care are proprietatea că

$$\mu([x_1, y_1] \times ... \times [x_n, y_n]) = \mu_1([x_1, y_1]) ... \mu_1([x_n, y_n]),$$

pentru orice produs de intervale $[x_i, y_i] \subset (a_i, b_i), i = 1, ..., n$. Această măsură se numeşte produsul măsurilor $\mu_1, ..., \mu_n$ şi se notează $\mu = \mu_1 \bigotimes ... \bigotimes \mu_n$. În cazul în care măsurile iniţiale au densităţi, $\mu_i = \rho_i(x)\, dx, i = 1, ..., n$, se dovedeşte că măsura produs are şi ea densitate şi anume aceasta este $\rho(x_1, ..., x_n) = \rho_1(x_1) ... \rho_n(x_n)$. Cu alte cuvinte măsura produs se exprimă printr-o integrală multiplă sub forma

$$\mu(A) = \int ... \int_A \rho_1(x_1) ... \rho_n(x_n)\, dx_1 ... dx_n,$$

pentru orice $A \in \mathcal{B}((a_1, b_1) \times ... \times (a_n, b_n))$.

Independenţa variabilelor aleatoare se caracterizează prin repartiţii la fel ca în cazul discret. Fie $X_1, ..., X_n$, variabile aleatoare reale pe spaţiul probabiliozat $(\Omega, \mathcal{F}, P)$, ale căror repartiţi admit densităţi: $P_{X_i} = \rho_i(x)\, dx, i = 1, ..., n$. Notăm cu $Z = (X_1, ..., X_n) : \Omega \to \mathbf{R}^n$, variabila aleatoare vectorială de componente $X_1, ..., X_n$. Se demonstrează că are loc egalitatea de σ -algebre

$$\sigma(Z) = \sigma(\sigma(X_1), ..., \sigma(X_n)),$$

iar independenţa variabilelor $X_1, ..., X_n$ este echivalentă cu egalitatea

$$P_Z = \rho_1(x_1) ... \rho_n(x_n)\, dx_1 ... dx_n.$$

1.2. Exemple de repartiţii cu densitate*. Repartiţia uniformă.

Fie $[0,a] \subset \mathbf{R}$ un interval mărginit, pe care-l vedem ca spaţiu măsurabil echipat cu σ -algebra mulţimilor boreliene $\mathcal{B}([0,a])$. Definim o măsură de probabilitate pe acest interval, pe care o notăm $\mathcal{U}_a$, punând

$$\mathcal{U}_a(A) = \frac{1}{a}\int_A dx,$$

pentru orice mulţime $A \in \mathcal{B}([0,a])$. Această măsură este numită *repartiţia uniformă* pe $[0,a]$. Putem scrie prescurtat $\mathcal{U}_a = \frac{1}{a}1_{[0,a]}(x)\,dx$.

Repartiţia exponenţială.

Fie $\alpha > 0$. Pe spaţiul măsurabil $((0,\infty), \mathcal{B}((0,\infty)))$ se defineşte măsura de probabilitate $\mathcal{E}_\alpha$ prin

$$\mathcal{E}_\alpha(A) = \alpha\int_A e^{-\alpha x}dx,$$

pentru orice mulţime $A \in \mathcal{B}((0,\infty))$. Această măsură este numită *repartiţia exponenţială* de parametru α. Prescurtat putem scrie $\mathcal{E}_\alpha = \alpha 1_{(0,\infty)}(x)\,dx$.

Exemplu de construcţie de variabilă repartizată exponenţial. Fie $\Omega = (0,1), \mathcal{F} = \mathcal{B}((0,1))$ şi să notăm cu P măsura Lebesgue pe $(0,1)$. Deci $(\Omega, \mathcal{F}, P)$ este un spaţiu probabilizat. Definim $X : \Omega \to (0,\infty)$, prin $X(x) = -\frac{1}{\alpha}\ln x$. Afirmăm că atunci X este o variabilă aleatoare repartizată exponenţial cu parametrul α.

Pentru a verifica acest fapt este suficient să aplicăm formula (*) din secţiunea anterioară. Luăm $h = X$ şi avem $h^{-1}(y) = e^{-\alpha y}$ şi $|h'(x)| = \frac{1}{\alpha x}$. Pe de altă parte, în cazul de faţă avem $\rho \equiv 1$ şi atunci $r(y) = \alpha e^{-\alpha y}$ şi de aceea putem scrie

$$P\left(X^{-1}(A)\right) = \int_A \alpha e^{-\alpha y}dy = \mathcal{E}_\alpha(A).$$

1.3. Paradoxul lui Bertrand*. Problema propusă de Bertrand se enunţă astfel. Intr-un cerc dat se duce la întâmplare o coardă. Să se afle probabilitatea ca acea coardă să aibă lungimea mai mare ca latura triunghiului echilateral înscris. Paradoxul provine din lipsa de precizie a termenului „se duce la întâmplare”. Există mai multe feluri de a duce la întâmplare corzi pe un cerc şi rezultatele sunt diferite pentru diversele procedee.

Vom descrie în continuare trei experimente probabiliste care răspund la această problemă producând fiecare alt rezultat. Primele două se bazează pe corespondenţa bijectivă care există între corzile unui cerc şi punctele din cerc, corespondenţă realizată asociind fiecărei corzi mijlocul său.

1. Presupunem că un cerc mare este desenat în curte pe pământ. După o scurtă ploaie de vară rămân marcate în interiorul cercului un număr mare de puncte de impact ale picăturilor de ploaie. Fiecare din

aceste puncte determină o coardă cu mijlocul în punct. Dacă se face o numărătoare, se constată că proporţia de corzi mai lungi decât latura triunghiului înscris este de $\frac{1}{4}$.

2. Vom descrie acum un alt procedeu aleator de construcţie a corzilor în cerc, utilizând o roată de ruletă cu care producem direcţii aleatoare şi un pachet de şireturi de diverse lungimi care ne dă distanţe aleatoare măsurate în centimetri, repartizate uniform pe intervalul $[0, 50]$.

Presupunem că un cerc cu raza 50 cm este desenat pe o planşă. Aşezăm roata de ruletă în interiorul cercului, concentric şi îi imprimăm o mişcare de rotaţie, după care marcăm direcţia la care s-a oprit. Extragem la întâmplare un şiret din pachet şi, pe direcţia respectivă, marcăm distanţa de la centrul cercului dată de lungimea şiretului. Punctul astfel marcat va fi centrul unei coarde în cerc. Proporţia de corzi, obţinute în acest fel, care vor fi mai lungi decât latura triunghiului înscris în cerc se constată a fi apropiată de valoarea $\frac{1}{2}$.

3. Un alt procedeu de a construi corzi aleatoare în cerc este bazat pe generarea fiecărui capăt al corzii cu o roată de ruletă. Cercul este desenat pe o planşă şi se pune ruleta concentric, în interiorul cercului. Se roteşte ruleta, iar la oprire se marchează pe cerc punctul din direcţia indicatorului roţii de ruletă. Acesta este un capăt de coardă. Se repetă această operaţie pentru a obţine şi cel de al doilea capăt al corzii. Proporţia de corzi cu lungimea mai mare decât latura triunghilui echilateral înscris se constată a fi apropiată de valoarea $\frac{1}{3}$.

Modelare matematică.

Vom modela acum cele trei experimente probabiliste.

1. Pentru primul experiment se impune modelul următor: Ω este discul mărginit de cercul dat; $\mathcal{F} = \mathcal{B}(\Omega)$ este familia mulţimilor boreliene din Ω, iar $P(A) = \frac{ariaA}{aria\Omega}$, pentru orice regiune $A \in \mathcal{F}$. Fiecare coardă este identificată cu mijlocul său, care este un punct din Ω. Dacă r este latura cercului, se ştie că latura triunghiului echilateral înscris în cerc este la distanţa $r/2$ de centrul cercului de rază r. (In figura 1 MNQ este triunghiul echilateral înscris în cerc. Punctul O este centrul cercului iar OE este perpendiculară pe MQ. Se ştie că $OE = EF$.)

Rezultă că dacă alegem un punct G din interiorul cercului de rază $\frac{r}{2}$, concentric cu cel iniţial, şi ducem raza ce trece prin G iar apoi ducem prin G o perpendiculară pe rază, aceasta va determina o coardă a cercului iniţial de lungime mai mare decât latura triunghiului lateral înscris. Pentru punctele din afara cercului de rază $\frac{r}{2}$, acelaşi procedeu conduce la corzi de lungime mai mică decât latura triunghiului echilateral înscris.

Fie Ω_0 mulţimea punctelor din cercul de rază $\frac{r}{2}$ şi concentric cu Ω. Aceasta este mulţimea a cărei probabilitate trebuie să o calculăm. Avem

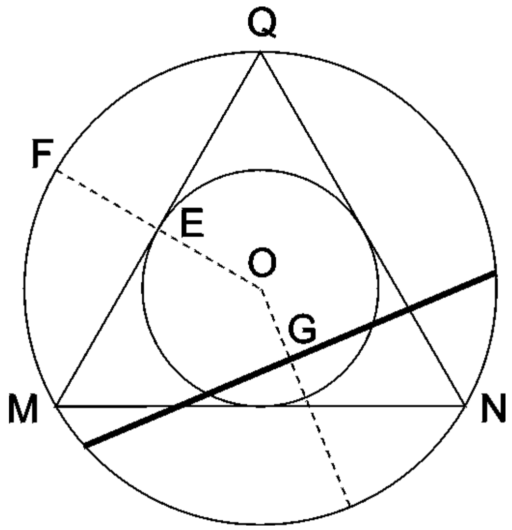


FIGURA 1. Experimentul 1.

$$P\left(\Omega_0\right)=\frac{\left|\Omega_0\right|}{|\Omega|}=\frac{\pi\left(\frac{r}{2}\right)^2}{\pi r^2}=\frac{1}{4},$$

valoare care este în concordanţă cu rezultatul experimental. □

2. Pentru a construi modelul pentru cel de al doilea experiment se porneşte de la faptul că experimentul are două componente aleatoare independente: alegerea direcţiei prin rotirea roatei de ruletă şi alegerea distanţei punctului faţă de centru prin extragerea unui şiret din pachet. Direcţia corzilor, cât şi a perpendicularelor pe acestea, este repartizată uniform în jurul centrului de cerc. La fel, distanţa de la centrul cercului la coardă este şi ea repartizată uniform pe intervalul $(0,50)$.

În figura 2, alăturată, am luat un sistem cartezian cu centrul în O. O coardă în cerc este unic determinată de unghiul $\theta\in[0,2\pi)$ pe care îl face segmentul ce uneşte O de mijlocul corzii G,cu direcţia Ox şi de distanţa $l\in(0,50)$ dintre G şi O.

Atunci se impune ca model dreptunghiul $\Omega=[0,2\pi)\times(0,50)$ cu $\mathcal{F}$ mulţimea părţilor boreliene din Ω. Cele două proiecţii pe componentele lui Ω, $X_1:\Omega\rightarrow[0,2\pi), X_2:\Omega\rightarrow(0,50)$, definite prin $X_1(\theta,l)=\theta, X_2(\theta,l)=l$, dorim să fie independente şi uniform repartizate. Pe intervalele $[0,2\pi)$ şi $(0,50)$ avem măsurile uniforme $\mathcal{U}_{2\pi}$ şi $\mathcal{U}_{50}$. Rezultă că măsura de probabilitate pe Ω o definim prin produsul $P=\mathcal{U}_{2\pi}\bigotimes\mathcal{U}_{50}$. La fel ca în cazul produselor de măsuri discrete, această măsură asigură independenţa celor două proiecţii, şi faptul că repartiţiile lor sunt $\mathcal{U}_{2\pi}$, respectiv, $\mathcal{U}_{50}$. Mulţimea $\Omega_0=[0,2\pi)\times[0,25]$ corespunde punctelor G care sunt mijloace ale corzilor de lungime mai mare ca latura triunghiului echilateral înscris. Probabilitatea acestei

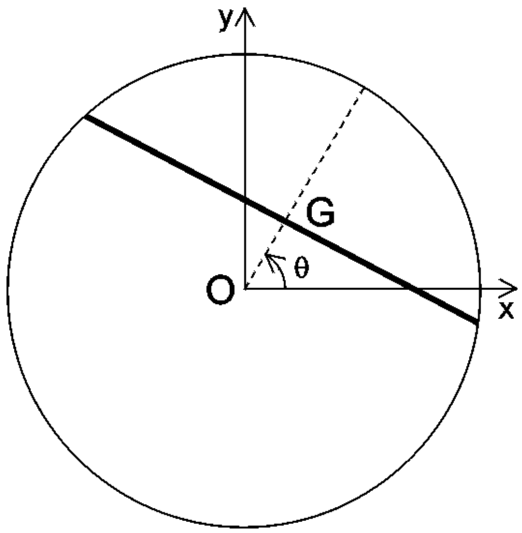


FIGURA 2. Experimentul 2.

mulţimi este

$$P(\Omega_0) = \frac{2\pi}{2\pi} \times \frac{25}{50} = \frac{1}{2},$$

care este în concordanţă cu rezultatul experimental. □

3. De data aceasta privim capetele P_1, P_2 ale corzii. Fiecare punct de pe cerc este determinat de unghiul pe care-l face raza ce trece prin el cu direcţia Ox, măsurat în sensul trigonometric, astfel că pentru model putem lua $\Omega = [0, 2\pi) \times [0, 2\pi)$. Coarda arbitrară de capete P_1, P_2 pe cerc este reprezentată de o pereche $(\theta_1, \theta_2) \in [0, 2\pi) \times [0, 2\pi)$, unde θ_1 este unghiul pe care-l face OP_1 cu Ox, şi θ_2 unghiul dintre OP_2 şi Ox. Luăm $\mathcal{F} = \mathcal{B}(\Omega)$ şi probabilitatea pe Ω o definim ca produsul $P = \mathcal{U}_{2\pi} \bigotimes \mathcal{U}_{2\pi}$, ceea ce asigură independenţa poziţiilor lui P_1 şi P_2 cât şi faptul că sunt repartizate uniform. Mulţimea corzilor ce ne interesează este descrisă de perechile de puncte (P_1, P_2) pe cerc cu proprietatea că unghiul dintre OP_1 şi OP_2, măsurat pe partea cea mai scurtă, este mai mare decât $\frac{2\pi}{3}$. Dacă (θ_1, θ_2) reprezintă o astfel de coardă şi presupunem că $\theta_1 \in \left[0, \frac{2\pi}{3}\right]$, atunci putem exprima condiţia aceasta privind unghiul prin

$$\theta_2 \in \left[\theta_1 + \frac{2\pi}{3}, \theta_1 + \frac{4\pi}{3}\right].$$

În cazul în care $\theta_1 \in \left[\frac{2\pi}{3}, \frac{4\pi}{3}\right]$, condiţia devine

$$\theta_2 \in \left[0, \theta_1 - \frac{2\pi}{3}\right] \cup \left[\theta_1 + \frac{2\pi}{3}, 2\pi\right].$$

În cazul în care $\theta_1 \in \left[\frac{4\pi}{3}, 2\pi\right]$, trebuie să fie satisfăcută relaţia

$$\theta_2 \in \left[\theta_1 - \frac{4\pi}{3}, \theta_1 - \frac{2\pi}{3}\right].$$

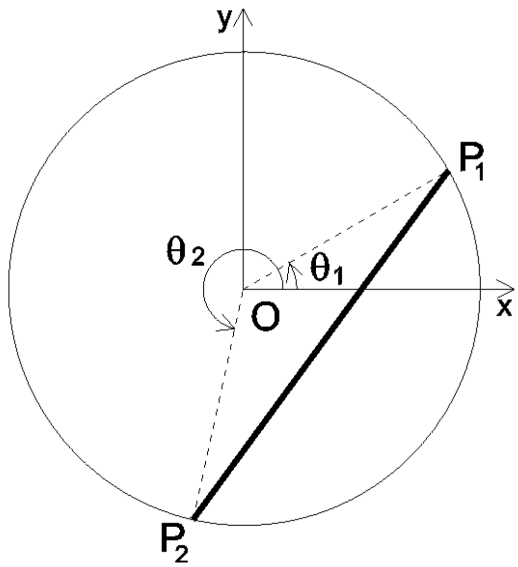


FIGURA 3. Experimentul 3.

Putem atunci scrie mulţimea care ne interesează sub forma

$$\Omega_0 = \Omega_1 \cup \Omega_2 \cup \Omega_3,$$

cu mulţimile din dreapta definite prin

$$\Omega_1 = \left\{ (\theta_1, \theta_2) \in \Omega / \theta_1 \in \left[0, \frac{2\pi}{3}\right], \theta_2 \in \left[\theta_1 + \frac{2\pi}{3}, \theta_1 + \frac{4\pi}{3}\right] \right\},$$

$$\Omega_2 = \left\{ (\theta_1, \theta_2) \in \Omega / \theta_1 \in \left(\frac{2\pi}{3}, \frac{4\pi}{3}\right), \theta_2 \in \left[0, \theta_1 - \frac{2\pi}{3}\right] \cup \left[\theta_1 + \frac{2\pi}{3}, 2\pi\right] \right\},$$

$$\Omega_3 = \left\{ (\theta_1, \theta_2) \in \Omega / \theta_1 \in \left[\frac{4\pi}{3}, 0\right], \theta_2 \in \left[\theta_1 - \frac{4\pi}{3}, \theta_1 - \frac{2\pi}{3}\right] \right\}.$$

Probabilitatea ce ne interesează este $P(\Omega_0) = P(\Omega_1) + P(\Omega_2) + P(\Omega_3)$. Pentru a calcula $P(\Omega_1)$ vom aplica formula lui Fubini:

$$P(\Omega_1) = \frac{1}{2\pi} \int_0^{\frac{2\pi}{3}} \mathcal{U}_{2\pi}(\Omega_{1,\theta_1})\, d\theta_1,$$

unde am notat $\Omega_{1,\theta_1} = \{\theta_2 \in [0, 2\pi) / \ (\theta_1, \theta_2) \in \Omega_0\}$, secţiunea prin θ_1 a lui Ω_1. Din expresia lui Ω_1 se vede că $\Omega_{0,\theta_1} = \left[\theta_1 + \frac{2\pi}{3}, \theta_1 + \frac{4\pi}{3}\right]$, ceea ce implică $\mathcal{U}_{2\pi}(\Omega_{0,\theta_1}) = \frac{1}{3}$ şi atunci $P(\Omega_1) = \left(\frac{1}{3}\right)^2$.

Un calcul similar conduce la $P(\Omega_2) = P(\Omega_3) = \left(\frac{1}{3}\right)^2$, astfel că obţinem $P(\Omega_0) = \frac{1}{3}$. Din nou avem un rezultat ce corespunde valorii obţinute experimental. $\square$

2. Medie şi dispersie*

În continuare vom presupune că $(\Omega, \mathcal{F}, P)$ este un spaţiu probabilizat fixat şi vom discuta despre extensia noţiunii de medie pe care am definit-o pentru variabile cu valori într-o mulţime cel mult numărabilă. Pentru aceasta este util să facem aproximarea variabilelor reale generale cu un şir de variabile care iau valori într-o mulţime cel mult numărabilă. Dacă X este o variabilă aleatoare reală astfel că $X \geq 0$, se defineşte şirul $(X_n)_{n\in\mathbf{N}^*}$ în felul următor:

$$X_n = \frac{k}{n},\ pe\ \left\{\frac{k}{n} \leq X < \frac{k+1}{n}\right\},\ k \in \mathbf{N}.$$

Este clar că fiecare din variabilele $X_n : \Omega \to \left\{\frac{k}{n}/k \in \mathbf{N}\right\}, n \in \mathbf{N}^*$ ia valori într-o mulţime cel mult numărabilă. O altă expresie pentru aceste variabile este

$$X_n = \frac{1}{n}\sum_{k=1}^{\infty} 1_{\left\{\frac{k}{n} \leq X\right\}}\ .$$

Se verifică punctual că au loc inegalităţile $X - \frac{1}{n} < X_n \leq X$. Rezultă că $\lim_{n\to\infty} X_n = X$. Dacă considerăm subşirul $(X_{2^l})_{l\in\mathbf{N}}$, acesta este crescător, după cum se poate uşor verifica.

Următoarea lemă va permite definirea mediei varibilei X.

Lema 6.1. *Dacă $EX_1 < \infty$, atunci $EX_n < \infty$, pentru orice $n \in \mathbf{N}$ şi au loc inegalităţile*

$$|EX_n - EX_m| \leq \frac{1}{n},\ \forall m \geq n.$$

Demonstraţie. Din inegalităţile $X - \frac{1}{n} < X_n \leq X$ şi $X - \frac{1}{m} < X_m \leq X$ rezultă

$$|X_n - X_m| \leq \frac{1}{n},\ (*)$$

dacă $m \geq n$. În particular, avem

$$X_m \leq X_1 + 1,$$

ceea ce implică $EX_m \leq EX_1 + 1$. Cu aceasta am demonstrat afirmaţia din enunţ referitoare la finitudinea mediilor. Pentru a demonstra inegalităţile din enunţ pornim cu următoarele inegalităţi, ce rezultă din (*),

$$X_n - \frac{1}{n} \leq X_m \leq X_n + \frac{1}{n}.$$

De aici se deduce

$$EX_n - \frac{1}{n} \leq EX_m \leq EX_n + \frac{1}{n},$$

care implică inegalităţile din enunţ. □

Din lema anterioară rezultă că dacă $EX_1 < \infty$, atunci şirul $EX_n, n \in \mathbf{N}$ este Cauchy şi, prin urmare, există $\lim_{n\to\infty} EX_n$. Prin definiţie punem $EX = \lim_{n\to\infty} EX_n$ şi spunem că acest număr reprezintă *media* variabilei X. Vom spune că X *admite medie* dacă este îndeplinită condiţia $EX_1 < \infty$.

Pentru o variabilă care nu este neapărat nenegativă vom face descompunerea $X = X^+ - X^-$, unde $X^+ = \max(X, 0)$ şi $X^- = \max(-X, 0)$ reprezintă părţile pozitivă, respectiv negativă, ale lui X. Vom spune că X *admite medie* dacă atât X^+ cât şi X^- admit medii şi în acest caz notăm $EX = EX^+ - EX^-$ şi vom spune că aceasta este *media* variabilei X.

Aceste definiţii sunt diferite de cele uzual utilizate în teoria măsurii, dar sunt de fapt echivalente, după cum un cititor avansat se poate singur convinge. În mod curent media este numită integrală în teoria măsurii iar notaţia pentru ea are versiunile următoare

$$EX = MX = \int_\Omega X(\omega) P(d\omega) = \int X dP.$$

Este vorba de integrala în sensul Lebesgue. Toate proprietăţile demonstrate pentru media variabilelor cu valori cel mult numărabile se extind la cazul variabilelor generale. Facem mai jos o recapitulare a acestora.

1. Dacă X este o variabilă ce admite medie şi $X \geq 0$, atunci $EX \geq 0$.

2. Variabila $X : \Omega \to E \subset \mathbf{R}$ admite medie, dacă şi numai dacă $|X|$ admite medie. În plus are loc inegalitatea $EX \leq E|X|$.

3. Dacă variabila X admite medie, atunci, pentru orice $a \in \mathbf{R}$, variabila aX admite de asemenea medie şi are loc relaţia $E(aX) = aEX$.

4. Dacă $X_1, ..., X_n$ sunt variabile aleatoare ce admit medie, atunci orice sumă de tipul $a_1X_1 + ... + a_nX_n$, cu $a_1, ..., a_n \in \mathbf{R}$, admite de asemenea medie şi are loc formula

$$E(a_1X_1 + ... + a_nX_n) = a_1EX_1 + ... + a_nEX_n.$$

5. Dacă X, Z sunt două variabile aleatoare astfel că Z este integrabilă şi are loc inegalitatea $|X| \leq Z$, atunci şi X este integrabilă şi $EX \leq E|X| \leq EZ$.

6. Inegalitatea lui *Schwartz* rămâne valabilă ca şi în cazul discret.

7. Pentru variabile independente care admit medie avem $E(XY) = (EX)(EY)$.

8. Dacă X este o variabilă aleatoare reală, iar $f : \mathbf{R} \to \mathbf{R}$ este o funcţie continuă pe porţiuni, notăm $Y = f \circ X$ şi aceasta este tot variabilă cu valori reale. Presupunem că repartiţia lui X admite densitatea $P_X = \rho(x) dx$, iar ρf este integrabilă. În aceste condiţii variabila aleatoare Y admite medie şi aceasta se calculează după formula

$$EY = \int f(x) P_X(dx) = \int \rho(x) f(x) dx.$$

Pentru variabile aleatoare care sunt de pătrat integrabile, adică astfel încât X^2 este integrabilă, se defineşte *dispersia* la fel ca pentru variabilele cu valori discrete

$$D^2X := E\left(X - EX\right)^2 = EX^2 - \left(EX\right)^2.$$

Numărul $DX := \sqrt{D^2X}$ este numit *deviaţia standard* a lui X. Variabila $(X - EX)$ care apare în definiţia dispersiei este numită versiune centrată a lui X. Ţinând cont de formula anterioară, pentru cazul în care repartiţia lui X admite densitatea ρ, putem scrie

$$EX = \int x\rho\left(x\right)dx, \quad E\left(X^2\right) = \int x^2 P_X\left(dx\right) = \int x^2\rho\left(x\right)dx,$$

$$D^2X = \int x^2\rho\left(x\right)dx - \left(\int x\rho\left(x\right)dx\right)^2.$$

10. Dacă variabilele $X_1, ..., X_n$ sunt de pătrat integrabile şi independente, atunci are loc formula

$$D^2\left(X_1 + ... + X_2\right) = D^2X_1 + ... + D^2X_n.$$

11. Inegalitatea lui *Cebâşev* rămâne neschimbată.

12. Legea numerelor mari este valabilă ca şi pentru variabile cu valori discrete.

În încheiere să notăm regulile după care se modifică media şi dispersia unei variabile când adunăm o constantă sau când o înmulţim cu o constantă. Calcule directe simple ne spun că

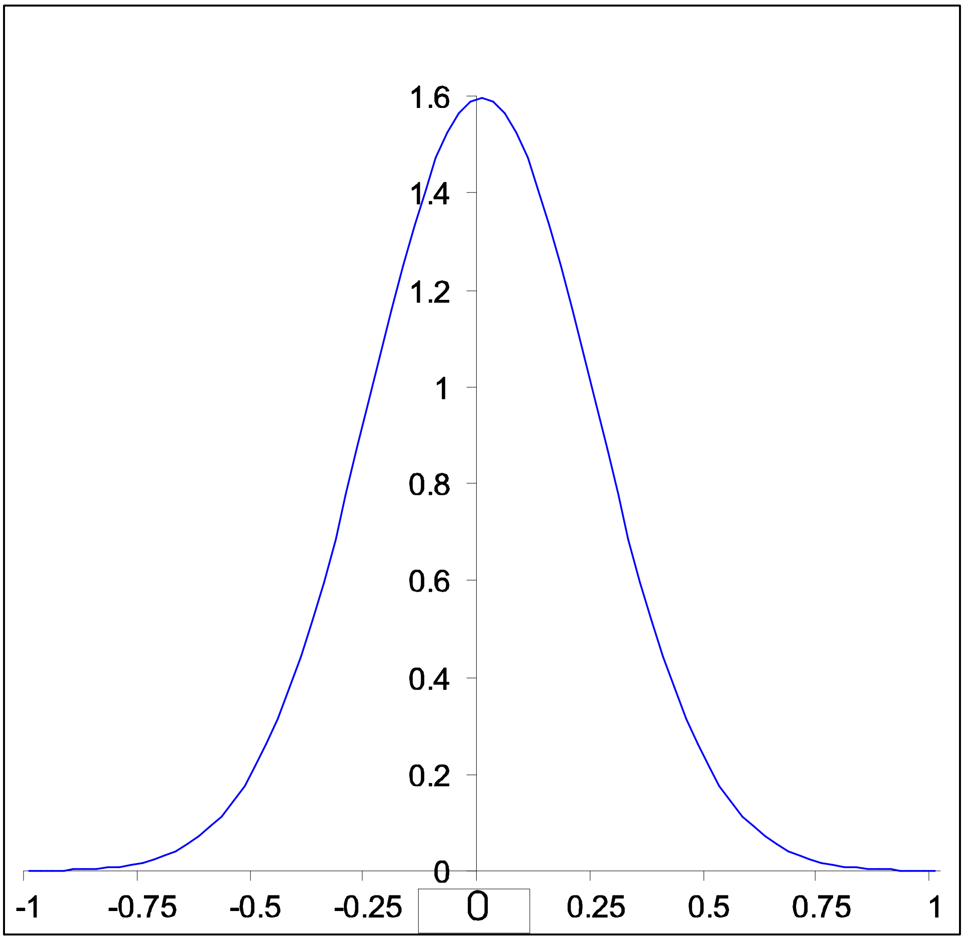
$$E\left(X + \xi\right) = EX + \xi, \quad D^2\left(X + \xi\right) = D^2X,$$
$$E\left(\theta X\right) = \theta EX, \quad D^2\left(\theta X\right) = \theta^2 D^2X.$$

3. Repartiţia normală

Funcţia $\varphi\left(x\right) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ joacă un rol deosebit în teoria probabilităţilor. Graficul acestei funcţii numit şi *clopotul lui Gauss* este reprezentat în figura 4, în partea de jos. Pentru trasarea graficului se ţine cont că $\varphi'\left(x\right) = -x\varphi\left(x\right)$ şi $\varphi''\left(x\right) = \left(x^2 - 1\right)\varphi\left(x\right)$. Este o curbă simetrică faţă de axa Oy, cu un maxim în $\varphi\left(0\right) = \frac{1}{\sqrt{2\pi}} \approx 0,3989$ şi punctele de inflexiune în -1 şi 1 : $\varphi\left(1\right) = \varphi\left(-1\right) \approx 0,2419$. Funcţia este descrescătoare pe intervalul $[0, \infty)$, iar limitele la infinit sunt $\lim\limits_{x\to-\infty}\varphi\left(x\right) = \lim_{x\to\infty}\varphi\left(x\right) = 0$. Ceea ce trebuie reţinut, însă, este faptul că are loc o convergenţă foarte rapidă. Viteza mare de convergenţă la zero este ilustrată, de exemplu, de faptul că pentru numere naturale are loc relaţia $\varphi\left(n+1\right) = \varphi\left(n\right)e^{-n-\frac{1}{2}} \leq \frac{1}{n+1}\varphi\left(n\right)$. Am utilizat aici binecunoscuta inegalitate $1 + x \leq e^x$.

Pe de altă parte, calcule numerice arată că

$$\begin{aligned}\varphi\left(-2\right) &= \varphi\left(2\right) \approx 0,0539, \varphi\left(-3\right) = \varphi\left(3\right) \approx 0,00443,\\ \varphi\left(-4\right) &= \varphi\left(4\right) \approx 0,000134.\end{aligned}$$



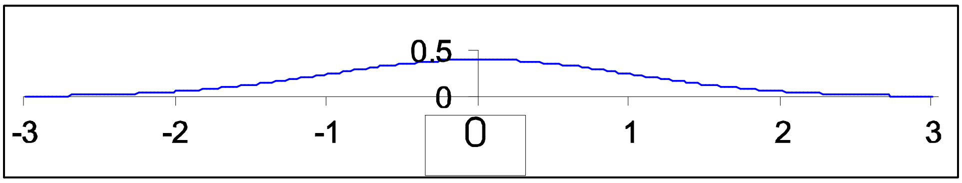


FIGURA 4. Jos este graficul lui φ. Sus este graficul densităţii repartiţiei normale cu $m = 0$ şi $\sigma = 0,25$.

Valorile acestei funcţii sunt tabelate pentru că funcţia este utilizată în probleme de statistică. Din punctul de vedere al majorităţii aplicaţiilor numerice, funcţia φ este considerată zero pentru $|x| \geq 4$.

Valoarea integralei următoare este cunoscută, fiind una din integralele speciale clasice,

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

Pentru completitudinea expunerii vom verifica această egalitate.

Se porneşte cu observaţia că, datorită simetriei funcţiei avem $\int\limits_{-\infty}^{\infty} e^{-\frac{x^2}{2}}dx = 2\int\limits_{0}^{\infty} e^{-\frac{x^2}{2}}dx$. Atunci vom nota $I = \int\limits_{0}^{\infty} e^{-\frac{x^2}{2}}dx$ şi vom scrie pătratul acestei integrale ca o integrală dublă,

$$I^2 = \left(\int\limits_{0}^{\infty} e^{-\frac{x^2}{2}}dx\right)\left(\int\limits_{0}^{\infty} e^{-\frac{y^2}{2}}dy\right) = \iint_{\mathbf{R}^2_+} e^{-\frac{x^2+y^2}{2}}dxdy = \int_0^{\frac{\pi}{2}}\int_0^{\infty} re^{-\frac{r^2}{2}}drd\theta$$
$$= \frac{\pi}{2}\int_0^{\infty} re^{-\frac{r^2}{2}}dr == \frac{\pi}{2}\lim_{a\to\infty}\left(-e^{-\frac{r^2}{2}}\Big|_0^a\right) = \frac{\pi}{2}.$$

La al treilea semn egal am utilizat trecerea la coordonate polare. Rezultă $I = \sqrt{\frac{\pi}{2}}$, ceea ce stabileşte formula dorită.

Relaţia ce tocmai am verificat-o arată că φ este densitatea unei măsuri de probabilitate pe $\mathbf{R}$. Mai precis, se defineşte o măsură pe $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$ prin

$$\mathcal{N}(A) = \int_A \varphi(x)\,dx,$$

iar relaţia anterioară arată că $\mathcal{N}(\mathbf{R}) = 1$. Această măsură de probabilitate este numită *repartiţia normală standard* sau *repartiţia gaussiană standard*. Proprietăţile funcţiei φ conduc la faptul că măsura P este concentrată practic pe intervalul $(-4, 4)$. Estimări numerice arată că au loc relaţiile aproximative următoare

$$\int\limits_{-1}^{1} \varphi(x)\,dx \approx 0,6826, \int\limits_{-2}^{2} \varphi(x)\,dx \approx 0,9546,$$
$$\int\limits_{-3}^{3} \varphi(x)\,dx \approx 0,9972, \int\limits_{-4}^{4} \varphi(x)\,dx \approx 0,9999.$$

Integralele funcţiei φ pe diverse intervale sunt utilizate în calcule statistice şi de aceea se defineşte funcţia următoare

$$\Phi(x) = \int_{-\infty}^{x} \varphi(y)\,dy,\ \forall x \in \mathbf{R},$$

care se numeşte *funcţia lui Laplace*. Prin intermediul acestei funcţii putem exprima orice integrală pe interval: $\int\limits_{a}^{b} \varphi(x)\,dx = \Phi(b) - \Phi(a)$. Pentru a facilita calculele, funcţia lui Laplace este tabelată de obicei cu valori ale lui x care cresc cu $0,01$ de la 0 până la 4. Cărţile de statistică conţin în mod curent, la sfârşit, un număr de tabele care permit efectuarea rapidă a anumitor calcule statistice des întâlnite în aplicaţii. Tabelul cu valorile funcţiei Laplace este nelipsit din orice carte de statistică.

Pentru valori negative ale argumentului se utilizează simetria funcţiei φ, care arată că

$$\Phi(-x) = \int_{-\infty}^{-x} \varphi(y)\, dy = \int_{x}^{\infty} \varphi(y)\, dy = 1 - \Phi(x).$$

În afară de repartiţia normală standard există o întreagă clasă de repartiţii care se numesc tot normale şi care sunt obţinute prin transformări afine din repartiţia standard. Mai precis, find date două constante $\sigma, m \in \mathbf{R}, \boldsymbol{\sigma} > \mathbf{0}$, aplicaţia $h(x) = \sigma x + m$ este o bijecţie pe $\mathbf{R}$. Se notează cu $\mathcal{N}(m, \sigma^2)$ măsura definită prin $\mathcal{N}(m, \sigma^2)(A) = \mathcal{N}(h^{-1}(A)), \forall A \in \mathcal{B}(\mathbf{R})$. Această măsură se poate exprima prin integrala

$$\mathcal{N}\left(m, \sigma^2\right)(A) = \int_B \varphi(x)\, dx,$$

unde am notat $B = h^{-1}(A)$. În cazul unui interval, să zicem $A = (a, b)$ şi corespunzător $B = \left(\frac{a-m}{\sigma}, \frac{b-m}{\sigma}\right)$, facem schimbarea de variabilă $x = h(y) = \frac{y-m}{\sigma}$ şi această integrală devine

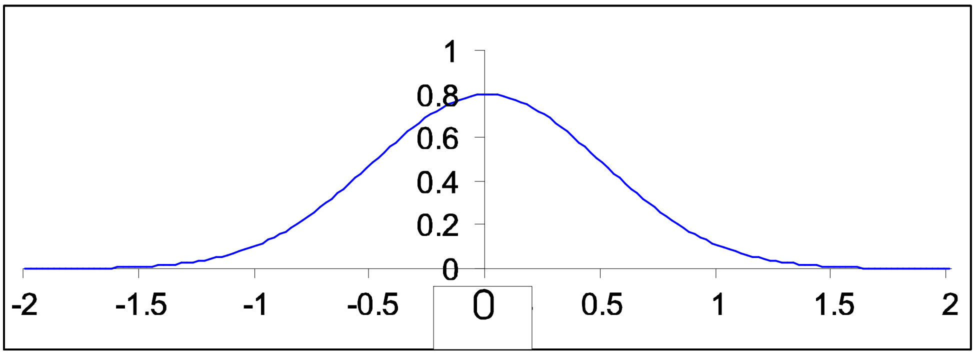
$$\int_{\frac{a-m}{\sigma}}^{\frac{b-m}{\sigma}} \varphi(x)\, dx = \int_A \varphi\left(\frac{y-m}{\sigma}\right) \frac{dy}{\sigma} = \frac{1}{\sigma\sqrt{2\pi}} \int_a^b e^{-\frac{(y-m)^2}{2\sigma^2}} dy.$$

Deci măsura $\mathcal{N}(m, \sigma^2)$ admite drept densitate funcţia

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

Cu această notaţie, măsura normală standard devine $\mathcal{N}(0, 1)$. În partea de sus a figurii 4 este trasat graficul corespunzător acestei funcţii în cazul $m = 0, \sigma = \frac{1}{4}$. În figura 5 sunt reprezentate graficele corespunzătoare cazurilor $m = 0, \sigma = \frac{1}{2}$ şi $m = 5, \sigma = \frac{1}{2}$.

Repartiţia $\mathcal{N}\left(0, \left(\frac{1}{4}\right)^2\right)$, se obţine din $\mathcal{N}$ prin transformarea $h_1(x) = \frac{1}{4}x$, care contractă dreapta reală. Aşa se explică forma primului grafic în raport cu graficul funcţiei φ, iniţiale. Masa măsurii $\mathcal{N}$ a fost strânsă în jurul lui zero. Repartiţia lui $\mathcal{N}\left(5, \left(\frac{1}{2}\right)^2\right)$ se obţine din $\mathcal{N}$ prin transformarea $h(x) = \frac{1}{2}x + 5$, care este contracţia cu $\frac{1}{2}$ urmată de o translaţie cu 5. Se mai poate spune şi că $\mathcal{N}\left(5, \left(\frac{1}{2}\right)^2\right)$ se obţine din $\mathcal{N}\left(0, \left(\frac{1}{2}\right)^2\right)$ prin transformarea $h_2(x) = x + 5$, care este o translaţie.



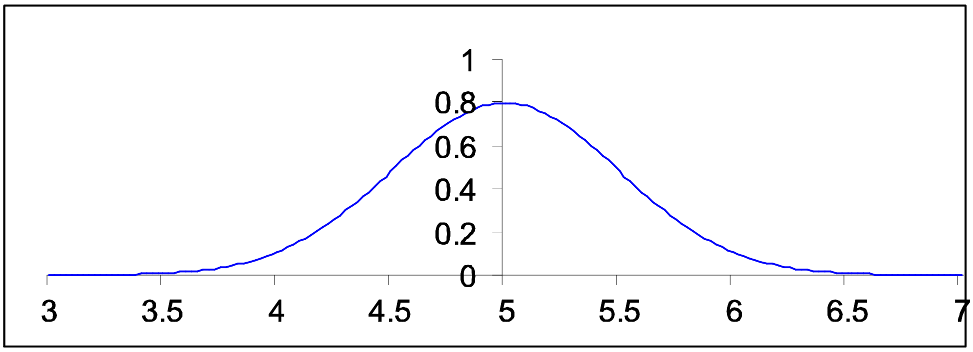


FIGURA 5. Sus este graficul densităţii repartiţiei normale cu $m = 0$ şi $\sigma = 0,5$ şi jos este graficul densităţii cu $m = 5$ şi $\sigma = 0,5$.

Să presupunem că X este o variabilă repartizată normal standard, adică $P_X = \mathcal{N}$. Atunci putem calcula media sa:

$$EX = \int_{\mathbf{R}} x P_X(dx) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-\frac{x^2}{2}} dx.$$

Funcţia care apare sub ultima integrală este integrabilă. Pe de altă parte ea este impară, şi de aceea integrala face zero. Deci media unei variabile repartizate normal standard este zero. Putem calcula şi dispersia variabilei X :

$$\begin{aligned} D^2X &= E\left(X - EX\right)^2 = EX^2 = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{2}} dx = \\ &= \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = 1. \end{aligned}$$

Să considerăm acum variabila $Y = \sigma X + m$, unde $\sigma, m \in \mathbf{R}, \boldsymbol{\sigma} > \mathbf{0}$. Media acestei variabile este $EY = E\left(\sigma X + m\right) = m$, iar dispersia ei este $D^2Y = E\left(Y - m\right)^2 = \sigma^2 EX^2 = \sigma^2$. Pentru a calcula repartiţia lui Y notăm $h\left(x\right) = \sigma x + m$, astfel că $Y = h\left(X\right)$, şi atunci putem scrie

$$\begin{aligned} P_Y\left(A\right) &= P\left(X^{-1}\left(h^{-1}\left(A\right)\right)\right) = P_X\left(h^{-1}\left(A\right)\right) = \\ &= \mathcal{N}\left(h^{-1}\left(A\right)\right) = \mathcal{N}\left(m, \sigma^2\right)\left(A\right). \end{aligned}$$

Putem deduce atunci că repartiţia lui Y este $\mathcal{N}\left(m, \sigma^2\right)$.

În concluzie, parametrul m reprezintă media, iar σ^2 dispersia unei variabile repartizate $\mathcal{N}\left(m, \sigma^2\right)$. De fapt, orice variabilă repartizată $\mathcal{N}\left(m, \sigma^2\right)$ are media m şi dispersia σ^2. Aceste fapte justifică denumirea de *repartiţia normală de medie* m *şi dispersie* σ^2, care este dată repartiţiei $\mathcal{N}\left(m, \sigma^2\right)$ (sau repartiţie gaussiană de medie m şi dispersie σ^2).

Repartiţia normală apare în realitatea fizică frecvent. Fenomene din astronomie, sociologie, antropologie, biologie, agricultură şi altele sunt modelate de repartiţia normală. Printre primele observaţii concrete se numără tabelul cu măsura circumferinţei pieptului la un număr de 5738 de soldati scoţieni (figura 6) remarcat de statisticianul belgian Lambert Adolphe Jacques Quetelet (1796-1874).

4. Exerciţii

Exerciţiul 6.1. *Două autobuze vin în aceeaşi staţie cu probabilitate uniformă între ora* 8 *şi* 8 : 30. *Primul stă în staţie* 10 *minute, iar al doilea* 5 *minute. (i) Care este probabilitatea ca primul să vină înaintea celui de al doilea? (ii) Care este probabilitatea ca cele două autobuze să se întâlnească? (iii) Care este probabilitatea ca ştiind că s-au întâlnit, primul să plece din staţie lăsându-l pe al doilea acolo?*

Measures de la poitrine	Nombre d'hommes	Nombre Proportionnel	Probabilité d'après L'observation	Rang dans La table	Rang d'après le calcul.	Probabilité d'après La table	Nombre d'observations calculé
Pouces							
33	3	5	0.5000			0.5000	7
34	18	31	0.4995	52	50	0.4993	29
35	81	141	0.4964	42.5	42.5	0.4964	110
36	185	322	0.4823	33.5	34.5	0.4854	323
37	420	732	0.4501	26.0	26.5	0.4531	732
38	749	1305	0.3769	18.0	18.5	0.3799	1333
39	1073	1867	0.2464	10.5	10.5	0.2466	1838
			0.0597	2.5	2.5	0.0628	
40	1079	1882	0.1285	5.5	5.5	0.1359	1987
41	934	1628	0.2913	13	13.5	0.3034	1675
42	658	1148	0.4061	21	21.5	0.4130	1096
43	370	645	0.4706	30	29.5	0.4690	560
44	92	160	0.4866	35	37.5	0.4911	221
45	50	87	0.4953	41	45.5	0.4980	69
46	21	38	0.4991	49.5	53.5	0.4996	16
47	4	7	0.4998	56	61.8	0.4999	3
48	1	2	0.5000			0.5000	1
	5738	1,0000					1,0000

FIGURA 6. Tabelul găsit de Quetelet cu dimensiunea toracelui pentru 5738 de soldaţi scoţieni.

CAPITOLUL 7

Teorema de Moivre-Laplace

În acest capitol presupunem că $n \in \mathbf{N}^*$ este ordinul şi $p \in (0,1)$ este parametrul unei repartiţii binomiale. Pentru o mai bună ilustrare a fenomenelor ce le vizăm vom mai presupune că avem un spaţiu probabilizat $(\Omega, \mathcal{F}, P)$ pe care este definită o variabilă aleatoare X cu repartiţia binomială de ordin n şi parametru p. Vom nota

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

Principala problemă a acestui capitol va fi să găsim o bună aproximare a acestor numere şi să aplicăm rezultatul la estimarea parametrului p prin metode statistice. Matematicienii francezi Abraham *de Moivre* (1667-1754) şi Pierre Simon *Laplace* (1749- 1827) au descoperit această aproximare, primul în cazul $p = \frac{1}{2}$, iar cel de al doilea în cazul general.

1. Aproximarea repartiţiei binomiale

Mai departe vom cerceta ordonarea după mărime a numerelor p_k. Un interes natural îl prezintă maximul acestor numere. Dacă p_{k_0} este un maxim şi toate celelalte numere sunt strict mai mici, atunci indicele său k_0 poate fi calificat drept valoarea lui X cu cea mai mare probabilitate de realizare. Un astfel de indice, sau mai corect, o astfel de valoare a lui X se numeşte *mod*. Vom introduce notaţia

$$m = [(n+1)p],$$

unde în partea dreaptă avem partea întreagă a numărului real $(n+1)p$ şi, în lema următoare, vom arăta că, cu unele excepţii, el reprezintă un mod.

Lema 7.1. *Şirul* $p_0, p_1, ..., p_{m-1}$ *este strict crescător, şirul* $p_m, p_{m+1}, ..., p_n$ *este strict descrescător şi* $p_{m-1} \leq p_m$. *Dacă* $(n+1)p$ *este întreg, atunci* $p_{m-1} = p_m$*, în caz contrar* $p_{m-1} < p_m$.

Demonstraţie. Trebuie să examinăm semnul diferenţei

$$p_{k+1} - p_k = p^k q^{n-k-1} \frac{n!}{(k+1)!\,(n-k)!} (pn - q - k).$$

Acest semn este dat de semnul expresiei $pn-q-k = p(n+1)-(k+1)$. Din această ultimă expresie se observă imediat că : 1^0 dacă $k+1 < p(n+1)$, atunci $p_k < p_{k+1}$; 2^0 dacă $p(n+1) < k+1$, atunci avem $p_k > p_{k+1}$; 3^0 dacă $p(n+1) = k+1$, atunci $p_{k+1} = p_k$. □

Observăm că modul m este apropiat de media np. Mai precis, din definiţia părţii întregi a unui număr, rezultă că au loc relaţiile

$$m \leq (n+1)\,p < m+1.$$

Prin urmare, dacă notăm $\delta_n = m - np$, rezultă că $\delta_n \in (-q, p]$. În continuare vom presupune că p **este constant**, dar vom face să varieze n **tinzând la infinit**. Cantitatea δ_n va fi mică în raport cu celelalte care intervin şi la limită va dispărea.

Mai departe facem o aproximare a numărului p_m utilizând următoarea formulă ce reprezintă un caz particular al *formulei lui Stirling*:

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{\theta_n},$$

unde θ_n este un factor de corecţie care este în intervalul $\theta_n \in \left(\frac{1}{12n+1}, \frac{1}{12n}\right)$. (Vezi [**10**] sau [**7**].) Următoarea propoziţie arată că probabilitatea p_m este asimptotic (când $n \to \infty$) de acelaşi ordin de mărime cu $\frac{1}{\sqrt{2\pi npq}}$.

Propoziţia 7.1. *Probabilitatea modului se exprimă prin formula*

$$p_m = \frac{1}{\sqrt{2\pi npq}}\gamma_n,$$

unde γ_n *este un factor de corecţie astfel încât* $\lim\limits_{n\to\infty} \gamma_n = 1$.

Demonstraţie. Vom exprima cu ajutorul formulei lui Stirling numerele $n!, m!$ şi $(n-m)!$ astfel că expresia lui p_m devine

$$p_m = \frac{n!}{m!\,(n-m)!}p^m q^{n-m} =$$

$$\begin{aligned} &= \frac{1}{\sqrt{2\pi}}\frac{n^n}{m^m\,(n-m)^{n-m}}\frac{\sqrt{n}}{\sqrt{m}\sqrt{n-m}}e^{\theta_n-\theta_m-\theta_{n-m}}p^m q^{n-m} = \\ &= \frac{1}{\sqrt{2\pi npq}}\alpha_1\sqrt{\alpha_2}\alpha_3, \end{aligned}$$

unde am notat $\alpha_1 = e^{\theta_n-\theta_m-\theta_{n-m}}, \alpha_2 = \frac{n^2pq}{m(n-m)}$ şi $\alpha_3 = \frac{n^n p^m q^{n-m}}{m^m(n-m)^{n-m}}$.

Ţinând cont că $m = np + \delta_n \geq np - 1$ şi $n - m = nq - \delta_n \geq nq - 1$ rezultă $\theta_m \leq \frac{1}{12(np-1)}, \theta_{n-m} \leq \frac{1}{12(nq-1)}$, ceea ce asigură că $\lim_{n\to\infty}\theta_m = \lim_{n\to\infty}\theta_{n-m} = 0$ şi atunci $\lim_{n\to\infty}\alpha_1 = 1$.

Trebuie să examinăm şi raportul

$$\alpha_2 = \frac{npnq}{(np+\delta_n)\,(nq-\delta_n)},$$

pentru care se vede clar că limita este 1, când $n \to \infty$.

Rămâne să mai studiem raportul

$$\alpha_3 = \frac{(np)^m}{m^m}\frac{(nq)^{n-m}}{(n-m)^{n-m}} = \left(1 - \frac{\delta_n}{np+\delta_n}\right)^{np+\delta_n}\left(1 + \frac{\delta_n}{nq-\delta_n}\right)^{nq-\delta_n}.$$

Pentru a trece la limită după n vom scrie ultima expresie sub forma

$$\left(\left(1-\frac{\delta_n}{np+\delta_n}\right)^{\frac{np+\delta_n}{\delta_n}}\left(1+\frac{\delta_n}{nq-\delta_n}\right)^{\frac{nq-\delta_n}{\delta_n}}\right)^{\delta_n},$$

Deoarece $x_n = \frac{\delta_n}{np+\delta_n}$ şi $y_n = \frac{\delta_n}{nq-\delta_n}$ tind ambele la 0, produsul din expresia anterioară are limita

$$z_n = (1-x_n)^{\frac{1}{x_n}}(1+y_n)^{\frac{1}{y_n}} \to 1.$$

Cum şirul (δ_n) este mărginit, rezultă că $\alpha_3 = z_n^{\delta_n} \to 1$.

Rezumând aceste considerente, putem spune că factorul $\gamma_n = \alpha_1\sqrt{\alpha_2}\alpha_3$ satisface condiţiile din enunţ. $\square$

Pentru probabilităţile p_k, cu indicele k nu prea depărtat de m, comportamentul asimptotic (când $n \to \infty$) este apropiat de comportamentul cantităţii $p_m e^{-\frac{(k-m)^2}{2npq}}$. Următoarea propoziţie exprimă mai precis acest fapt.

Propoziţia 7.2. *Fie $\epsilon_{n,k}$ numărul determinat de relaţia*

$$p_k = p_m e^{-\frac{(k-m)^2}{2npq}} e^{\epsilon_{n,k}} \quad , \quad k \in \{0,1,...,n\}.$$

Atunci, pentru orice constantă $c > 0$, are loc relaţia

$$\lim_{n\to\infty} \sup_{|l|\le c\sqrt{n}} |\epsilon_{n,m+l}| = 0.$$

Demonstraţie. Pornim de la egalitatea

$$p_k = p_m \frac{m!}{k!}\frac{(n-m)!}{(n-k)!}p^{k-m}q^{m-k}.$$

Notăm $l = k - m$ şi, pentru a face o alegere, presupunem $l > 0$. Atunci vom avea

$$p_{m+l} = p_m \frac{(n-m-l+1)\,...\,(n-m)}{(m+1)\,...\,(m+l)}\frac{p^l}{q^l} =$$

$$= p_m \frac{(nq-\delta_n-l+1)\,...\,(nq-\delta_n)}{(np+\delta_n+1)\,...\,(np+\delta_n+l)}\frac{p^l}{q^l} =$$

$$= p_m \frac{\left(1-\frac{\delta_n+l-1}{nq}\right)...\left(1-\frac{\delta_n}{nq}\right)}{\left(1+\frac{\delta_n+1}{np}\right)...\left(1+\frac{\delta_n+l}{np}\right)} = p_m\frac{A}{B},$$

unde am notat cu A şi B numărătorul şi numitorul fracţiei penultime. Vom aproxima mai departe produsele de la A şi de la B prin aplicarea *lemei* ... Pentru A avem

$$A = \exp - \left(\sum_{i=0}^{l-1}\frac{\delta_n+i}{nq}\right)\exp r_1,$$

unde factorul de corecţie este estimat prin

$$-\sum_{i=0}^{l-1} \frac{\left(\frac{\delta_n+i}{nq}\right)^2}{1-\frac{\delta_n+i}{nq}} \leq r_1 \leq 0.$$

Pentru a exprima mai bine estimarea lui r_1 vom observa că $\frac{\delta_n+i}{nq} \leq \frac{l}{nq} \leq \frac{c\sqrt{n}}{nq} = \frac{c}{q\sqrt{n}}$, dacă presupunem că $l \leq c\sqrt{n}$, aşa cum apare în relaţia din enunţ. Deducem astfel că, pentru n mare, vom avea $\frac{\delta_n+i}{nq} < \frac{1}{2}$, ceea ce ne permite să estimăm numitorul din suma anterioară prin $1-\frac{\delta_n+i}{nq} > \frac{1}{2}$ şi apoi să deducem

$$|r_1| \leq 2\sum_{i=0}^{l-1}\left(\frac{\delta_n+i}{nq}\right)^2 \leq \frac{2}{n^2q^2}\sum_{i=0}^{l-1}(1+i)^2 = \frac{2}{n^2q^2}\frac{l\,(l+1)\,(2l+1)}{6} \leq$$

$$\leq \frac{2}{3q}\frac{l\,(l+1)^2}{n^2} \leq \frac{2}{3q}\frac{c\sqrt{n}\,(c\sqrt{n}+1)^2}{n^2} \to 0$$

Notăm că în ultima estimare am obţinut o expresie ce nu depinde de l, deci are loc relaţia

$$\lim_{n\to\infty}\sup_{|l|\leq c\sqrt{n}} |r_1| = 0.$$

La fel se procedează şi cu expresia B :

$$B = \exp\sum_{i=1}^{l}\frac{\delta_n+i}{np}\exp r_2.$$

Un calcul similar cu cel anterior conduce la

$$|r_2| \leq \sum_{i=1}^{l}\left(\frac{\delta_n+i}{np}\right)^2 \leq \frac{(l+1)\,(l+2)\,(2l+3)}{6n^2p^2}$$

şi apoi la

$$\lim_{n\to\infty}\sup_{|l|\leq c\sqrt{n}} |r_2| = 0.$$

Revenind la expresia ce ne interesează, ea arată astfel

$$p_{m+l} = p_m\left(\exp-\sum_{i=0}^{l-1}\frac{\delta_n+i}{nq}\right)(\exp r_1)\left(\exp-\sum_{i=1}^{l}\frac{\delta_n+i}{np}\right)(\exp -r_2)\,.$$

Cele două sume de la exponent se pot pune sub forma

$$-\sum_{i=0}^{l-1}\frac{\delta_n+i}{nq}-\sum_{i=1}^{l}\frac{\delta_n+i}{np} = -\frac{l\delta_n}{npq}-\frac{l\,(l-1)}{2nq}-\frac{l\,(l+1)}{2np} =$$

$$= -\frac{l^2}{2npq}-\frac{l\delta_n}{npq}+\frac{l\,(p-q)}{2npq}.$$

Suma ultimilor doi termeni o notăm r_3 şi putem face majorarea, uniform în raport cu l,

$$r_3 = \left| -\frac{l\delta_n}{npq} + \frac{l\,(p-q)}{2npq} \right| \leq \frac{c\sqrt{n}}{npq} + \frac{c\sqrt{n}}{2npq} \to 0.$$

Revenind la termenul ce ne interesează avem

$$p_{m+l} = \left(\exp -\frac{l^2}{2npq} \right) \exp\left(r_1 - r_2 + r_3\right)$$

şi expresia $\epsilon_{n,m+l} = r_1 - r_2 + r_3$ satisface relaţia din enunţ, conform celor arătate. Cu aceasta am încheiat demonstraţia. □

Cele două proprietăţi anterioare pot fi rezumate în forma rezultatului clasic următor.

Teorema 7.1. *()*

$$p_k = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(k-\mu)^2}{2\sigma^2}} e^{\theta_{n,k}},$$

unde coeficienţii $\theta_{n,k}$ *satisfac relaţia următoare*

$$\lim_{n\to\infty} \sup_{|k-\mu|\leq c\sqrt{n}} \theta_{n,k} = 0,$$

cu orice constantă $c > 0$, *iar* $\mu = np$, $\sigma^2 = npq$ *sunt media, respectiv dispersia variabilei aleatoare repartizate binomial cu parametrii* n, p.

Demonstraţie.* Ţinând cont că $m = \mu + \delta_n$, putem scrie

$$\frac{(k-m)^2}{2npq} = \frac{(k-\mu)^2}{2npq} - \frac{(k-\mu)\,\delta_n}{npq} + \frac{\delta_n^2}{2npq}.$$

Coeficientul $\theta_{n,k}$ se exprimă în funcţie de coeficienţii introduşi în propoziţiile anterioare astfel

$$\theta_{n,k} = \ln \gamma_n + \epsilon_{n,k} + \frac{(k-\mu)\,\delta_n}{npq} - \frac{\delta_n^2}{2npq}.$$

Condiţia $|k-\mu| \leq c\sqrt{n}$ implică evident $|k-m| \leq (c+1)\sqrt{n}$ şi deci putem utiliza relaţia din propoziţia anterioară. □

Comparaţia grafică a repartiţiei binomiale cu funcţia φ.

Aproximaţia dată de teorema de Moivre-Laplace este ilustrată grafic în figurile 1,2 şi 3. Sunt reprezentate histogramele unor repartiţii binomiale de parametrii n şi p, iar prin linie sunt trasate graficele corespunzătoare funcţiei $\varphi_{m,\sigma}$, cu $m = np$ şi $\sigma = \sqrt{npq}$.

Se observă că în cazul $n = 10$ şi $p = 0,5$ cele două reprezentări sunt foarte apropiate. Graficul intersectează fiecare dreptunghi al histogramei foarte aproape de mijlocul bazei superioare. În cazul $n = 10$ şi $p = 0,1$ apar diferenţe simţitoare între reprezentări. În general, pentru $n = 10$ se observă o evoluţie a diferenţelor pentru diverse valori ale

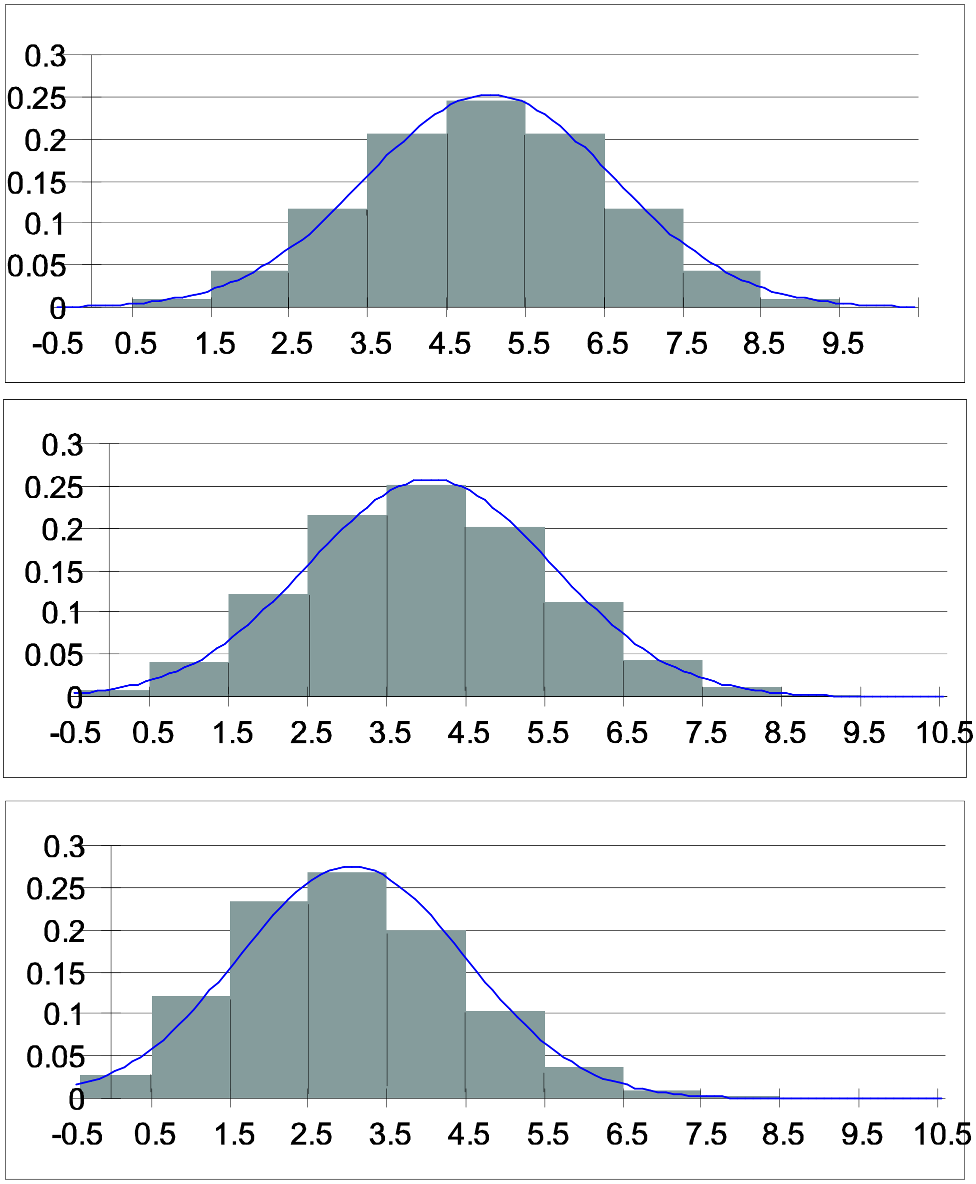


FIGURA 1. Histogramele repartiţiei binomiale cu $n = 10$ şi $p = 0,5$, apoi $p = 0,4$ şi $p = 0,3$ şi graficele densităţilor $\varphi_{\mu,\sigma}$ corespunzătoare, suprapuse.

lui p. De fapt, în aplicaţii, diferenţele de aproximare a repartiţiei binomiale prin repartiţia normală se fac simţite nu atât prin distanţa dintre mijloacele dreptunghiurilor histogramei şi valoarea densităţii $\varphi_{m,\sigma}$ corespunzătoare, cât prin diferenţa dintre ariile dreptunghiurilor şi ariile porţiunilor corespunzătoare de sub graficul densităţii. Acest lucru rezultă din teorema limită centrală ce va fi prezentată în secţiunea care urmează.

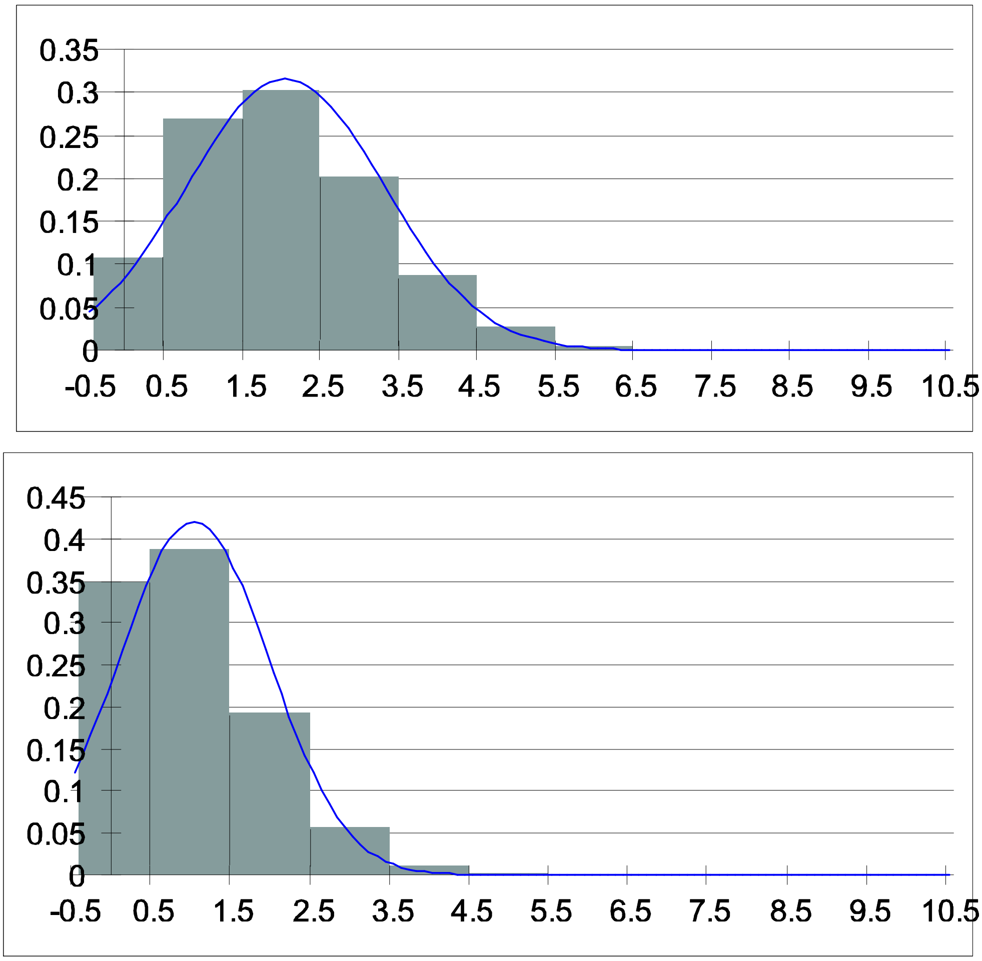


FIGURA 2. Histogramele repartiţiei binomiale cu $n = 10$ şi $p = 0,2$, apoi $p = 0,1$ şi graficele densităţilor $\varphi_{\mu,\sigma}$ corespunzătoare, suprapuse.

Pentru graficele cu $n = 100$, de asemenea, se observă o evoluţie a gradului de apropiere între histograme şi grafice. Dar de data aceasta, chiar şi în cazul deosebirilor maxime, care este $p = 0,1$, se observă că diferenţele măsurate prin arii sunt, proporţional, mai mici decât în cazul $n = 10$ şi $p = 0,1$.

2. Teorema limită centrală

Teorema de Moivre-Laplace are o serie de aplicaţii interesante în statistică. Modul în care se aplică această teoremă va fi pus în lumină de următoarea teoremă cunoscută sub numele de "teorema limită centrală". (În fapt, aici avem un caz particular al acesteia, dar semnificativ şi netrivial.) Formularea acestei teoreme se face de obicei într-un cadru puţin diferit faţă de cel anterior, în sensul că nu se adresează

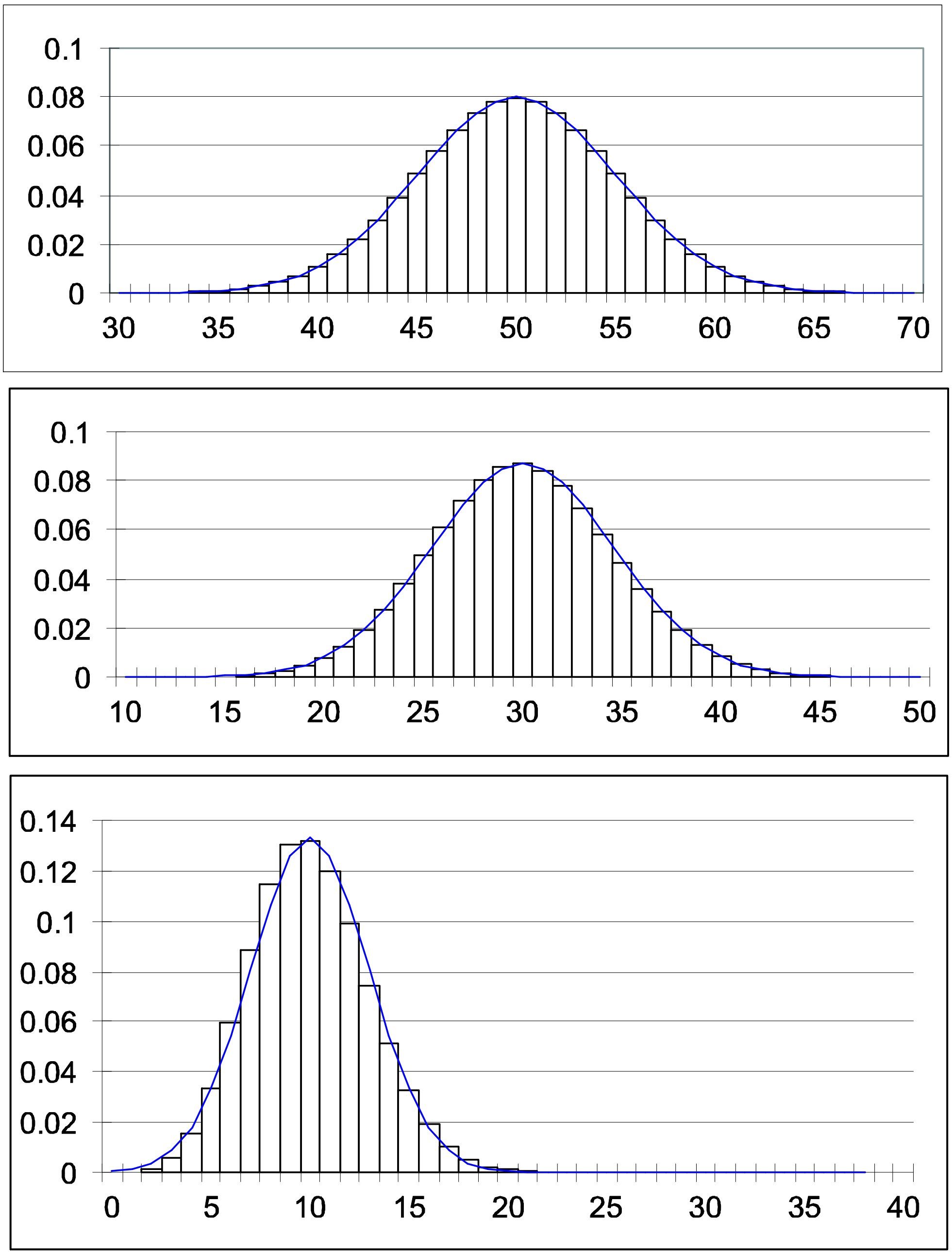


FIGURA 3. Histogramele repartiţiei binomiale cu $n = 100$ şi $p = 0,5$, apoi $p = 0,3$ şi $p = 0,1$ şi graficele densităţilor $\varphi_{\mu,\sigma}$ corespunzătoare, suprapuse.

direct unui şir de repartiţii binomiale ci unui şir de variabile aleatoare repartizate binomial. Exemplul tipic care conduce la astfel de variabile este dat de sumele parţiale ale unui şir de variabile independente. Fie

$(A_n)_{n\in\mathbf{N}^*}$ un şir de evenimente independente şi cu aceeaşi probabilitate $p = P(A_n) \in (0,1)$, pentru orice $n \in \mathbf{N}^*$. Notăm $X_n = 1_{A_n}$ şi $S_n = \sum\limits_{i=1}^{n} X_n$. Variabila aleatoare S_n va fi repartizată binomial cu parametrii n, p şi notăm $q = 1-p$. Forma sub care am prezentat teorema de Moivre-Laplace arată că variabilele $S_n, n \in \mathbf{N}$ au un anumit comportament regulat când $n \to \infty$. Dar se poate arăta că $\lim_{n\to\infty} S_n = \infty$, aproape sigur. Mai mult, legea tare a numerelor mari afirmă că de fapt are loc relaţia limită $\lim_{n\to\infty} \frac{S_n}{n} = p$, aproape sigur. Acest rezultat nu intră în obiectivele acestui curs.

Metoda optimă prin care se exploatează aproximarea obţinută în *teorema* 7.1 constă în normalizarea sumelor astfel

$$\frac{S_n - np}{\sqrt{npq}}.$$

Aceste variabile au toate media 0 şi dispersia 1. Ele nu converg în sensul punctual ci numai repartiţiile lor converg la rapartiţia normală într-un anumit sens. Aici, în abordarea noastră, teorema următoare trebuie privită doar ca producând o aproximare utilă în calcule statistice, ceea ce a fost şi intenţia iniţiatorilor de Moivre şi Laplace.

Teorema 7.2. *Dacă $a < b$ sunt două numere reale, atunci are loc relaţia*

$$\lim_{n\to\infty} P\left(a < \frac{S_n - np}{\sqrt{npq}} < b\right) = \frac{1}{\sqrt{2\pi}} \int\limits_a^b e^{-\frac{x^2}{2}} dx.$$

DEMONSTRAŢIE.* Ca să uşurăm scrierea vom utiliza notaţia $\mu = np$ şi $\sigma = \sqrt{npq}$ pentru media şi respectiv, deviaţia standard a lui S_n. Mai reamintim notaţia pentru valorile repartiţiei binomiale

$$p_k = P(S_n = k) = C_n^k p^k q^{n-k}, k = 0, ..., n.$$

Probabilitatea pe care vrem să o estimăm poate fi scrisă astfel:

$$P\left(a < \frac{S_n - \mu}{\sigma} < b\right) = P(a\sigma < S_n - \mu < b\sigma) = \sum_{k: a\sigma < k-\mu < b\sigma} p_k.$$

Ţinând cont de teorema de Moivre -Laplace şi comparând histograma repartiţiei binomiale cu graficul funcţiei

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

apare în mod natural ideea de a aproxima fiecare număr p_k prin integrala

$$I_k = \int_{k-\frac{1}{2}}^{k+\frac{1}{2}} \varphi_{\mu,\sigma}(x)\, dx.$$

Pentru a estima diferenţa $p_k - I_k$ vom scrie pe I_k într-o altă formă exprimând exponentul din expresia funcţiei $\varphi_{\mu,\sigma}$ sub forma

$$\frac{(x-\mu)^2}{2\sigma^2} = \frac{(k-\mu)^2}{2\sigma^2} + \frac{(x-k)^2 + 2(x-k)(k-\mu)}{2\sigma^2}.$$

Rezultă că avem

$$I_k = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(k-\mu)^2}{2\sigma^2}} \int_{k-\frac{1}{2}}^{k+\frac{1}{2}} e^{-\frac{(x-k)^2+2(x-k)(k-\mu)}{2\sigma^2}} dx.$$

Ţinând cont că $|x-k| \leq \frac{1}{2}$ şi că $|k-\mu| \leq c\sigma$, unde $c = \max(-a, b)$, exponentul din această ultimă integrală se majorează astfel:

$$\left| \frac{(x-k)^2 + 2(x-k)(k-\mu)}{2\sigma^2} \right| \leq \frac{\frac{1}{4} + c\sigma}{2\sigma^2}.$$

Din teorema lui Lagrange rezultă că integrala anterioară poate fi scrisă sub forma

$$\int_{k-\frac{1}{2}}^{k+\frac{1}{2}} e^{-\frac{(x-k)^2+2(x-k)(k-\mu)}{2\sigma^2}} dx = e^{\eta_{n,k}},$$

unde $\left|\eta_{n,k}\right| \leq \frac{1+4c\sigma}{8\sigma^2}$. În particular, avem

$$\lim_{n\to\infty} \sup_{k:|k-np|<c\sqrt{n}} \left|\eta_{n,k}\right| \leq \lim_{n\to\infty} \frac{1+4c\sqrt{npq}}{8npq} = 0.$$

Cu notaţia introdusă la *propoziţia* 7.2 putem scrie $I_k = p_k e^{-\theta_{n,k}+\eta_{n,k}}$ şi apoi estimăm

$$|p_k - I_k| = p_k \left|1 - e^{-\theta_{n,k}+\eta_{n,k}}\right| \leq 2p_k \left(|\theta_{n,k}| + \left|\eta_{n,k}\right|\right), \ (*)$$

utilizând inegalitatea $|1-e^x| \leq 2x$, valabilă pentru $|x| \leq 2$.

Mai departe notăm

$$k_1 = \inf\left\{k / a\sqrt{npq} + np < k\right\}, \ k_2 = \sup\left\{k / k < np + b\sqrt{npq}\right\}$$

astfel că putem scrie

$$P\left(a < \frac{S-\mu}{\sigma} < b\right) = \sum_{k=k_1}^{k=k_2} p_k.$$

Suma din dreapta relaţiei este aproximată, ţinând cont de relaţia (*), prin

$$\left| \sum_{k=k_1}^{k=k_2} p_k - \int_{k_1-\frac{1}{2}}^{k_2+\frac{1}{2}} \varphi_{\mu,\sigma}(x)\, dx \right| \leq 2 \sum_{k=k_1}^{k=k_2} p_k \left(|\theta_{n,k}| + \left|\eta_{n,k}\right|\right). \ (**)$$

Cum $\sum_{k=k_1}^{k=k_2} p_k \leq 1$, ultima expresie se majorează cu

$$2 \sup_{k_1 \leq k \leq k_2} \left(|\theta_{n,k}| + \left|\eta_{n,k}\right|\right),$$

iar această cantitate tinde la zero când n tinde la infinit.

Este momentul să ne întoarcem acum la integrala ce apare la limita din enunţ şi să o punem sub o formă apropiată de ceea ce am obţinut mai sus:

$$\int_a^b \varphi(x)\,dx = \int_{\mu+a\sigma}^{\mu+b\sigma} \varphi_{\mu,\sigma}(x)\,dx.$$

Pentru a modifica capetele intervalului de integrare al ultimei expresii ţinem cont de estimarea generală

$$\left|\int_u^v f - \int_{u'}^{v'} f\right| \le \int_{\min(u,u')}^{\max(u,u')} |f| + \int_{\min(v,v')}^{\max(v,v')} |f|$$

şi de faptul că $\left|k_1 - \frac{1}{2} - \mu - a\sigma\right| \le \frac{1}{2}$ şi $\left|k_2 + \frac{1}{2} - \mu - b\sigma\right| \le \frac{1}{2}$, astfel că obţinem

$$\left|\int_{k_1-\frac{1}{2}}^{k_2+\frac{1}{2}} \varphi_{\mu,\sigma}(x)\,dx - \int_{\mu+a\sigma}^{\mu+b\sigma} \varphi_{\mu,\sigma}(x)\,dx\right| \le \frac{1}{\sqrt{2\pi npq}}.$$

Concluzionând, inegalitatea (**) şi cu cea de mai sus ne conduc la

$$\left|P\left(a < \frac{S_n - \mu}{\sigma} < b\right) - \int_a^b \varphi(x)\,dx\right| \le 2 \sup_{k_1 \le k \le k_2} \left(|\theta_{n,k}| + |\eta_{n,k}|\right) + \frac{1}{\sqrt{2\pi npq}}.$$

Expresia din dreapta tinde la zero când n tinde la infinit, ceea ce permite a se deduce relaţia dorită. □

O formă frecventă de aplicare a acestei teoreme este următoarea formulă aproximativă

$$P\left(np - 2\sqrt{npq} < S_n < np + 2\sqrt{npq}\right) \approx 0,9546,$$

unde s-a luat $a = -2$ şi $b = 2$ şi s-a utilizat valoarea cunoscută pentru repartiţia normală, neglijând complet eroarea de aproximare. Aceasta este utilizată în cazul unei evaluări preliminare sumare dar rapide. Similar apare utilizată formula

$$P\left(np - 3\sqrt{npq} < S_n < np + 3\sqrt{npq}\right) \approx 0,9972.$$

Exemplu (acumularea erorilor).

La o fabrică de chibrituri se pun câte 50 de beţe într-o cutie de chibrituri şi, cu o probabilitate de 50% se face o eroare de ± un băţ. Cutiile se ambalează cu hârtie în pachete de câte 12 şi aceste pachete se pun în pachete mari de carton, câte 12. Deci un pachet de carton conţine $12 \times 12 = 144$ cutii de chibrituri. Ne interesează care este numărul de beţe ce se află într-un pachet de carton.

Soluţie. Eroarea de umplere a cutiilor este descrisă de o repartiţie care are masa $\frac{1}{2}$ în punctul 0, masa $\frac{1}{4}$ în punctul 1 şi masa $\frac{1}{4}$ în punctul -1. Observăm că o repartiţie asemănătoare are suma $X + Y$, dacă variabilele X şi Y sunt independente şi ambele repartizate Bernoulli cu parametrul $p = \frac{1}{2}$. Anume repartiţia lui $X + Y$ are masa $\frac{1}{2}$ în 1, masa $\frac{1}{4}$ în 0 şi masa $\frac{1}{4}$ în 2. Rezultă că variabila $Z = X + Y - 1$ are exact repartiţia erorii de umplere a unei cutii de chibrituri.

Aşadar, pentru a descrie 144 de erori de umplere independente, vom presupune că pe un spaţiu $(\Omega, \mathcal{F}, P)$ avem construite variabilele aleatoare $X_1, ..., X_{144}, Y_1, ..., Y_{144}$, independente, identic repartizate Bernoulli, cu parametrul $p = \frac{1}{2}$. Variabila $Z_i = X_i + Y_i - 1$ descrie eroarea de umplere la cea de a i -a cutie, pentru fiecare $i = 1, ..., 144$. Eroarea care se acumulează la cele 144 de cutii este reprezentată de variabila

$$T = \sum_{i=1}^{144} Z_i = \sum_{i=1}^{144} X_i + \sum_{i=1}^{144} Y_i - 144.$$

Repartiţia variabilei $U = T - 144 = \sum\limits_{i=1}^{144} X_i + \sum\limits_{i=1}^{144} Y_i$ este o repartiţie binomială de parametri $p = \frac{1}{2}$ şi $n = 288$. Avem $\sqrt{npq} = \sqrt{72} \approx 8,485$. Aplicăm aproximarea dată de teorema limită centrală pentru $S_n = U$:

$$P(-16 \leq T \leq 16) = P(144 - 16 \leq U \leq 144 + 16) \approx 0,95,$$

$$P(-25 \leq T \leq 25) = P(144 - 25 \leq U \leq 144 + 25) \approx 0,99.$$

Deci, cu o mare probabilitate, la numărul mediu de $144 \times 50 = 7200$ beţe într-un carton, pot fi în plus sau în minus doar 16 şi aproape întotdeauna eroarea va fi sub 25. □

Acurateţea aproximării date de teorema limită centrală.*

În aplicaţii nu este necesară evaluarea fiecărui număr p_k aproximat prin teorema de Moivre-Laplace ci este cel mai adesea nevoie de o aproximare de tipul celei date în teorema limită centrală. Această aproximare este utilizată în practică chiar şi pentru valori mici ale lui n. Pentru ca aceasta să se facă cât mai bine există câteva corecţii care se fac şi pe care le vom prezenta, fără demonstraţii mai jos. Pentru două numere k_1 şi $k_2 \in \mathbf{N}$ astfel încât $0 \leq k_1 < k_2 \leq n$ vom nota

$$p_{k_1,k_2} = P(k_1 \leq S_n \leq k_2) = P\left(\frac{k_1 - np}{\sqrt{npq}} \leq \frac{S_n - np}{\sqrt{npq}} \leq \frac{k_2 - np}{\sqrt{npq}}\right).$$

Teorema limită centrală ne spune că această probabilitate este aproximată de diferenţa $\Phi\left(\frac{k_2-np}{\sqrt{npq}}\right) - \Phi\left(\frac{k_1-np}{\sqrt{npq}}\right)$. Se dovedeşte însă că numărul

$$N(k_1, k_2) = \Phi\left(\frac{k_2 - np + \frac{1}{2}}{\sqrt{npq}}\right) - \Phi\left(\frac{k_1 - np - \frac{1}{2}}{\sqrt{npq}}\right)$$

aproximează mai bine probabilitatea p_{k_1,k_2}. Adăugarea fracţiei $\frac{1}{2}$ în această expresie se spune că introduce o corecţie de continuitate pentru aproximarea noastră. (Numele sugerează faptul că repartiţia binomială, care este discontinuă, este aproximată mai bine de repartiţia normală, care este continuă, dacă se face această corecţie.)

Dacă parametrul p este $\frac{1}{2}$, atunci estimarea erorilor de aproximare conduce la concluzia că diferenţa dintre p_{k_1,k_2} şi $N(k_1, k_2)$ este mai mică dacât $0,01$ pentru $n \geq 10$ şi mai mică decât $0,005$ pentru $n \geq 20$.

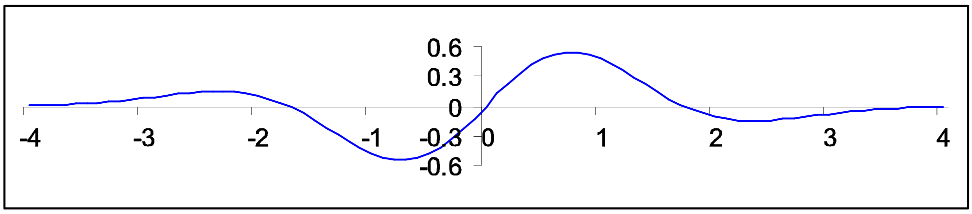


FIGURA 4. Graficul funcţiei φ'''.

Pentru cele mai multe aplicaţii practice o astfel de eroare este neglijabilă.

În cazul $p \neq \frac{1}{2}$ eroarea este evaluată pornind de la următoarea formulă

$$p_{k_1,k_2} \approx \int_a^b \left(\varphi(x) + \frac{1}{6} \frac{p-q}{\sqrt{npq}} \varphi'''(x) \right) dx \ , \tag{2.1}$$

unde $a = \frac{k_1 - np - \frac{1}{2}}{\sqrt{npq}}$, $b = \frac{k_2 - np + \frac{1}{2}}{\sqrt{npq}}$, $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, iar φ''' este derivata de ordinul trei a lui φ. Eroarea cu care este verificată egalitatea aproximativă de mai sus este inferioară lui $0,013$ dacă $\sqrt{npq} \geq 5$. (Vezi Uspenski [**21**] teorema din cap.VII, sec.11, pg. 129. Este de notat că în lucrarea [**17**] la pag. 103 şi la pag. 107 se afirmă în cursul unor comentarii fără demonstraţie că eroarea ar fi de fapt mai mică de $0,005$ chiar sub condiţia mai slabă $\sqrt{npq} \geq 3$.)

Termenul care integrează φ''' se numeşte corecţia de asimetrie. Faptul că $p \neq \frac{1}{2}$ produce o asimetrie între termenii p_{m+k} şi p_{m-k}, asimetrie ce este preluată de corecţia de asimetrie în formula de aproximare.

Graficul funcţiei $\varphi'''(x) = (3x - x^3)\varphi(x)$ este uşor de trasat şi arată ca în figura 4. Calculând derivata φ^{iv} se găsesc extremele. Un punct de maxim este în $\sqrt{3-\sqrt{6}}$ şi un punct minim este în $\sqrt{3+\sqrt{6}}$.

Valoarea maximă pentru integrala $\int_a^b \varphi'''(x)\,dx$ este $0,577$ şi se obţine când $a = 0$ şi $b = \sqrt{3}$. Iar vloarea minimă este $-0,577$ şi se obţine când $a = -\sqrt{3}$ şi $b = 0$. Ţinând cont de presupunerea $\sqrt{npq} \geq 5$, rezultă următoarea estimare uniformă pentru corecţia de asimetrie,

$$\left| \frac{(p-q)}{6\sqrt{npq}} \int_a^b \varphi'''(x)\,dx \right| \leq \frac{0,577}{30} \leq 0,0193. \tag{2.2}$$

Putem deci utiliza aproximarea fără corecţia de asimetrie: $p_{k_1,k_2} \approx N(k_1, k_2)$, cu o eroare de cel mult $0,013 + 0,0193 \leq 0,033$ (bineînţeles sub condiţia $\sqrt{npq} \geq 5$).

Pentru cazul unor intervale simetrice faţă de media np corecţia de asimetrie devine foarte mică datorită faptului că funcţia φ''' este antisimetrică. Mai concret, să zicem că pornim la evaluarea următoarei probabilităţi

$$P\left(\left|\frac{S_n-np}{\sqrt{npq}}\right|<\alpha\right)=P\left(np-\alpha\sqrt{npq}<S_n<np+\alpha\sqrt{npq}\right)=p_{k_1,k_2},$$

unde α este un număr dat, iar umerele k_1 şi k_2 sunt asociate lui astfel

$$k_1=\inf\left\{k\in\mathbf{N}/\ k>np-\alpha\sqrt{npq}\right\},$$
$$k_2=\sup\left\{k\in\mathbf{N}/\ k<np+\alpha\sqrt{npq}\right\}.$$

Rezultă că aceste ultime numere se pot exprima sub forma $k_1=np-\alpha\sqrt{npq}+\delta_1$, $k_1=np+\alpha\sqrt{npq}-\delta_2$, cu $\delta_1,\delta_2\in(0,1]$. Trecem acum la evaluarea integralei $\int_a^b\varphi'''(x)\,dx$, unde

$$a=\frac{k_1-np-\frac{1}{2}}{\sqrt{npq}}=-\alpha+\frac{\delta_1}{\sqrt{npq}}-\frac{1}{2\sqrt{npq}},$$
$$b=\frac{k_2-np+\frac{1}{2}}{\sqrt{npq}}=\alpha-\frac{\delta_2}{\sqrt{npq}}+\frac{1}{2\sqrt{npq}}.$$

Ţinem cont că funcţia φ''' este antisimetrică şi obţinem

$$\int_a^b\varphi'''(x)\,dx=\int_0^b\varphi'''(x)\,dx-\int_0^{-a}\varphi'''(x)\,dx=\int_{-a}^b\varphi'''(x)\,dx.$$

Pentru evaluarea ultimei integrale ţinem cont că valoarea lui φ''' în minimul local cel mai din dreapta este $\varphi'''\left(\sqrt{3+\sqrt{6}}\right)=-0.149$ iar în $1,5$ avem $\varphi'''(1,5)=0,145$. Rezultă că $|\varphi'''(x)|\leq 0,15$ pentru $x\geq 1,5$.

Vom impune condiţia $\alpha\geq 1,6$, astfel ca să avem $-a,b\geq\alpha-\frac{1}{2\sqrt{npq}}\geq 1,5$, (ţinând cont şi de condiţia $\sqrt{npq}\geq 5$). Atunci avem $|b-a|\leq|\delta_1-\delta_2|$ şi prin urmare,

$$\left|\int_a^b\varphi'''(x)\,dx\right|\leq 0,15\times\frac{|\delta_1-\delta_2|}{\sqrt{npq}}\leq 0,03,$$

ceea ce conduce la

$$\frac{1}{6}\left|\frac{p-q}{\sqrt{npq}}\int_a^b\varphi'''(x)\,dx\right|\leq\frac{1}{30}\times 0,03=0,001.$$

Putem concluziona că pentru orice $\alpha\geq 1,6$ are loc estimarea

$$P\left(\left|\frac{S_n-np}{\sqrt{npq}}\right|<\alpha\right)\approx\int_a^b\varphi(x)\,dx=\Phi(b)-\Phi(a),\tag{2.3}$$

cu o eroare mai mică decât $0,014$, unde a şi b sunt definite mai sus.

Exemplu.

O companie aeriană a constatat, după o lungă experienţă, că pasagerii care fac rezervare pentru un zbor renunţă după aceea în proporţie de 10% la efectuarea zborului. De aceea compania decide să accepte 445 de rezervări pentru un avion care are doar 420 de locuri. Care este probabilitatea ca cel puţin un pasager să fie respins, deşi are rezervare? Care este probabilitatea ca un avion pentru care au fost rezervate 445 locuri să plece cu mai mult de 10 locuri libere?

Soluţie. Este vorba de a modela pe un spaţiu probabilizat $(\Omega, \mathcal{F}, P)$ cele 445 de rezervări ca evenimente independente, $A_1, ..., A_{445}$, fiecare cu probabilitatea de realizare $P(A_i) = 0,9$. Suma $S = \sum_{i=1}^{445} 1_{A_i}$ ne dă numărul de pasageri care se prezintă la avion pentru a efectua zborul. Ea este repartizată binomial cu ordinul $n = 445$ şi parametrul $p = 0,9$. Pentru a estima probabilitatea $P(S > 420)$ putem încerca mai întâi să aplicăm teorema limită centrală. Pentru aceasta calculăm media $np = 445 \times 0,9 = 400,5$ şi deviaţia standard $\sqrt{npq} = 6,328$. Apoi scriem evenimentul care ne interesează sub forma

$$\{S > 420\} = \left\{ \frac{S - 400,5}{6,328} > 3,08 \right\}.$$

Aplicând teorema limită centrală fără să ţinem seama de eroarea de aproximare se obţine

$$P(S > 420) \approx 1 - \Phi(3,08) \approx 1 - 0,9989 = 0,0011.$$

Acesta este un număr foarte mic şi ne trezeşte suspiciunea unei erori la aproximare. La evaluarea erorii prin estimarea dată de Uspensky nu putem apela pentru că valoarea $0,013$ este prea mare în raport cu probabilitatea care se prefigurează. De aceea apelăm la calculator şi scoatem valoarea exactă

$$P(S > 420) = 0,000334,$$

care este şi mai mică. Aceasta este probabilitatea de a avea mai mulţi pasageri care se prezintă la îmbarcare decât locuri în avion.

Probabilitatea ca numărul de locuri libere să fie mai mare de 10 este

$$P(S < 410) = 0,9260. \square$$

Regula radicalului.

Să presupunem că avem în faţă o monedă îndoită, care, desigur, ne aşteptăm să cadă cu probabilităţi diferite pe cele două feţe. Încercăm să stabilim care este probabilitatea de a cădea cifra prin experimentarea aruncărilor repetate. Dar problema nu este aşa de simplă cum pare. Autorul a încercat acest lucru cu diverse obiecte, precum monede îndoite, nasturi de diverse forme, o lentilă de lupă foarte bombată, şi a avut surpriza să constate că după serii de câte 100 de aruncări nu se obţin rezultate concludente. În sensul că numerele de apariţii ale celor două feţe au fost aproximativ egale, aşa cum se întâmplă lucrurile în cazul aruncării cu o monedă perfectă. Mai precis, fiecare faţă a ieşit între 40 şi 60 de ori, aşa cum ar corespunde repartiţiei binomiale cu $n = 100$ şi $p = 0,5$. Desigur că aceste experimente au arătat că, deşi diferit de $0,5$, parametrul p care corespunde acestor obiecte nu este foarte depărtat de această valoare. Figura 5 pune în evidenţă imposibilitatea determinării parametrului, atunci când avem de decelat dacă este $p = 0,5$ sau $p = 0,45$, în urma unui experiment cu $n = 100$.

Se pune atunci întrebarea de câte ori ar trebui să aruncăm cu o monedă strâmbă, de exemplu care are parametrul $p = 0,45$, pentru a obţine rezultate net diferite de cele corespunzătoare aruncărilor, de acelaşi număr de ori, cu o monedă perfectă.

Pentru a face estimări numerice revenim la relaţia lui Uspenski notată cu (*) mai sus. Observăm că funcţia φ''' este impară şi de aceea integrala sa pe intervale simetrice de tipul $(-b, b)$ este zero: $\int_{-b}^{b} \varphi'''(x)\,dx = 0$. Deci relaţia (*), în cazul $a = -b$ şi presupunând că $\sqrt{npq} \geq 5$, ne spune că avem $p_{k_1,k_2} \approx N(k_1, k_2)$, cu o eroare de cel mult $0,013$. Dacă în plus avem $b \geq 2$, atunci tabelul cu valorile funcţiei lui Laplace ne spune că $N(k_1, k_2) \geq 0,9544$. În acest caz vom avea în mod sigur $p_{k_1,k_2} \geq 0,9544 - 0,013 \geq 0,94$, care este o probabilitate destul de mare.

Să vedem cum putem construi intervalele $[k_1, k_2]$, cât mai mici, dar în aşa fel încât să îndeplinească condiţiile de mai sus. Condiţia de simetrie $a = -b$, revine la $k_2 - np = np - k_1$, adică intervalul $[k_1, k_2]$ trebuie să aibă mijlocul np. Condiţia $b \geq 2$ revine la

$$k_2 \geq np + 2\sqrt{npq} - \frac{1}{2}.$$

Putem deci alege $k_2 = k_2(n, p)$ să fie cel mai mic întreg care satisface această inegalitate şi să punem apoi $k_1 = 2np - k_2$. Cu aceste numere vom avea $p_{k_1,k_2} \geq 0,94$. Putem spune că intervalul astfel construit, $[k_1, k_2]$, reprezintă un „interval de atenţie" pentru repartiţia binomială de rang n şi parametru p. Aruncând de n ori o monedă care are probabilitatea p de a cădea cu cifra în sus, vom obţine rezultate cuprinse în intervalul $[k_1, k_2]$, cu probabilitate mai mare de $0,94$. Distanţa dintre centrul intervalului şi extremităţi este $d := np - k_1 = k_2 - np$ şi, ţinând cont de definiţia lui k_2, avem $2\sqrt{npq} - \frac{1}{2} \leq d \leq 2\sqrt{npq} + \frac{1}{2}$. Cum întotdeauna avem $pq \leq \frac{1}{4}$, rezultă $d \leq \sqrt{n} + \frac{1}{2}$.

Să presupunem că avem de comparat două repartiţii binomiale de acelaşi rang n şi de parametrii $p' < p$. Distanţa dintre mediile corespunzătoare este $n(p - p')$. Când n este mare distanţa aceasta depăşeşte substanţial lărgimea intervalului de atenţie. Avem atunci următoarea concluzie, numită „regula lui $\sqrt{n}$" : intervalele de atenţie corespunzătoare lui p şi p' au lărgimea de ordinul lui $\sqrt{n}$ în timp ce distanţa dintre centrele acestor intervale este de ordinul lui n. În consecinţă, pentru n mare, intervalele de atenţie devin disjuncte.

În cazul problemei concrete $p = 0,5$ şi $p' = 0,45$ avem următoarele rezultate numerice:

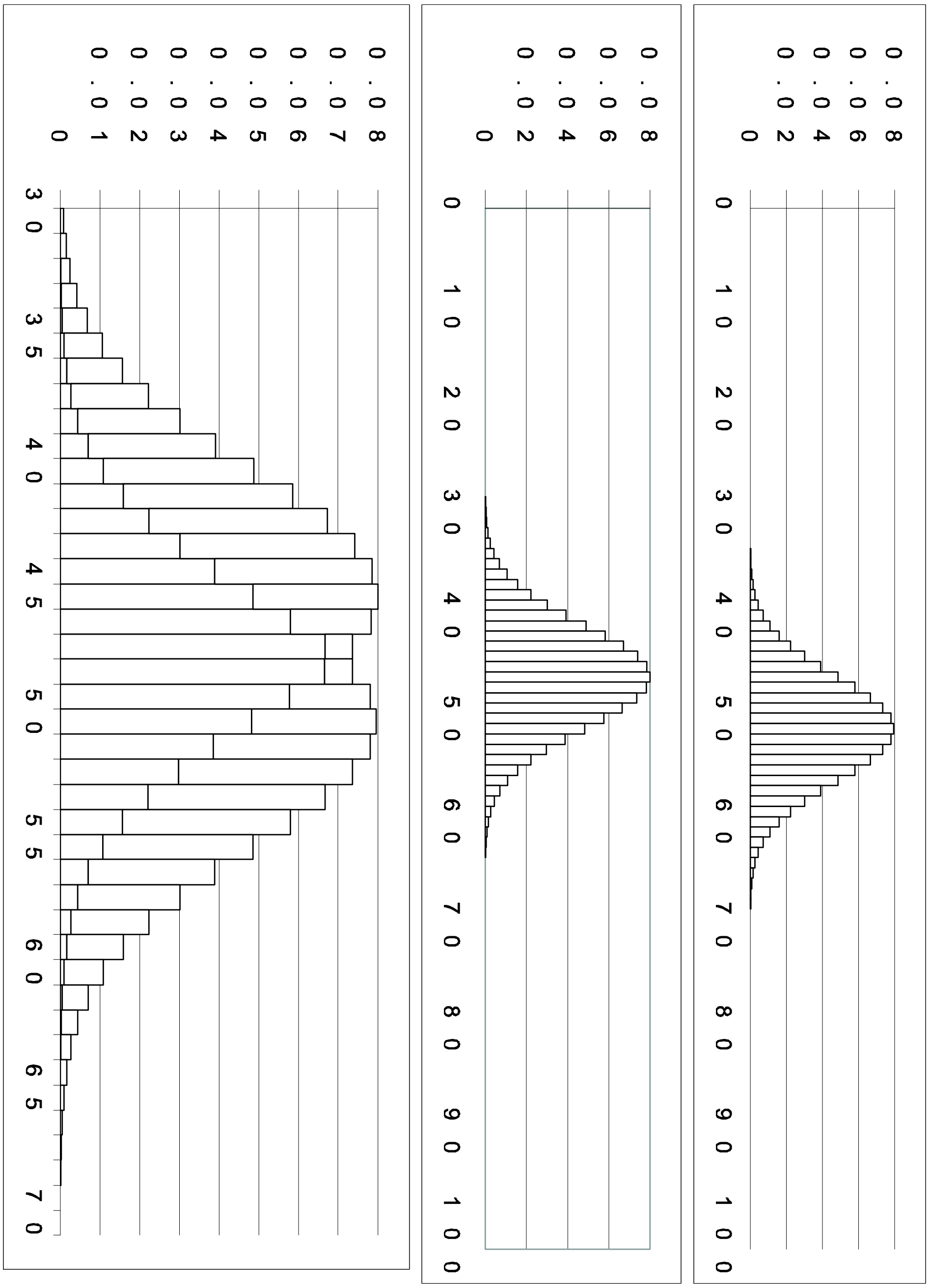


Figura 5. Histogramele repartiţiei binomiale cu $n = 100$ şi $p = 0,5$ şi cu $p = 0,45$. Jos este pusă în evidenţă suprapunerea celor două reparițiii. Pentru claritate am facut reprezentarea la o altă scară.

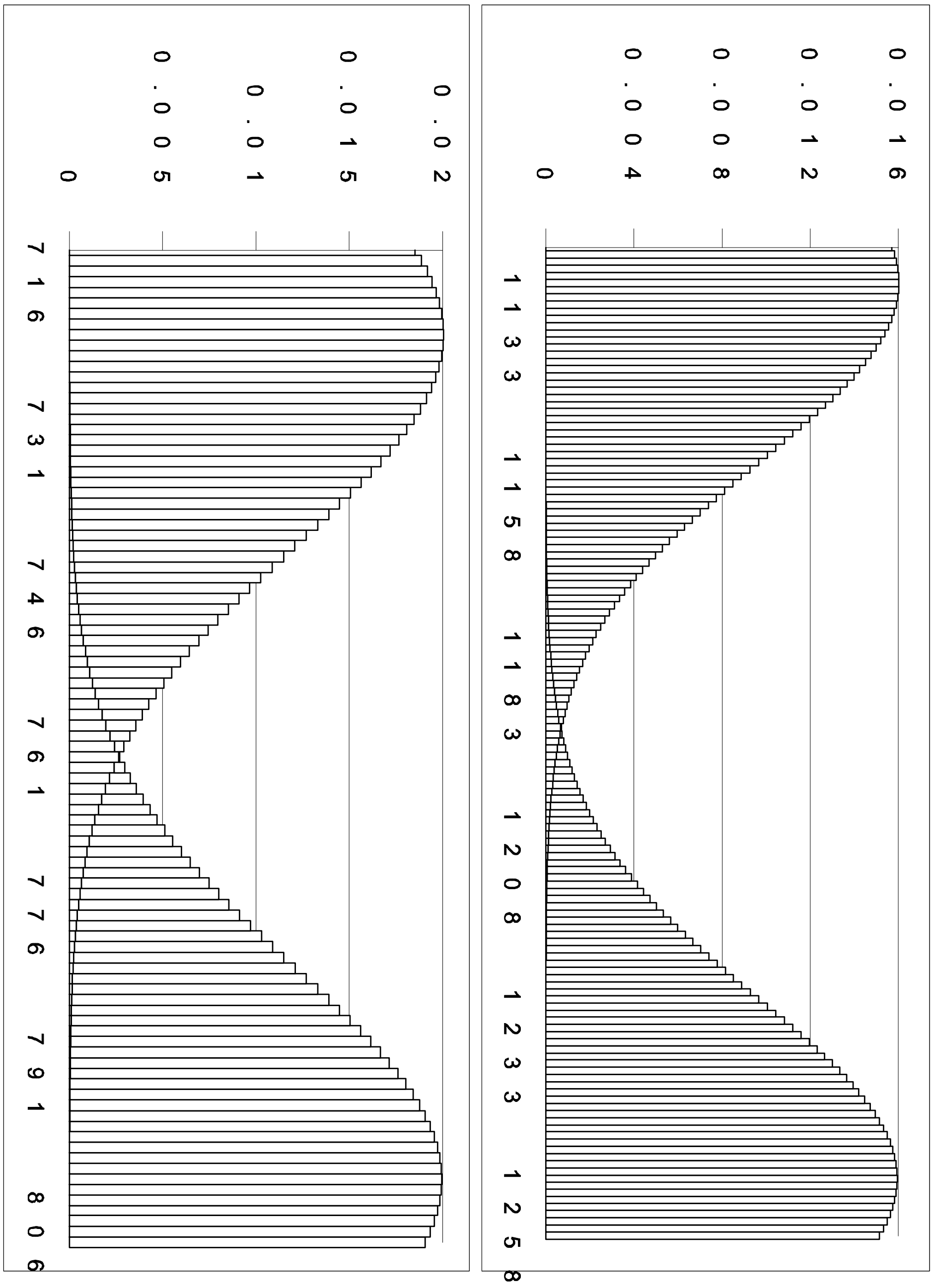


FIGURA 6. În graficele de jos se vede suprapunerea repartiţiilor binomiale cu $n = 1600$ şi $p = 0,5$ şi respectiv $p = 0,45$. Graficele de sus arată practic despărţirea celor două repartiţii binomiale cu $n = 2500$ şi $p = 0,5$ şi respectiv $p = 0,45$.

p	n	$\sqrt{n}$	np	k_1	k_2
$p = 0,5$	100	10	50	40	60
$p = 0,45$	100	10	45	35	55
$p = 0,5$	1600	40	800	760	840
$p = 0,45$	1600	40	720	680	760
$p = 0,5$	2500	50	1250	1200	1300
$p = 0,45$	2500	50	1125	1075	1175

Să presupunem că aruncăm o monedă de 1600 ori şi o considerăm echilibrată dacă ea cade în sus cu aceeaşi faţă de un număr de ori care este cuprins în intervalul de atenţie $[760, 840]$. Pe de altă parte, dacă o variabilă X' este repartizată binomial de rang n şi parametru $0,45$, atunci ea poate avea valori mai mari de 760 cu o probabilitate ce poate fi calculată cu calculatorul şi este

$$P(X' \geq 760) \leq 0,023.$$

Deci la un astfel de test, ne putem înşela şi acceptăm o monedă cu $p = 0,45$ drept echilibrată, cu o probabilitate mai mică decât $0,023$.

În mod asemănător putem calcula pentru o variabilă, X'', repartizată binomial de rang $n = 2500$ şi parametru $p = 0,45$:

$$P(X'' \geq 1200) \leq 0,0014.$$

Deci putem stabili testul următor: se aruncă o monedă de 2500 ori şi o considerăm echilibrată dacă fiecare faţă cade în sus de un număr de ori cuprins în intervalul de atenţie $[1200, 1300]$. Posibiltatea ca având o monedă cu $p = 0,45$ să obţinem un rezultat mai mare de 1200 pentru una din feţe are probabilitatea mai mică de $0,0014$. Deci probabilitatea de a ne înşela cu un astfel de test, pentru o monedă cu $p = 0,45$, este foarte mică.

În figurile 5 şi 6 sunt reprezentate repartiţiile binomiale despre care am discutat. Reamintim că repartiţia binomială are valori strict pozitive în fiecare punct din suportul său. Totuşi marea majoritate a acestor valori sunt atât de mici încât în graficele noastre nu apar. Numai valorile din jurul modului sunt vizibile în graficele noastre şi asta pentru că axa valorilor verticale este dilatată.

3. Noţiuni de estimare statistică*

În această secţiune vom presupune că pe un spaţiu probabilizat, $(\Omega, \mathcal{F}, P)$, sunt date variabilele aleatoare reale $X_1, ..., X_n$, despre care ştim că sunt independente şi că au aceaşi distribuţie Bernoulli, dar este necunoscut parametrul p. Presupunem că parametrul se ştie că se află într-un interval $I \subset [0, 1]$. Variabilele iau valori atunci în mulţimea $E = \{0, 1\}$. E^n este spaţiul în care ia valori vectorul $X = (X_1, ..., X_n)$. Spunem că E^n este *spaţiul eşantioanelor* sau *spaţiul de selecţie*, iar X este o *selecţie aleatoare* sau un *eşantion aleator* de volum n. Conform cu *propoziţia* 4.4, repartiţia lui X este măsura produs

$$\mu_p = (q\delta_0 + p\delta_1) \bigotimes ... \bigotimes (q\delta_0 + p\delta_1),$$

care este o măsură de probabilitate pe E^n, iar $q = 1-p$. Se urmăreşte estimarea parametrului necunoscut $p \in I$ prin metode statistice pornind de la rezultatele unui experiment care oferă realizări concrete pentru selecţia aleatoare, adică pentru variabilele $X_1, ..., X_n$.

Exemplul 1.

Acesta este modelul potrivit pentru cazul problemei controlului de calitate. Se caută determinarea proporţiei p de piese defecte care sunt produse într-un proces de fabricaţie. Pentru aceasta, la punctul de ieşire din procesul de fabricaţie sunt testate un număr de n piese. Secvenţa $X_1(\omega), ..., X_n(\omega)$, pentru $\omega \in \Omega$, reprezintă un "eşantion" sau o "selecţie" făcută printre obiectele produse. În cazul controlului de calitate notăm cu 1 rezultatul testului în cazul în care piesa este defectă şi cu 0 rezultatul în cazul în care piesa este corespunzătoare din punct de vedere calitativ. Efectuarea unui control la un eşantion de n piese conduce la consemnarea unei secvenţe de tipul $0, 0, 1, 0, 1, ...1, 0, 0,$ în care sunt n cifre de zero sau unu. În continuare vom vedea cum se prelucrează această secvenţă de numere pentru a determina parametrul p necunoscut. De fapt ceea ce vom utiliza concret va fi doar un număr, anume numărul care reprezintă de câte ori apare cifra 1 în secvenţa de n cifre.

3.1. Intervale de încredere.

Definiţia 7.1. *Fiind dată o aplicaţie* $f : E^n \to I$, *vom spune că variabila aleatoare compusă* $f(X_1, ..., X_n)$ *reprezintă un* estimator. *Fiind daţi doi estimatori* $\theta_i = f_i(X_1, ..., X_n)$, $i = 1, 2$, *se spune că intervalul aleator* (θ_1, θ_2) *este un* interval de încredere *pentru p. Aplicaţia* $\gamma : I \to [0, 1]$ *definită prin*

$$\gamma(p) = P(\theta_1 < p < \theta_2),$$

este numită coeficientul de încredere *asociat intervalului de încredere* (θ_1, θ_2).

Desigur că această definiţie este cam vagă, în sensul că obiectele sunt prea puţin fixate. În cazuri concrete trebuiesc considerate intervale de încredere care au relevanţă, adică sunt rezonabil construite pentru a prinde valorile lui p în interior. Repartiţiile estimatorilor θ_1 şi θ_2 depind de repartiţia variabilelor X_i şi de aceea γ este o funcţie de p în general. Pentru a avea relevanţă, un coeficient de încredere trebuie să fie cât mai apropiat de 1, independent de $p \in (0, 1)$. În mod practic, uneori se reuşeşte construirea unui interval de încredere care dă un coeficient de încredere constant. Alteori, intervalele de încredere se construiesc în aşa fel încât se poate obţine o margine inferioară, uniformă şi apropiată de 1, pentru coeficientul de încredere. Mai observăm că dacă (θ_1, θ_2) şi

$\left(\theta_1^{'}, \theta_2^{'}\right)$ sunt două intervale de încredere astfel încât $\theta_1^{'} \leq \theta_1$ şi $\theta_2 \leq \theta_2^{'}$, atunci coeficienţii de încredere corespunzători γ şi $\gamma^{'}$ satisfac relaţia

$$\gamma\left(p\right) = P\left(\theta_1 < p < \theta_2\right) \leq P\left(\theta_1^{'} < p < \theta_2^{'}\right) = \gamma^{'}\left(p\right).$$

Cu cât este mai mic intervalul de încredere cu atât estimarea lui p este mai precisă, dar în acelaşi timp coeficientul de încredere se micşorează. Alegerea unui interval de încredere are de rezolvat şi acest aspect contradictoriu: intervalul trebuie să dea o precizie suficientă în determinarea lui p, dar în acelaşi timp să asigure un coeficient de încredere ridicat.

Media empirică.

Variabila aleatoare $\overline{X} = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$ poartă numele de media empirică. Acesta este un estimator care estimează media p a variabilelor date. (Legea numerelor mari dă un sens mai precis cuvântului „estimează”, afirmând că media empirică converge la valoarea medie, aproape sigur, pentru $n \to \infty$.) Funcţia $f : E^n \to \mathbf{R}$, definită prin $f\left(x_1, ..., x_n\right) = \frac{1}{n}\sum\limits_{i=1}^{n} x_i$ satisface relaţia $\overline{X} = f\left(X_1, ..., X_n\right)$.

Deoarece suma $X_1 + ... + X_n$ este repartizată binomial cu ordinul n şi parametrul p, rezultă că media mediei empirice este $E\overline{X} = \frac{np}{n} = p$, iar dispersia este $D^2\overline{X} = \frac{1}{n^2}D^2\left(X_1 + ... + X_n\right) = \frac{1}{n^2}npq = \frac{pq}{n}$.

3.2. Evaluarea coeficientului de încredere. Vom utiliza acum estimarea lui Uspensky sub forma 2.3 ca să obţinem o margine inferioară pentru coeficientul de încredere asociat unui interval de încredere centrat în jurul mediei empirice.

Propoziţia 7.3. *Dacă* $I \subset (0,1)$, $\sqrt{pq} \leq c$ *şi* $\sqrt{npq} \geq 5$, *pentru orice* $p \in I$, *atunci intervalul de încredere* $\left(\overline{X} - \frac{3c}{\sqrt{n}}, \overline{X} + \frac{3c}{\sqrt{n}}\right)$ *are un coeficient de încredere mai mare de* $0,98$, *iar intervalul* $\left(\overline{X} - \frac{2c}{\sqrt{n}}, \overline{X} + \frac{2c}{\sqrt{n}}\right)$ *are un coeficient de încredere mai mare de* $0,94$.

Demonstraţie. Deoarece $S_n = \sum\limits_{i=1}^{n} X_n = n\overline{X}$ şi $\sqrt{pq} \leq c$, putem scrie

$$\left\{\overline{X} - 3\frac{c}{\sqrt{n}} < p < \overline{X} + 3\frac{c}{\sqrt{n}}\right\} \supset \left\{\overline{X} - 3\frac{\sqrt{pq}}{\sqrt{n}} < p < \overline{X} + 3\frac{\sqrt{pq}}{\sqrt{n}}\right\} =$$

$$= \left\{\left|\frac{S_n - np}{\sqrt{npq}}\right| < 3\right\}.$$

Pentru evaluarea ultimului eveniment aplicăm 2.3:

$$P\left(\left|\frac{S_n - np}{\sqrt{npq}}\right| < 3\right) \approx \int_{-3}^{3} \varphi\left(x\right)dx = 0,9972$$

cu o eroare mai mică de $0,014$, ceea ce implică afirmaţia din enunţ referitoare la primul coeficient de încredere.

Pentru cel de al doilea coeficient de încredere aplicăm acelaşi raţionament şi aproximarea

$$P\left(\left|\frac{S_n - np}{\sqrt{npq}}\right| < 2\right) \approx \int_{-2}^{2} \varphi(x)\, dx = 0,9546. \square$$

Exemplul 1.

Un laborator îşi propune să determine coeficientul p care reprezintă raportul dintre numărul de pui masculi şi numărul total de pui, la naştere, pentru o anumită specie de iepuri. În acest scop se face un recensământ la fermele de iepuri înregistrându-se 56.680 familii de iepuri fiecare cu câte 8 pui, în total $56680 \times 8 = 429.440$ de pui din care 221.023 sunt masculi.

<u>*Soluţie.*</u> Modelăm problema în felul următor: variabilele $X_1, ..., X_n$ reprezintă rezultatele uneia din cele $n = 429.440$ naşteri, iar media empirică calculată pentru eşantionul nostru este

$$\overline{X} = 221.023 : 429.440 = 0,5146.$$

Mai departe calculăm $\sqrt{n} = 655$ şi, dată fiind situaţia concretă, putem face presupunerea că p aparţine intervalului $I = \left(\frac{1}{4}, \frac{3}{4}\right)$. Prin urmare este satisfăcută cerinţa $\sqrt{npq} > 5$. Mai facem majorarea $\sqrt{pq} = \sqrt{p(1-p)} \leq \frac{1}{2}$ şi atunci propoziţia anterioară ne asigură că intervalul

$$\left(0,5146 - \frac{3}{2 \times 655};\ \ 0,5146 + \frac{3}{2 \times 655}\right) = (0,51230;\ \ 0,51689)$$

are un coeficient de încredere mai mare decât $0,98$. $\square$

Exemplul 2.

Observăm că în exemplul anterior parametrul p a fost determinat cu două zecimale exacte. Se pune întrebarea: care ar fi numărul minim, n, de pui ce ar trebui luat în calcul pentru a obţine trei zecimale exacte în estimarea parametrului p, cu aceeaşi margine pentru coeficientul de încredere? Şi care ar fi numărul minim pentru a avea două zecimale exacte?

<u>*Soluţie.*</u> Pentru a obţine trei zecimale exacte trebuie ca marginile intervalului de încredere $\left(\overline{X} - \frac{3}{2\sqrt{n}}, \overline{X} + \frac{3}{2\sqrt{n}}\right)$ să difere printr-o valoare mai mică decât 10^{-3}. Deci trebuie să avem $\frac{3}{\sqrt{n}} < \frac{1}{1.000}$, adică $n > 9.000.000$. Un raţionament similar arată că pentru a avea două zecimale exacte, trebuie ca n să satisfacă $\frac{3}{\sqrt{n}} < \frac{1}{100}$, adică să avem $n > 90.000$.$\square$

3.3. Alt interval de încredere. Vom examina mai întâi proprietăţile de monotonie în raport cu parametrul ale probabilităţilor asociate unei repartiţii binomiale. Vom presupune $n \in \mathbf{N}^*$ fixat şi $p \in (0,1)$

variabil. Pentru fiecare $k = 0, ..., n$ notăm

$$F(p,k) = \sum_{l=0}^{k} C_n^l p^l (1-p)^{n-l},$$

aşa numita *probabilitate cumulată* de la 0 la k. Următoarea lemă rezultă din studiul derivatei funcţiei corespunzătoare.

Lema 7.2. *Dacă $p \leq \frac{k}{n}$ şi $p' < p$, sau dacă $\frac{k}{n} \leq p$ şi $p < p'$, în ambele cazuri are loc inegalitatea*

$$p^k (1-p)^{n-k} > (p')^k (1-p')^{n-k}.$$

Lema 7.3. *Fie $0 \leq k < n$ fixe. Aplicaţia $p \to F(p,k)$ este strict descrescătoare.*

Demonstraţie. Fie $p < p'$. Vom analiza trei cazuri.

Cazul 1. Presupunem că $\frac{k}{n} \leq p$. Atunci fiecare termen din expresia lui $F(p,k)$ se compară cu termenul de acelaşi rang din expresia lui $F(p',k)$, după cum precizează lema anterioară:

$$C_n^l p^l (1-p)^{n-l} > C_n^l (p')^l (1-p')^{n-l}, l = 0, ..., k.$$

De aici se deduce $F(p,k) > F(p',k)$.

Cazul 2. Presupunem $p' \leq \frac{k}{n}$ şi atunci raţionăm exact ca în cazul anterior, dar pentru termenii expresiei

$$1 - F(p,k) = \sum_{l=k+1}^{n} C_n^l p^l (1-p)^{n-l}.$$

Avem

$$C_n^l p^l (1-p)^{n-l} < C_n^l (p')^l (1-p')^{n-l}, l = k+1, ..., n,$$

ceea ce implică $1 - F(p,k) < 1 - F(p',k)$, adică inegalitatea dorită.

Cazul 3. Presupunem că $p < \frac{k}{n} < p'$ şi atunci ne bazăm pe cazurile anterioare pentru a deduce $F(p,k) > F\left(\frac{k}{n}, k\right) > F(p',k)$, ceea ce încheie demonstraţia. $\square$

Înainte de a construi un nou interval de încredere, mai precis şi cu coeficientul mai bine estimat, vom pune în evidenţă noua metodă de analiză pe un exemplu.

Exemplul 3.

Pentru a determina calitatea unui produs se testează 2500 de bucăţi. Dintre acestea 95 sunt găsite cu defect. Proporţia de produse defecte pare a fi de $\frac{95}{2500} = 0,038$ sau, în procente, $3,8\%$.

Soluţie. Pentru a determina un interval în care se află raportul p dintre numărul produselor defecte şi numărul total al produselor fabricate modelăm problema ca mai înainte. Avem $n = 2500, S_n = 95$ şi $\overline{X} = \frac{95}{2500} = 0,038$. Deci se conturează concluzia că valoarea lui p

trebuie să fie aproape de numărul $0,038$. Scoatem din calculator următoarele date, pentru $p = 0,049$ şi pentru $p = 0,0288$:

$$F(0,049;95) = \sum_{k=0}^{95} C_{2500}^{k} (0,049)^k (0,951)^{2500-k} = 0,004892$$

$$1 - F(0,0288;94) = \sum_{k=95}^{2500} C_{2500}^{k} (0,0288)^k (0,9712)^{2500-k} = 0,004829.$$

Aplicând lema precedentă pentru $p = 0,049$ şi $p' \geq 0,049$ se deduce pentru cazul că variabilele $X_1, ..., X_n$ sunt de parametru p'

$$P(S_{2500} \leq 95) = F(p', 95) \leq F(0,049;95) < 0,005.$$

Prin urmare, probabilitatea de a greşi eliminând din discuţie cazul $p \geq 0,049$ este foarte mică.

La fel, pentru $p = 0,0288$ şi $p' \leq p$ se deduce, pentru acelaşi caz,

$$P(S_{2500} \geq 95) = 1 - F(p', 94) \leq 1 - F(0,0288;94) < 0,005.$$

Aceste inegalităţi fac neacceptabile valorile lui p din afara intervalului $(0,0288;0,049)$ şi suntem conduşi la concluzia că este mai verosimil ca parametrul p să se afle în acest interval. $\square$

Trecem acum la exploatarea ideii din soluţia de mai sus într-un sens mai general. Anume, vom arăta că intervalul anterior este un interval de încredere cu un bun coeficient de încredere.

Vom presupune în continuare că avem $0 \leq k < n$. Funcţia $p \to F(p,k)$ este continuă şi are limitele

$$\lim_{p \to 0} F(p,k) = 1, \quad \lim_{p \to 1} F(p,k) = 0,$$

deci ea poate fi inversată. Vom nota cu $g(\alpha, k)$ inversa aplicaţiei $p \to F(p,k)$. Pentru fiecare $\alpha \in (0,1)$, numărul $g(\alpha,k) \in (0,1)$ este determinat de egalitatea $F(g(\alpha,k),k) = \alpha$. Pentru $k = n$ punem $g(\alpha, n) = 1$.

Dacă $0 < k \leq n$, vom defini $\overline{g}(\beta,k) = g(1-\beta, k-1)$. Pentru fiecare $\beta \in (0,1)$, numărul $\overline{g}(\beta,k)$ este determinat de relaţia $F(\overline{g}(\beta,k), k-1) = 1-\beta$, sau altfel exprimat, acest număr este determinat de relaţia $P_{\overline{g}(\beta,k)}(S_n \geq k) = \beta$, în care se presupune că repartiţia variabilelor $X_1, ..., X_n$ are parametrul $\overline{g}(\beta,k)$. Pentru $k = 0$ definim $\overline{g}(\beta, 0) = 0$.

Cu aceste notaţii putem enunţa următorul rezultat.

Teorema 7.3. *Fie $\alpha, \beta \in (0,1)$. Atunci intervalul de încredere $(\overline{g}(\beta, S_n), g(\alpha, S_n))$ are un coeficient de încredere mai mare decât $1 - \alpha - \beta$, pentru $p \in (0,1)$.*

Demonstraţie. Are loc egalitatea

$$P_p(\overline{g}(\beta,S_n) < p < g(\alpha,S_n)) \geq P_p(\overline{g}(\beta,S_n) < p) + P_p(p < g(\alpha,S_n)) - 1$$

şi evaluăm separat cele două probabilităţi din membrul drept.

Începem prin a observa că inegalitatea $p < g(\alpha, l)$ este echivalentă cu $F(p, l) > \alpha$ şi atunci putem scrie

$$P_p(p < g(\alpha, S_n)) = P_p(F(p, S_n) > \alpha).$$

Pentru a estima ultima probabilitate vom introduce numărul

$$h = h(p, \alpha) = \inf\{k/F(p, k) > \alpha\},$$

definit pentru orice $p, \alpha \in (0, 1)$. Deoarece aplicaţia $k \to F(p, k)$ este crescătoare, rezultă că inegalitatea $F(p, l) > \alpha$ este echivalentă cu $l \geq h$ şi că, în caz că avem $h \geq 1$, are loc inegalitatea

$$F(p, h-1) \leq \alpha.$$

De aceea putem exprima probabilitatea de dinainte sub forma

$$P_p(p < g(\alpha, S_n)) = P_p(S_n \geq h),$$

care este 1 în cazul $h = 0$, iar dacă $h \geq 1$, această probabilitate devine

$$= 1 - P_p(S_n \leq h-1) = 1 - F(p, h-1) \geq 1 - \alpha.$$

Am obţinut, deci,

$$P_p(p < g(\alpha, S_n)) \geq 1 - \alpha. \quad (*)$$

Mai departe, trecem la evaluarea celeilalte probabilităţi care minorează expresia coeficientului de încredere. Inegalitatea $\overline{g}(\beta, l) < p$ este echivalentă cu

$$1 - \beta = F(\overline{g}(\beta, l), l-1) > F(p, l-1).$$

Facem convenţia $F(p, -1) = 0$ şi notăm

$$\overline{h} = \overline{h}(p, \beta) = \sup\{k/1 - F(p, k-1) > \beta\}$$

şi observăm că are loc inegalitatea $1 - \beta > F(p, k-1)$, pentru orice $k \leq \overline{h}$, iar dacă $\overline{h} \leq n-1$ pentru $k = \overline{h}+1$ avem $1 - \beta \leq F\left(p, \overline{h}\right)$. De aceea, putem spune că inegalitatea $\overline{g}(\beta, l) < p$ este echivalentă cu $l \leq \overline{h}$ şi putem exprima evenimentul ce ne interesează astfel:

$$\{\overline{g}(\beta, S_n) < p\} = \left\{S_n \leq \overline{h}\right\}.$$

Rezultă că

$$P_p(\overline{g}(\beta, S_n) < p) = P_p\left(S_n \leq \overline{h}\right) = F\left(p, \overline{h}\right)$$

şi această expresie face 1 dacă $\overline{h} = n$, iar dacă $\overline{h} \leq n-1$, atunci avem $F\left(p, \overline{h}\right) \geq 1 - \beta$. În orice caz, am dedus

$$P_p(\overline{g}(\beta, S_n) < p) \geq 1 - \beta.$$

Această inegalitate împreună cu (*) permit estimarea membrului drept al minorării coeficientului de încredere cu care am început:

$$P_p(\overline{g}(\beta, S_n) < p < g(\alpha, S_n)) \geq 1 - \alpha - \beta,$$

ceea ce termină demonstraţia. □

Întorcându-ne la exemplul 3. dinainte, pentru $\alpha = \beta = 0,005$, putem aplica teorema anterioară şi deducem că intervalul $(0,0288; 0,049)$ este un interval de încredere cu coeficientul de încredere mai mare decât $1 - 0,005 - 0,005 = 0,99$. Astfel, concluzia teoriei noastre este: cu un coeficient de încredere mai mare decât $0,99$ proporţia de produse defecte este în jurul lui $3,8\%$, cu o aproximaţie de cel mult $-0,92\%$ sau $+1,1\%$.

Privitor la aplicabilitatea ultimei teoreme, trebuie notat că este esenţial să se poată calcula uşor funcţiile g şi $\overline{g}$. În cazul exemplului 3. calculul a fost posibil prin determinarea probabilităţilor cumulate utilizând un PC. Dacă n este mult mai mare acest lucru poate deveni dificil.

În concluzie, putem spune că estimările statistice permit punerea în evidenţă a gradului de relevanţă pe care îl are o medie empirică pe care o face orice persoană neavizată. Rezultatele, intervalul de încredere şi coeficientul de încredere, nu dau informaţii sigure în mod absolut, ci numai cuantificări numerice pentru fenomene probabiliste. Pentru a se înţelege utilitatea, eficienţa practică a acestor rezultate, este nevoie de o experienţă a măsurării fenomenelor aleatorii. De exemplu, cine a aruncat cu banul de mai multe ori a înţeles că rezultatele experimentului nu sunt arbitrare. Ele sunt aleatorii dar respectă legile teoriei probabilităţilor.

Bibliografie

[1] J. Bertoin, Cours de Probabilités, Internet, 1999.
[2] P. Billingsley, Probability and Measure, ediţia a doua, John Wiley & Sons, New York - Chichester - Brisbane - Toronto - Singapore, 1986..
[3] K.L. Chung, Elementary Probability Theory with Stochastic Processes, ediţia a doua, Springer, New York - Heidelberg - Berlin, 1975.
[4] Gh. Ciucu, C. Tudor, Teoria Probabilităţilor şi Aplicaţii, Editura ştiinţifică şi enciclopedică, Bucureşti, 1983.
[5] I. Cuculescu, Teoria probabilităţilor, Editura All, Bucureşti, 1998.
[6] D. Dacunha -Castelle, M. Duflo, Probabilités et statistiques, tome 1, Masson, Paris - Milan - Barcelone - Mexico, 1990.
[7] W. Feller, An Introduction to Probability Theory and Its Applications, vol.1, ediţia a treia, John Wiley & Sons, New York - London - Sidney, 1968.
[8] D. Guinin, B. Joppin, Mathématiques, Terminale S, Bréal, Rosny, 1998.
[9] A. Gut, An Intermediate Course in Probability, Springer, New York -Berlin -Heidelberg..., 1995.
[10] M. Iosifescu, Gh. Mihoc, R. Theodorescu, Teoria Probabilităţilor şi Statistică Matematică, Editura tehnică, Bucureşti, 1966.
[11] J. Jacod, Cours de Probabilités, Internet, 1999.
[12] R.J. Larsen, M.L.Marx, Mathematical Statistics and Its Applications, Prentice - Hall, Englewood Cliffs, New Jersey, 1986.
[13] O. Onicescu, Calculul probabilităţilor, Editura tehnică, Bucureşti, 1956.
[14] O. Onicescu, Gh. Mihoc, Calculul Probabilităţilor, Bucureşti, 1939.
[15] O. Onicescu, Gh. Mihoc, C.T. Ionescu Tulcea, Calculul probabilităţilor şi aplicaţii, Editura Academiei, Bucureşti, 1956.
[16] E. Partzen, Modern Probability Theory and Its Applications, John Wiley & Sons, New York - London, 1960.
[17] J. Pitman, Probability, ediţia a şasea, Springer Texts in Statistics, Springer, 1997.
[18] A. Rényi, Probability Theory, Académiai Kiadó, Budhapest, 1970.
[19] D. Stirzaker, Probability and Random Variables, Cambridge University Press, 1999.
[20] C. Tudor, Teoria Probabilităţilor, Editura Universităţii din Bucureşti, 2004.
[21] J.V. Uspenski, Introduction to Mathematical Probability, McGrow-Hill Book Company, New York - Toronto - London, 1937.

Index