Curs 2

Cristian Niculescu

1 Probabilitatea condiţionată, independenţa şi teorema lui Bayes

1.1 Scopurile învăţării

1. Să ştie definiţiile probabilităţii condiţionate şi independenţei evenimentelor.
2. Să poată calcula probabilitatea condiţionată direct din definiţie.
3. Să poată folosi regula mutiplicării pentru a calcula probabilitatea totală a unui eveniment.
4. Să poată verifica dacă 2 evenimente sunt independente.
5. Să poată folosi formula lui Bayes pentru a ”inversa” probabilităţile condiţionate.
6. Să poată organiza calculul probabilităţilor condiţionate folosind arbori şi tabele.
7. Să înţeleagă complet eroarea ratei de bază.

1.2 Probabilitatea condiţionată

Probabilitatea condiţionată răspunde la întrebarea ”cum se schimbă probabilitatea unui eveniment dacă avem informaţie suplimentară”.
Exemplul 1. Aruncăm o monedă corectă de 3 ori.
(a) Care este probabilitatea a 3 aversuri?
Răspuns: Spaţiul probelor $\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$.
Toate cazurile sunt egal probabile, deci $P(3 \text{ aversuri}) = 1/8$.
(b) Presupunem că ni se spune că prima aruncare a fost avers. Fiind dată această informaţie, cum ar trebui să calculăm probabilitatea a 3 aversuri?
Răspuns: Avem un nou spaţiu al probelor (redus): $\Omega' = \{HHH, HHT, HTH, HTT\}$.
Toate cazurile sunt egal probabile, deci

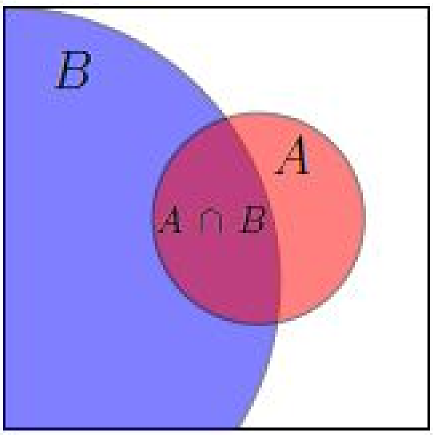
$$P(3 \text{ aversuri dat fiind că prima aruncare este avers}) = 1/4.$$

Aceasta este numită probabilitate condiţionată, deoarece ia în considerare condiţii suplimentare. Pentru a dezvolta notaţia, reformulăm (b) în termeni

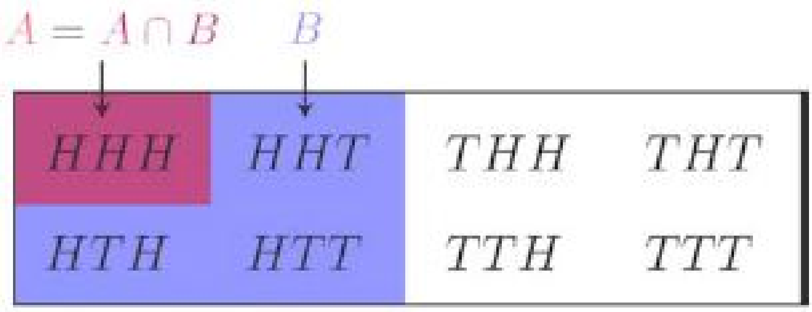
de *evenimente*.
(b) reformulat. Fie A evenimentul "toate 3 aruncările sunt aversuri" $= \{HHH\}$.
Fie B evenimentul "prima aruncare este avers" $= \{HHH, HHT, HTH, HTT\}$.
Probabilitatea condiţionată a lui A cunoscând că B a avut loc este scrisă

$$P(A|B).$$

Aceasta se citeşte "probabilitatea condiţionată a lui A dat fiind B" sau "probabilitatea lui A condiţionată de B" sau, mai simplu, "probabilitatea lui A dat fiind B".
Putem vizualiza probabilitatea condiţionată după cum urmează. Gândiţi $P(A)$ ca proporţia ariei ocupată de A din *tot* spaţiul probelor. Pentru $P(A|B)$ ne restrângem atenţia la B. Adică, $P(A|B)$ este proporţia ocupată de A din aria lui B, i.e. $P(A \cap B)/P(B)$.



Probabilitatea condiţionată: vizualizare abstractă



Probabilitatea condiţionată: exemplul cu moneda

Observăm că $A \subset B$ în ultima figură, deci apar doar 2 culori.
Definiţia formală a probabilităţii condiţionate prinde esenţa exemplului şi

vizualizărilor de mai sus.
Definiţia formală a probabilităţii condiţionate
Fie A şi B evenimente. Definim probabilitatea condiţionată a lui A dat fiind B ca

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \text{ dacă } P(B) \neq 0. \tag{1}$$

În continuare, când scriem o probabilitate condiţionată, presupunem implicit că probabilitatea evenimentului care condiţionează este nenulă, fără a mai scrie aceasta.
Să refacem exemplul cu aruncarea monezii folosind definiţia din relaţia (1). Reamintim că A = "3 aversuri" şi B = "prima aruncare este avers". Avem $P(A) = 1/8$ şi $P(B) = 1/2$. Deoarece $A \cap B = A$, avem de asemenea $P(A \cap B) = 1/8$. Acum, din (1), $P(A|B) = \frac{1/8}{1/2} = 1/4$, la fel ca răspunsul din exemplul 1b.

1.3 Regula multiplicării

Următoarea formulă este numită regula multiplicării.

$$P(A \cap B) = P(A|B) \cdot P(B). \tag{2}$$

Aceasta este doar o rescriere a definiţiei din relaţia (1) a probabilităţii condiţionate. Regula multiplicării este doar o versiune a regulii produsului.
Începem cu un exemplu simplu unde putem verifica direct toate probabilităţile numărând.
Exemplul 2. Tragem 2 cărţi fără revenire dintr-un pachet de cărţi de joc fără jokeri. Definim evenimentele: S_1 = "prima carte este o pică" şi S_2 = "a 2-a carte este o pică". Cât este $P(S_2|S_1)$?
Răspuns: Putem face asta direct prin numărare: dacă prima carte este o pică, atunci din cele 51 de cărţi rămase, 12 sunt pici.

$$P(S_2|S_1) = 12/51 = 4/17.$$

Acum, să recalculăm asta folosind formula (1). Trebuie să calculăm $P(S_1)$, $P(S_2)$ şi $P(S_1 \cap S_2)$. Ştim că $P(S_1) = 1/4$ deoarece sunt 52 de moduri egal posibile de a trage prima carte şi 13 din ele sunt pici. Aceeaşi logică spune că sunt 52 de moduri egal posibile de a trage a 2-a carte, deci $P(S_2) = 1/4$.
Comentariu: Probabilitatea $P(S_2) = 1/4$ poate părea surprinzătoare deoarece valoarea primei cărţi afectează cu siguranţă probabilităţile pentru a 2-a carte. Dar, dacă ne uităm la *toate* secvenţele posibile de 2 cărţi, vom vedea că fiecare carte din pachet are probabilitate egală de a fi a 2-a carte. Deoarece 13 din

cele 52 de cărţi sunt pici obţinem $P(S_2) = 13/52 = 1/4$. Un alt fel de a spune asta este: dacă nu avem dată valoarea pentru prima carte, atunci trebuie să considerăm toate posibilităţile pentru a 2-a carte.
Continuând, vedem că

$$P(S_1 \cap S_2) = \frac{13 \cdot 12}{52 \cdot 51} = 1/17.$$

(Aceasta a fost aflată numărând modurile de a trage o pică urmată de a 2-a pică şi împărţind prin numărul de moduri de a trege orice carte urmată de orice altă carte.) Acum, folosind (1) obţinem

$$P(S_2|S_1) = \frac{P(S_2 \cap S_1)}{P(S_1)} = \frac{1/17}{1/4} = 4/17.$$

În sfârşit, verificăm regula multiplicării calculând ambii membri din (2).

$$P(S_1 \cap S_2) = \frac{13 \cdot 12}{52 \cdot 51} = \frac{1}{17} \text{ şi } P(S_2|S_1) \cdot P(S_1) = \frac{4/17}{1/4} = \frac{1}{17}, \text{ q.e.d.}$$

Gândiţi: Pentru S_1 şi S_2 din exemplul precedent, cât este $P(S_2|S_1^c)$?

1.4 Legea probabilităţii totale

Legea probabilităţii totale ne permite să folosim regula multiplicării pentru a afla probabilităţi în exemple mai interesante. Necesită multă notaţie, dar ideea este destul de simplă. Enunţăm legea când spaţiul probelor este împărţit în 3 părţi. Este o chestiune simplă a extinde regula când sunt mai mult de 3 părţi.

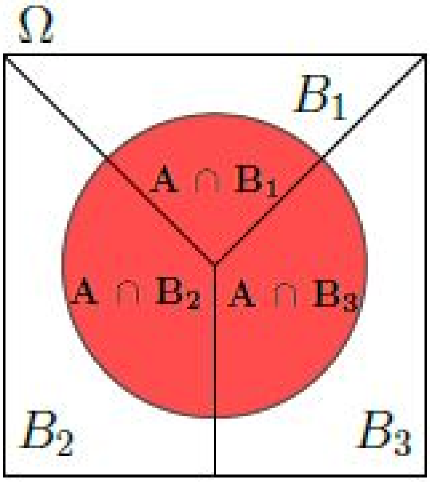
Legea probabilităţii totale

Presupunem că spaţiul probelor Ω este împărţit în 3 evenimente disjuncte 2 câte 2 B_1, B_2, B_3 (vezi figura de mai jos). Atunci pentru orice eveniment A:

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + P(A \cap B_3);$$

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + P(A|B_3)P(B_3). \qquad (3)$$

Prima relaţie spune "dacă A este împărţit în 3 părţi disjuncte 2 câte 2, atunci $P(A)$ este suma probabilităţilor părţilor". Relaţia (3) se numeşte legea probabilităţii totale. Ea este doar o rescriere a relaţiei de deasupra folosind regula multiplicării.



Spaţiul probelor Ω şi evenimentul A sunt fiecare împărţite în 3 părţi disjuncte 2 câte 2.

Legea este valabilă dacă împărţim Ω în orice număr de evenimente, atât timp cât ele sunt *disjuncte 2 câte 2* şi *acoperă* tot Ω (adică reuniunea lor este Ω). O astfel de împărţire este numită adesea o *partiţie* a lui Ω.
Primul nostru exemplu va fi unul unde ştim deja răspunsul şi putem verifica legea.
Exemplul 3. O urnă conţine 5 bile roşii şi 2 bile verzi. 2 bile sunt scoase una după alta, fără revenire. Care este probabilitatea ca a 2-a bilă să fie roşie?
Răspuns: Spaţiul probelor este $\Omega = \{rr, rg, gr, gg\}$. (Am pus ”r” pentru roşu şi ”g” pentru verde.)
Fie R_1 evenimentul ”prima bilă este roşie”, G_1 = ”prima bilă este verde”, R_2 = ”a 2-a bilă este roşie”, G_2 = ”a 2-a bilă este verde”. Ni se cere să aflăm $P(R_2)$.
Modul rapid de a calcula aceasta este la fel ca $P(S_2)$ din exemplul cu cărţi de joc de mai sus. Fiecare bilă este egal probabilă să fie a 2-a bilă. Deoarece 5 din 7 bile sunt roşii, $P(R_2) = 5/7$.
Să calculăm această valoare folosind legea probabilităţii totale (3). Întâi, aflăm probabilităţile condiţionate. Acesta este un exerciţiu simplu de numărare.

$$P(R_2|R_1) = 4/6 = 2/3, \quad P(R_2|G_1) = 5/6.$$

Deoarece R_1 şi G_1 partiţionează Ω, legea probabilităţii totale spune

$$\begin{aligned} P(R_2) &= P(R_2|R_1)P(R_1) + P(R_2|G_1)P(G_1) \qquad (4) \\ &= \frac{2}{3}\cdot\frac{5}{7} + \frac{5}{6}\cdot\frac{2}{7} = \frac{10}{21} + \frac{5}{21} = \frac{15}{21} = \frac{5}{7}. \end{aligned}$$

Urne de probabilitate
Exemplul precedent a folosit urne de probabilitate.
În probabilităţi şi statistică, o problemă a urnei este un exerciţiu mental idealizat în care unele obiecte de real interes (ca atomi, oameni, maşini, etc.) sunt reprezentate ca bile colorate într-o urnă sau alt container. Se pretinde a se extrage una sau mai multe bile din urnă; scopul este a se determina probabilitatea extragerii unei culori sau a alteia, sau alte proprietăţi. Un parametru cheie este dacă fiecare bilă este returnată în urnă sau nu după fiecare extragere.

Nu ne ia mult să facem un exemplu unde (4) este într-adevăr cel mai bun mod de a calcula probabilitatea. Iată un joc cu reguli puţin mai complicate.
Exemplul 4. O urnă conţine 5 bile roşii şi 2 bile verzi. O bilă este extrasă. Dacă este verde, o bilă roşie este adăugată în urnă, iar dacă este roşie, o bilă verde este adăugată în urnă. (Bila originală nu este returnată în urnă.) Apoi o a 2-a bilă este extrasă. Care este probabilitatea ca a 2-a bilă să fie roşie?
Răspuns. Cu notaţiile de la exemplul 3, legea probabilităţii totale spune că $P(R_2)$ poate fi calculată folosind relaţia (4). Doar valorile probabilităţilor condiţionate se schimbă. Avem

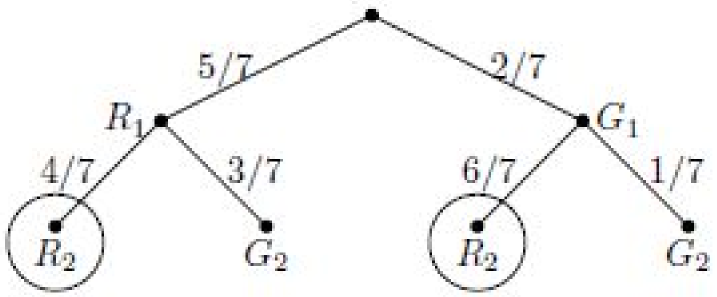
$$P(R_2|R_1) = 4/7, \quad P(R_2|G_1) = 6/7.$$

De aceea,

$$P(R_2) = P(R_2|R_1)P(R_1) + P(R_2|G_1)P(G_1) = \frac{4}{7} \cdot \frac{5}{7} + \frac{6}{7} \cdot \frac{2}{7} = \frac{32}{49}.$$

1.5 Utilizarea arborilor pentru a organiza calculul

Arborii sunt o metodă de a organiza calcule cu probabilităţi condiţionate şi legea probabilităţii totale. Ca la regula produsului, organizăm procesul de bază într-o secvenţă de acţiuni.
Începem refăcând exemplul 4. Secvenţa de acţiuni este: întâi extragem bila 1 (şi adăugăm bila corespunzătoare în urnă) şi apoi extragem bila 2.



Interpretăm arborele după cum urmează. Fiecare punct este numit nod. Arborele este organizat pe nivele. Nodul de la vârf (nodul rădăcină) este la nivelul 0. Următorul strat de mai jos este nivelul 1 şi aşa mai departe. Fiecare nivel arată cazurile la un stadiu al jocului. Nivelul 1 arată cazurile posibile la prima extragere. Nivelul 2 arată cazurile posibile la a 2-a extragere, plecând din fiecare nod din nivelul 1.
Probabilităţile sunt scrise de-a lungul ramurilor. Probabilitatea lui R_1 (bilă roşie la prima extragere) este $5/7$. Ea este scrisă de-a lungul ramurii de la nodul rădăcină la nodul etichetat R_1. La următorul nivel punem probabilităţile condiţionate. Probabilitatea de-a lungul ramurii de la R_1 la R_2 este $P(R_2|R_1) = 4/7$. Ea reprezintă probabilitatea de a merge în nodul R_2 dat fiind că suntem deja în R_1.
Regula multiplicării spune că probabilitatea de a ajunge în orice nod este tocmai produsul probabilităţilor aflate de-a lungul drumului de a ajunge acolo din nodul rădăcină. De exemplu, nodul etichetat R_2 din extrema stângă reprezintă în realitate evenimentul $R_1 \cap R_2$ deoarece vine din nodul R_1. Regula multiplicării spune acum

$$P(R_1 \cap R_2) = P(R_1) \cdot P(R_2|R_1) = \frac{5}{7} \cdot \frac{4}{7},$$

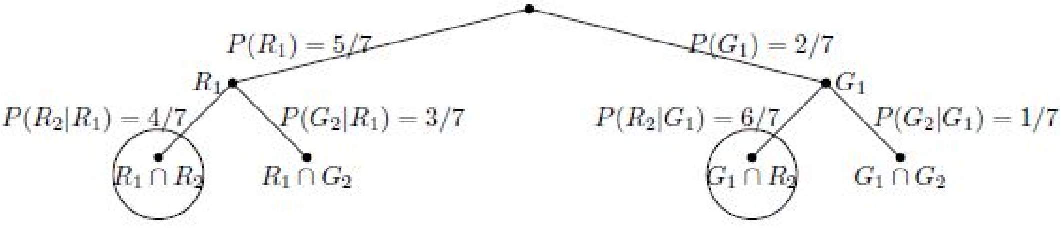
ceea ce este exact înmulţirea probabilităţilor aflate de-a lungul drumului de la nodul rădăcină la nodul R_2 din extrema stângă.
Legea probabilităţii totale este chiar afirmaţia că $P(R_2)$ este suma probabilităţilor tuturor drumurilor conducând de la nodul rădăcină la R_2 (cele 2 noduri încercuite din figură). În acest caz,

$$P(R_2) = \frac{5}{7} \cdot \frac{4}{7} + \frac{2}{7} \cdot \frac{6}{7} = \frac{32}{49},$$

exact ca în exemplul precedent.

1.5.1 Arbori prescurtaţi vs. precişi

Arborele dat mai sus implică o prescurtare. De exemplu, nodul marcat R_2 din extrema stângă reprezintă în realitate evenimentul $R_1 \cap R_2$, deoarece termină drumul de la rădăcină prin R_1 la R_2. Iată acelaşi arbore etichetat precis.



După cum se vede, acest arbore este mai greu de făcut. De obicei folosim versiunea prescurtată a arborilor.

1.6 Independenţă

2 evenimente sunt independente dacă informaţia că unul a avut loc nu schimbă probabilitatea ca celălalt să fi avut loc. Informal, evenimentele sunt independente dacă ele nu se influenţează unul pe celălalt.

Exemplul 5. Aruncăm o monedă de 2 ori. Ne aşteptăm ca rezultatele celor 2 aruncări să fie independente unul faţă de celălalt. În experimentele reale acest lucru trebuie totdeauna verificat. Dacă moneda mea aterizează în miere şi nu mă deranjez s-o curăţ, atunci a 2-a aruncare poate fi afectată de rezultatul primei aruncări.

Independenţa experimentelor poate fi subminată de eşecul curăţării sau recalibrării echipamentului între experimente, sau al izolării observatorilor presupuşi independenţi unul faţă de altul sau de o influenţă comună. Toţi am experimentat aflarea aceluiaşi "fapt" de la oameni diferiţi. Aflându-l din surse diferite tindem să-i dăm crezare până aflăm că toţi l-au aflat de la o sursă comună. Adică, sursele noastre nu erau independente.

Traducerea descrierii verbale în simboluri dă

$$A \text{ este independent de } B \iff P(A|B) = P(A). \tag{5}$$

Adică, cunoaşterea faptului că B a avut loc nu schimbă probabilitatea ca A să fi avut loc. În termeni de evenimente ca submulţimi, cunoaşterea faptului că rezultatul realizat este în B nu schimbă probabilitatea că el este în A.

Dacă A şi B sunt independente în sensul de mai sus, atunci regula multiplicării dă $P(A \cap B) = P(A|B) \cdot P(B) = P(A) \cdot P(B)$. Aceasta justifică următoarea definiţie tehnică a independenţei.

Definiţia formală a independenţei: 2 evenimente A şi B sunt independente $\iff$

$$P(A \cap B) = P(A) \cdot P(B). \tag{6}$$

Această definiţie simetrică clarifică faptul că A este independent de $B \iff B$ este independent de A. Spre deosebire de relaţia (5), această definiţie are sens şi când $P(B) = 0$. În termeni de probabilităţi condiţionate, avem:

1. Dacă $P(B) \neq 0$, atunci A şi B sunt independente $\iff P(A|B) = P(A)$.
2. Dacă $P(A) \neq 0$, atunci A şi B sunt independente $\iff P(B|A) = P(B)$.

Evenimentele independente apar de obicei ca probe diferite într-un experiment, ca în exemplul următor.

Exemplul 6. Aruncăm o monedă corectă de 2 ori. Fie $H_1 =$ "avers la prima

aruncare" şi $H_2 =$ "avers la a 2-a aruncare". Sunt H_1 şi H_2 independente?
Răspuns. Deoarece $H_1 \cap H_2$ este evenimentul "ambele aruncări sunt aversuri", avem

$$P(H_1 \cap H_2) = 1/4 = (1/2) \cdot (1/2) = P(H_1)P(H_2).$$

De aceea evenimentele H_1 şi H_2 sunt independente.

Putem întreba despre independenţa oricăror 2 evenimente, ca în următoarele 2 exemple.
Exemplul 7. Aruncăm o monedă corectă de 3 ori. Fie $H_1 =$ "avers la prima aruncare" şi $A =$ "2 aversuri în total". Sunt H_1 şi A independente?
Răspuns. Ştim că $P(A) = 3/8$. Deoarece $P(H_1) = 1/2 \neq 0$, putem verifica dacă formula din relaţia (5) este îndeplinită. $H_1 = \{HHH, HHT, HTH, HTT\}$ conţine exact 2 cazuri (HHT, HTH) din A, deci avem $P(A|H_1) = 2/4 = 1/2$. Deoarece $P(A|H_1) \neq P(A)$, evenimentele H_1 şi A nu sunt independente.
Exemplul 8. Tragem o carte dintr-un pachet de cărţi de joc standard (adică fără jokeri). Să examinăm independenţa 2 câte 2 a 3 evenimente: "cartea este as", "cartea este cupă" şi "cartea este roşie".
Notăm evenimentele $A =$ "as", $H =$ "cupă", $R =$ "roşie".
a) Ştim că $P(A) = 4/52 = 1/13, P(A|H) = 1/13$. Deoarece $P(A) = P(A|H)$, A este independent de H.
b) $P(A|R) = 2/26 = 1/13 = P(A)$. Deci A este independent de R.
c) În sfârşit, ce putem spune despre H şi R? Deoarece $P(H) = 13/52 = 1/4$ şi $P(H|R) = 13/26 = 1/2$, avem $P(H) \neq P(H|R)$, deci H şi R nu sunt independente. Puteam de asemenea vedea asta invers: $P(R) = 26/52 = 1/2$ şi $P(R|H) = 13/13 = 1$, deci $P(R) \neq P(R|H)$, de aceea H şi R nu sunt independente.

1.6.1 Paradoxurile independenţei

Un eveniment A cu probabilitatea 0 este independent de el însuşi, deoarece în acest caz ambii membri ai relaţiei (6) sunt 0. Aceasta apare paradoxal, deoarece cunoaşterea faptului că A a avut loc sigur dă informaţie despre faptul dacă A a avut loc. Rezolvăm paradoxul observând că deoarece $P(A) = 0$, afirmaţia "A a avut loc" este fără sens.
Gândiţi: Pentru ce altă valoare a lui $P(A)$ este A independent de el însuşi?

1.7 Teorema lui Bayes

Teorema lui Bayes este foarte importantă pentru probabilităţi şi statistică. Pentru 2 evenimente A şi B, teorema lui Bayes (de asemenea numită regula lui Bayes şi formula lui Bayes) spune

$$P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A)}. \tag{7}$$

Comentarii: 1. Regula lui Bayes ne spune cum să "inversăm" probabilităţile condiţionate, i.e. să aflăm $P(B|A)$ din $P(A|B)$.
2. În practică, $P(A)$ este adesea calculată folosind legea probabilităţii totale.
Demonstraţia regulii lui Bayes
Punctul cheie este că $A \cap B$ este simetrică în A şi B. Deci regula multiplicării spune

$$P(B|A) \cdot P(A) = P(A \cap B) = P(A|B) \cdot P(B).$$

Acum împărţim cu $P(A)$ pentru a obţine regula lui Bayes.

O greşeală obişnuită este a confunda $P(A|B)$ cu $P(B|A)$. Ele pot fi foarte diferite. Aceast fapt este ilustrat în următorul exemplu.
Exemplul 9. Aruncăm o monedă corectă de 5 ori. Fie H_1 = "prima aruncare este avers" şi H_A = "toate cele 5 aruncări sunt aversuri". Atunci $P(H_1|H_A) = 1$, dar $P(H_A|H_1) = 1/16$.
Pentru practică, să folosim teorema lui Bayes pentru a calcula $P(H_1|H_A)$ din $P(H_A|H_1)$. Avem $P(H_A|H_1) = 1/16, P(H_1) = 1/2, P(H_A) = 1/32$. Deci,

$$P(H_1|H_A) = \frac{P(H_A|H_1)P(H_1)}{P(H_A)} = \frac{(1/16) \cdot (1/2)}{1/32} = 1,$$

de acord cu calculul nostru anterior.

1.7.1 Eroarea ratei de bază

Eroarea ratei de bază este unul din multele exemple care arată că este uşor de confundat semnificaţia lui $P(B|A)$ cu $P(A|B)$ când o situaţie este descrisă în cuvinte.
Exemplul 10. Eroarea ratei de bază
Considerăm un test de rutină pentru o boală. Presupunem că frecvenţa bolii în populaţie (rata de bază) este 0.5%. Testul are o rată "fals pozitiv" de 5% şi o rată "fals negativ" de 10%.
Faceţi testul şi iese pozitiv. Care este probabilitatea să aveţi boala?
Răspuns. Vom face calculul de 3 ori: folosind arbori, tablele şi simboluri. Vom utiliza următoarele notaţii pentru evenimentele relevante:
$D+$ = "aveţi boala",
$D-$ = "nu aveţi boala",
$T+$ = "aţi fost testat pozitiv",

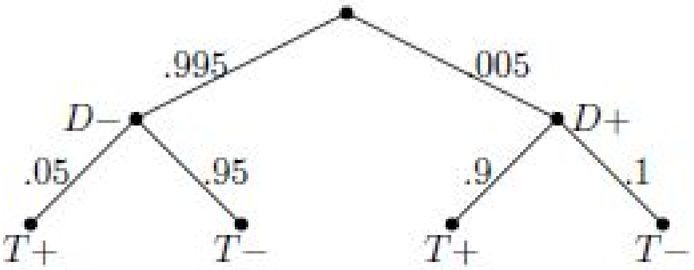
$T-$ = "aţi fost testat negativ".
Avem dat $P(D+) = 0.005$ şi de aceea $P(D-)$=0.995. Ratele "fals pozitiv" şi "fals negativ" sunt (prin definiţie) probabilităţi condiţionate.

$P(\text{fals pozitiv}) = P(T+|D-) = 0.05$ şi $P(\text{fals negativ}) = P(T-|D+) = 0.1$.

Probabilităţile complementare sunt cunoscute ca ratele "adevărat negativ" şi "adevărat pozitiv":

$P(T-|D-) = 1-P(T+|D-) = 0.95$ şi $P(T+|D+) = 1-P(T-|D+) = 0.9$.

Arbori: Toate aceste probabilităţi pot fi prezentate într-un arbore.



Se cere probabilitatea să aveţi boala dat fiind că aţi fost testat pozitiv, i.e. $P(D+|T+)$. Nu avem dată această valoare, dar ştim $P(T+|D+)$, deci putem utiliza teorema lui Bayes.

$$P(D+|T+) = \frac{P(T+|D+) \cdot P(D+)}{P(T+)}.$$

Cele 2 probabilităţi de la numărător sunt date. Calculăm numitorul folosind legea probabilităţii totale. Folosind arborele, trebuie doar să adunăm probabilităţile pentru fiecare din nodurile etichetate $T+$

$$P(T+) = 0.995 \cdot 0.05 + 0.005 \cdot 0.9 = 0.05425.$$

Astfel,

$$P(D+|T+) = \frac{0.9 \cdot 0.005}{0.05425} \approx 0.08294931 \approx 8.3\%.$$

Observaţii. Aceasta este numită eroarea ratei de bază deoarece rata de bază a bolii în populaţie este atât de mică încât marea majoritate a oamenilor care fac testul sunt sănătoşi şi chiar cu un test precis majoritatea celor testaţi pozitiv vor fi oameni sănătoşi.
Pentru a rezuma eroarea ratei de bază cu numere particulare
faptul că 95% din toate testele sunt precise nu implică faptul că 95% din testele pozitive sunt precise.

Alte moduri de a lucra exemplul 10

Tabele: Alt truc util pentru a calcula probabilităţi este a face un tabel. Să refacem exemplul precedent folosind un tabel construit pentru un total de 10000 de oameni împărţiţi conform probabilităţilor din acest exemplu.

Construim tabelul după cum urmează. Alegem un număr, să zicem 10000 de oameni, şi îl plasăm ca marele total în dreapta jos. Folosind $P(D+) = 0.005$ calculăm că 50 din cei 10000 de oameni sunt bolnavi ($D+$). Analog 9950 de oameni sunt sănătoşi. În acest moment tabelul arată aşa:

	$D+$	$D-$	total
$T+$			
$T-$			
total	50	9950	10000

Folosind $P(T+|D+) = 0.9$ putem calcula că numărul oamenilor bolnavi care au fost testaţi pozitiv este 90% din 50, adică 45. Folosind $P(T+|D-) = 0.05$ putem calcula că numărul oamenilor sănătoşi care au fost testaţi pozitiv este 5% din 9950, adică aproximativ 498. În acest moment tabelul arată aşa:

	$D+$	$D-$	total
$T+$	45	498	
$T-$	5	9452	
total	50	9950	10000

În sfârşit, adunăm elementele de pe liniile $T+$ şi $T-$ pentru a completa coloana din dreapta.

	$D+$	$D-$	total
$T+$	45	498	543
$T-$	5	9452	9457
total	50	9950	10000

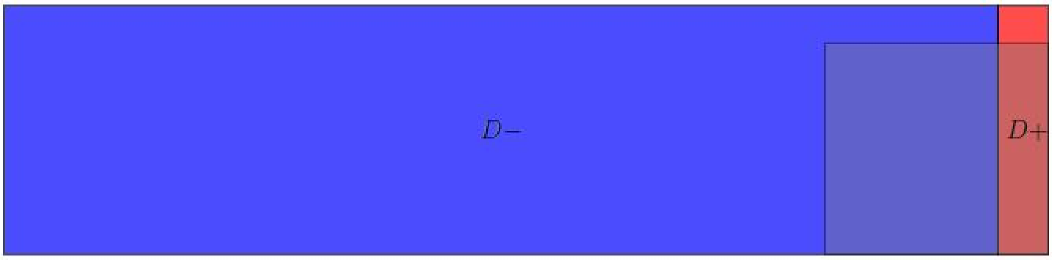
Folosind tabelul complet putem calcula

$$P(D+|T+) = \frac{|D+\cap T+|}{|T+|} = \frac{45}{543} \approx 8.3\%.$$

Simboluri: Pentru completitudine, arătăm soluţia scrisă direct cu simboluri.

$$\begin{aligned} P(D+|T+) &= \frac{P(T+|D+) \cdot P(D+)}{P(T+)} \\ &= \frac{P(T+|D+) \cdot P(D+)}{P(T+|D+) \cdot P(D+) + P(T+|D-) \cdot P(D-)} \\ &= \frac{0.9 \cdot 0.005}{0.9 \cdot 0.005 + 0.05 \cdot 0.995} \\ &\approx 8.3\%. \end{aligned}$$

Vizualizare: Figura de mai jos ilustrează eroarea ratei de bază. Aria albastră mare reprezintă toţi oamenii sănătoşi. Aria roşie mult mai mică reprezintă oamenii bolnavi. Dreptunghiul mai închis reprezintă oamenii care au fost testaţi pozitiv. Aria mai închisă acoperă cea mai mare parte a ariei roşii şi doar o mică parte din aria albastră. Chiar şi aşa, cea mai mare parte a ariei mai închise este peste albastru. Adică, cele mai multe teste pozitive sunt ale oamenilor sănătoşi.



2 Variabile aleatoare discrete

2.1 Scopurile învăţării

1. Să ştie definiţia unei variabile aleatoare discrete.
2. Să ştie distribuţiile (repartiţiile) Bernoulli, binomială şi geometrică şi exemple de ce modelează ele.
3. Să poată să descrie funcţia masă de probabilitate şi funcţia de distribuţie cumulativă folosind tabele şi formule.
4. Să poată să construiască noi variabile aleatoare din cele vechi.
5. Să ştie cum să calculeze media.

2.2 Variabile aleatoare

Cuvintele cheie sunt
1. Variabilă aleatoare

2. Funcţie masă de probabilitate (pmf)
3. Funcţie de distribuţie cumulativă (cdf), numită şi funcţie de repartiţie.

2.2.1 Recapitulare

Un spaţiu al probelor discret Ω este o mulţime finită sau numărabilă de cazuri (rezultate) $\{\omega_1, \omega_2, ...\}$. Probabilitatea cazului ω este notată $P(\omega)$.
Un eveniment E este o submulţime a lui Ω. Probabilitatea evenimentului E este $P(E) = \sum_{\omega \in E} P(\omega)$.

2.2.2 Variabilele aleatoare ca funcţii de plată

Exemplul 1. Un joc cu 2 zaruri corecte.
Aruncăm 2 zaruri corecte şi înregistrăm cazurile ca (i, j), unde i este rezultatul primului zar şi j rezultatul celui de-al 2-lea. Putem lua spaţiul probelor

$$\Omega = \{(1,1), (1,2), (1,3), ..., (6,6)\} = \{(i,j)|i,j = 1, ..., 6\}.$$

Funcţia de probabilitate este $P(i, j) = 1/36$.
În acest joc, câştigăm 500 de dolari dacă suma este 7 şi pierdem 100 altfel.
Dăm acestei funcţii de plată numele X şi o descriem formal prin

$$X(i,j) = \begin{cases} 500 & \text{dacă } i + j = 7 \\ -100 & \text{dacă } i + j \neq 7. \end{cases}$$

Exemplul 2. Putem schimba jocul folosind o funcţie de plată diferită. De exemplu

$$Y(i,j) = ij - 10.$$

În acest exemplu, dacă dăm (6,2), atunci câştigăm 2 dolari. Dacă dăm (2,3), atunci câştigăm -4 dolari (adică pierdem 4 dolari).

Aceste funcţii de plată sunt exemple de variabile aleatoare. O variabilă aleatoare atribuie un număr fiecărui caz dintr-un spaţiu al probelor. Mai formal:
Definiţie. Fie Ω un spaţiu al probelor. O variabilă aleatoare discretă este o funcţie

$$X : \Omega \to \mathbb{R},$$

care ia o mulţime discretă de valori. (Reamintim că $\mathbb{R}$ este mulţimea numerelor reale.)
De ce este X numită variabilă aleatoare? Este "aleatoare" deoarece valoarea ei depinde de un caz aleator al unui experiment. Şi tratăm pe X ca pe o variabilă uzuală: putem s-o adunăm la alte variabile aleatoare, s-o ridicăm la pătrat, ş.a.m.d.

2.2.3 Evenimente şi variabile aleatoare

Pentru orice valoare a scriem $X = a$ pentru evenimentul format din toate cazurile ω cu $X(\omega) = a$.
Exemplul 3. În exemplul 1 am aruncat 2 zaruri corecte şi X a fost variabila aleatoare

$$X(i,j) = \begin{cases} 500 & \text{dacă } i+j=7 \\ -100 & \text{dacă } i+j \neq 7. \end{cases}$$

Evenimentul $X = 500$ este mulţimea $\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$, i.e. mulţimea tuturor cazurilor cu suma componentelor 7. $P(X = 500) = 6/36 = 1/6$.
Permitem lui a să ia orice valoare, chiar şi valori pe care X nu le ia niciodată. În exemplul 1, am putea considera evenimentul $X = 1000$. Deoarece X nu este egal niciodată cu 1000, acesta este tocmai evenimentul vid (sau mulţimea vidă).
"$X = 1000$" $= \emptyset$.
$P(X = 1000) = 0$.

2.2.4 Funcţia masă de probabilitate

Devine obositor şi greu să citim şi să scriem $P(X = a)$ pentru probabilitatea lui $X = a$. Când ştim că vorbim despre X vom scrie simplu $p(a)$. Dacă vrem să-l facem explicit pe X vom scrie $p_X(a)$.
Definiţie. Funcţia masă de probabilitate (pmf) a unei variabile aleatoare discrete este funcţia $p : \mathbb{R} \to \mathbb{R}, p(a) = P(X = a)$.
Observăm:
1. Avem totdeauna $0 \leq p(a) \leq 1$.
2. Permitem lui a să fie orice număr real. Dacă a este o valoare pe care X n-o ia niciodată, atunci $p(a) = 0$.
Exemplul 4. Fie Ω spaţiul probelor de mai devreme pentru aruncarea a 2 zaruri corecte. Definim variabila aleatoare M ca valoarea maximă a celor 2 zaruri:

$$M(i,j) = \max(i,j).$$

De exemplu, aruncarea (3,5) are maximul 5, i.e. $M(3,5) = 5$.
Putem descrie o variabilă aleatoare listând valorile ei posibile şi probabilităţile asociate acestor valori. Pentru exemplul de mai sus avem:

valoarea	a :	1	2	3	4	5	6
pmf	$p(a)$:	1/36	1/12	5/36	7/36	1/4	11/36

Pe scurt, scriem

$$M \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1/36 & 1/12 & 5/36 & 7/36 & 1/4 & 11/36 \end{pmatrix}$$

De exemplu, $p(2) = 1/12$.
Întrebare: Cât este $p(8)$?
Răspuns: $p(8) = 0$.
Gândiţi: Care este pmf pentru $Z(i, j) = i + j$? Arată familiar?

2.2.5 Evenimente şi inegalităţi

Inegalităţile cu variabile aleatoare descriu evenimente. De exemplu, $X \leq a$ este mulţimea tuturor cazurilor ω a.î. $X(\omega) \leq a$.
Exemplul 5. Dacă spaţiul probelor este mulţimea tuturor perechilor (i, j) provenind din aruncarea a 2 zaruri corecte şi $Z(i, j) = i + j$ este suma zarurilor, atunci

$$Z \leq 4 = \{(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (3, 1)\}.$$

2.2.6 Funcţia de distribuţie cumulativă (cdf)

Definiţie. Funcţia de distribuţie cumulativă (cdf) a unei variabile aleatoare X este funcţia $F : \mathbb{R} \to \mathbb{R}$ dată de $F(a) = P(X \leq a)$. Adesea vom prescurta aceasta funcţia de repartiţie.
Observăm că definiţia lui $F(a)$ utilizează simbolul mai mic sau egal. Acest fapt va fi important pentru a face calculele exact.
Exemplu. Continuând exemplul cu M, avem

valoarea	a :	1	2	3	4	5	6
pmf	$p(a)$:	1/36	1/12	5/36	7/36	1/4	11/36
cdf	$F(a)$:	1/36	1/9	1/4	4/9	25/36	1

F e numită funcţie de distribuţie cumulativă deoarece $F(a)$ dă totalul probabilităţii care se acumulează adunând probabilităţile $p(b)$ când b merge de la $-\infty$ la a. De exemplu, în tabelul de mai sus, elementul 4/9 din coloana 4 pentru cdf este suma valorilor pmf de la coloana 1 la coloana 4. Scriem:

$$\text{"}M \leq 4\text{"} = \{(i, j) \in \Omega | M(i, j) \leq 4\}; F(4) = P(M \leq 4) = 1/36+1/12+5/36+7/36 = 4/9.$$

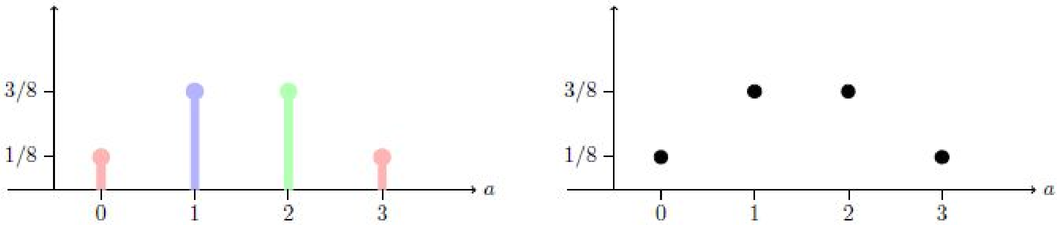
La fel ca funcţia masă de probabilitate, $F(a)$ este definită pentru toate valorile reale a. În exemplul de mai sus, $F(8) = 1, F(-2) = 0, F(2.5) = 1/9$ şi $F(\pi) = 1/4$.

2.2.7 Graficele lui p şi F

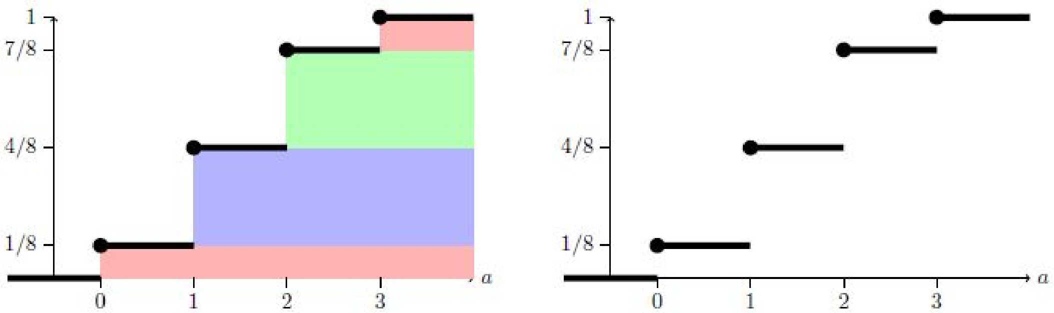
Putem vizualiza pmf şi cdf cu grafice. De exemplu, fie X numărul de aversuri din 3 aruncări ale unei monede corecte:

valoarea a :	0	1	2	3
pmf $p(a)$:	1/8	3/8	3/8	1/8
cdf $F(a)$:	1/8	1/2	7/8	1

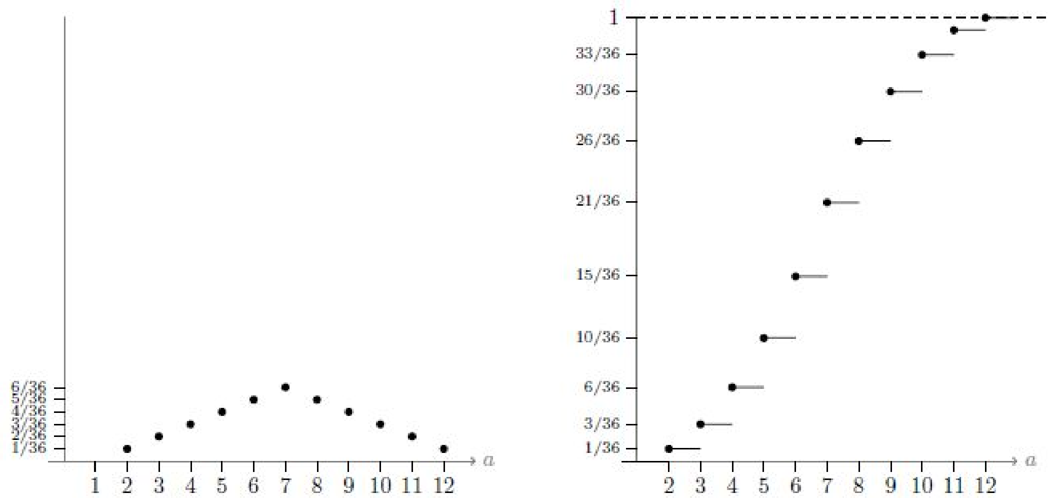
Graficele colorate arată cum funcţia de distribuţie cumulativă este construită prin acumularea probabilităţii când a creşte. Graficele alb-negru sunt prezentări standard.



Funcţia masă de probabilitate pentru X



Funcţia de distribuţie cumulativă a lui X



Pmf şi cdf pentru suma a 2 zaruri corecte (Exemplul 5)

2.2.8 Proprietăţi ale cdf F

Cdf F a unei variabile aleatoare satisface câteva proprietăţi:
1. F este crescătoare. Adică, graficul ei niciodată nu se duce în jos, sau, simbolic, dacă $a \leq b$, atunci $F(a) \leq F(b)$.
2. $0 \leq F(a) \leq 1, \forall a \in \mathbb{R}$.
3. $\lim\limits_{a \to -\infty} F(a) = 0, \quad \lim\limits_{a \to \infty} F(a) = 1$.

În cuvinte, (1) spune că $F(a)$ creşte sau rămâne constantă când a creşte, dar niciodată nu descreşte; (2) spune că probabilitatea acumulată este totdeauna între 0 şi 1; (3) spune că atunci când a ajunge negativ foarte mic, devine din ce în ce mai sigur faptul că $X > a$, iar când a ajunge foarte mare, devine din ce în ce mai sigur faptul că $X \leq a$.
Gândiţi: De ce o cdf satisface fiecare dintre aceste proprietăţi?

2.3 Repartiţii (distribuţii) clasice

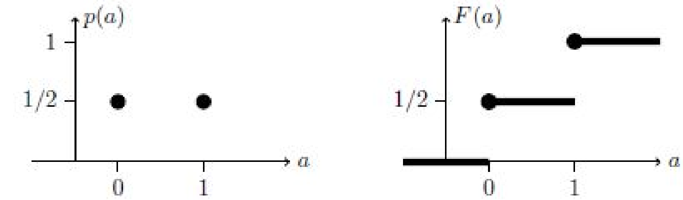
2.3.1 Repartiţii Bernoulli

Modelul: Repartiţia Bernoulli modelează o probă într-un experiment care poate avea ca rezultat ori succes ori eşec. O variabilă aleatoare X are o repartiţie Bernoulli cu parametrul $p \iff$
1. X ia valorile 0 şi 1.
2. $P(X = 1) = p$ şi $P(X = 0) = 1 - p$.

Vom scrie $X \sim \text{Bernoulli}(p)$ sau $\text{Ber}(p)$, ceea ce se citeşte "X are o repartiţie Bernoulli cu parametrul p".
Un model simplu pentru repartiţia Bernoulli este aruncarea unei monede cu

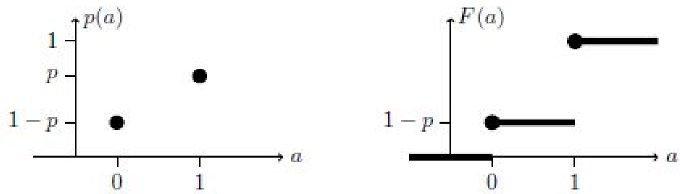
probabilitatea p a aversului, cu $X = 1$ pe avers şi $X = 0$ pe revers. Terminologia generală este a spune că X este 1 la succes şi 0 la eşec, cu succesul şi eşecul definite de context.
Multe decizii pot fi modelate ca o alegere binară, cum ar fi voturile pentru sau împotriva unei propuneri. Dacă p este proporţia din populaţia care votează care favorizează propunerea, atunci votul unui individ aleator este modelat prin Bernoulli(p).
Iată tabelul şi graficele pentru pmf şi cdf ale repartiţiei Bernoulli(1/2) şi mai jos cele pentru repartiţia generală Bernoulli(p).

value a:	0	1
pmf $p(a)$:	1/2	1/2
cdf $F(a)$:	1/2	1



Tabelul, pmf şi cdf pentru repartiţia Bernoulli(1/2)

values a:	0	1
pmf $p(a)$:	1-p	p
cdf $F(a)$:	1-p	1



Tabelul, pmf şi cdf pentru repartiţia Bernoulli(p)

2.3.2 Repartiţii binomiale

Repartiţia binomială Binomial(n, p), sau Bin(n, p) modelează numărul de succese din n probe independente Bernoulli(p).
Aici este o ierarhie. O singură probă Bernoulli este, să zicem, o aruncare a unei monede. O singură probă binomială constă din n probe Bernoulli. Pentru aruncări de monede, spaţiul probelor Bernoulli este $\{H, T\}$. Spaţiul probelor binomiale este format din toate secvenţele de aversuri şi reversuri de lungime n. O variabilă aleatoare Bernoulli ia valorile 0 şi 1. O variabilă aleatoare binomială ia valorile $0, 1, 2, ..., n$.
Exemplul 6. Binomial($1, p$) este la fel ca Bernoulli(p).
Exemplul 7. Numărul de aversuri din n aruncări ale unei monede cu probabilitatea p a aversului are o repartiţie Binomial(n, p).
Descriem $X \sim$ Binomial(n, p) dând valorile şi probabilităţile ei. Considerăm $k \in \{0, 1, ..., n\}$.
Reamintim că numărul de moduri de a alege k obiecte dintr-o colecţie de n

obiecte este ”combinări de n luate câte k” $= C_n^k$ şi are formula

$$C_n^k = \frac{n!}{k!(n-k)!}. \tag{8}$$

(Este numit şi coeficient binomial.) Iată un tabel pentru pmf a unei variabile aleatoare Binomial(n, k).

valorile a :	0	1	2	...	k	...	n
pmf $p(a)$:	$(1-p)^n$	$C_n^1 p^1 (1-p)^{n-1}$	$C_n^2 p^2 (1-p)^{n-2}$	...	$C_n^k p^k (1-p)^{n-k}$	...	p^n

Exemplul 8. Care este probabilitatea a 3 sau mai multe aversuri în 5 aruncări ale unei monede corecte?
Răspuns. Coeficienţii binomiali asociaţi cu $n = 5$ sunt

$$C_5^0 = 1,\ C_5^1 = \frac{5!}{1!4!} = 5,\ C_5^2 = \frac{5!}{2!3!} = \frac{5 \cdot 4}{2} = 10,$$

şi analog

$$C_5^3 = 10,\ C_5^4 = 5,\ C_5^5 = 1.$$

Folosind aceste valori obţinem următorul tabel pentru $X \sim$ Binomial$(5, p)$.

valorile a :	0	1	2	3	4	5
pmf $p(a)$:	$(1-p)^5$	$5p(1-p)^4$	$10p^2(1-p)^3$	$10p^3(1-p)^2$	$5p^4(1-p)$	p^5

Ni s-a spus că $p = 1/2$, deci

$$P(X \geq 3) = 10\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^2 + 5\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^1 + \left(\frac{1}{2}\right)^5 = \frac{16}{32} = \frac{1}{2}.$$

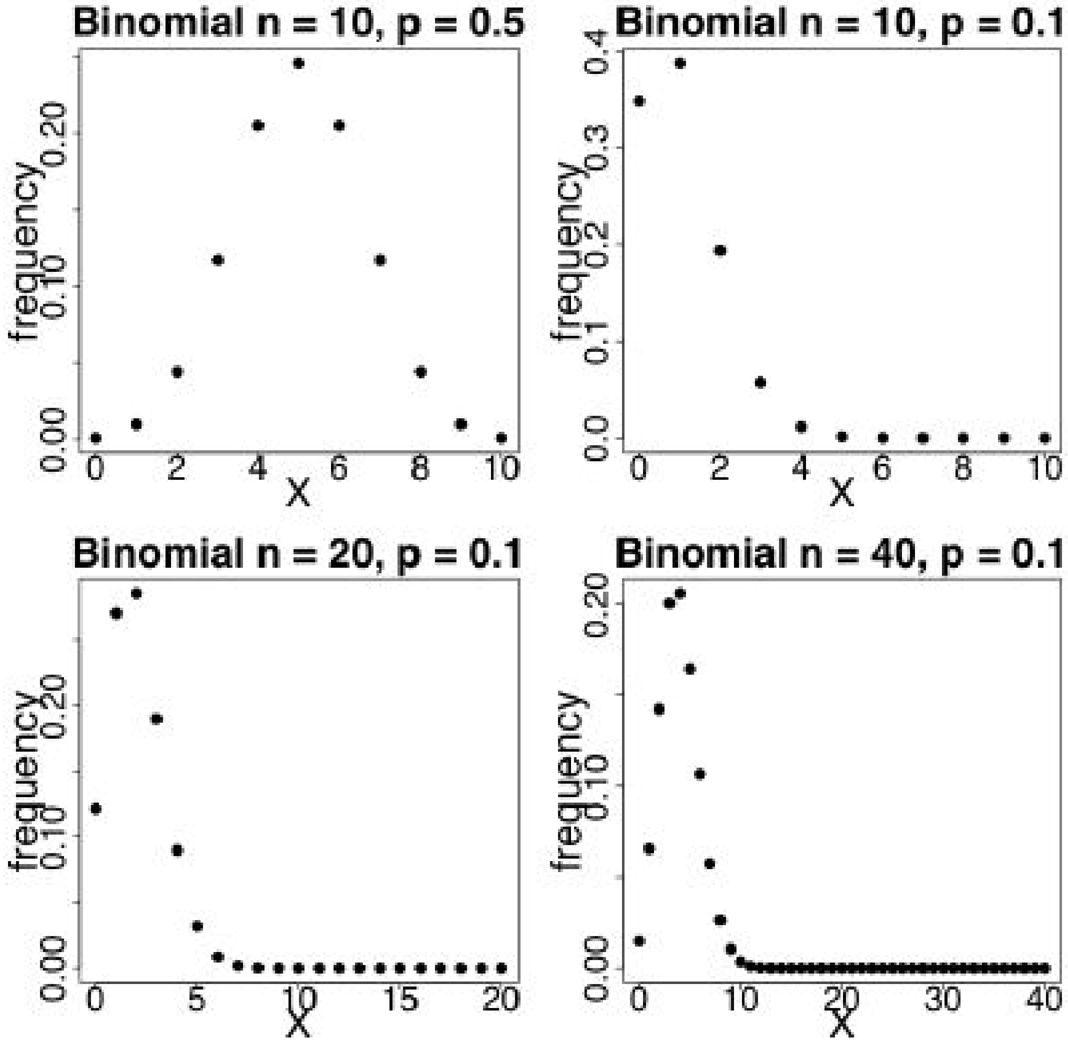
2.3.3 Explicarea probabilităţilor binomiale

Pentru concreteţe, fie $n = 5$ şi $k = 2$ (argumentul pentru n şi k arbitrari este identic). Deci $X \sim$ Binomial$(5, p)$ şi vrem să calculăm $p(2)$. Calea lungă de a calcula $p(2)$ este a lista toate modurile de a obţine exact 2 aversuri în 5 aruncări de monedă şi a aduna probabilităţile lor. Lista are 10 elemente:
HHTTT, HTHTT, HTTHT, HTTTH, THHTT, THTHT, THTTH, TTHHT, TTHTH, TTTHH.
Fiecare element are aceeaşi probabilitate de a avea loc, şi anume

$$p^2(1-p)^3.$$

Aceasta este deoarece fiecare dintre cele 2 aversuri are probabilitatea p şi fiecare dintre cele 3 reversuri are probabilitatea $1 - p$. Deoarece aruncările

individuale sunt independente, putem înmulţi probabilităţile. De aceea, probabilitatea totală a exact 2 aversuri este suma celor 10 probabilităţi identice, i.e. $p(2) = 10p^2(1-p)^3$, aşa cum se arată în tabel.
Aceasta ne ghidează la calea scurtă de a face calculul. Trebuie să calculăm numărul de secvenţe cu exact 2 aversuri. Pentru a face aceasta avem nevoie să alegem 2 dintre aruncări să fie aversuri şi cele 3 rămase vor fi reversuri. Numărul de astfel de secvenţe este numărul de moduri de a alege 2 din 5 obiecte, i.e. C_5^2. Deoarece fiecare astfel de secvenţă are aceeaşi probabilitate, $p^2(1-p)^3$, obţinem probabilitatea a exact 2 aversuri $p(2) = C_5^2 p^2(1-p)^3$.
Iată câteva pmf binomiale (aici, "frequency" înseamnă "probabilitate").



2.3.4 Repartiţii geometrice

O repartiţie geometrică modelează numărul de reversuri dinaintea primului avers într-o secvenţă de aruncări de monedă (probe Bernoulli).

Exemplul 9. (a) Aruncăm repetat o monedă. Fie X numărul de reversuri dinaintea primului avers. Deci, X poate fi 0, i.e. prima aruncare este avers, 1, 2, ... În principiu poate lua orice valoare naturală.

(b) Dăm unui revers valoarea 0 şi unui avers valoarea 1. În acest caz, X este numărul de 0-uri dinainte de primul 1.

(c) Dăm unui revers valoarea 1 şi unui avers valoarea 0. În acest caz, X este numărul de 1-uri dinainte de primul 0.

(d) Numim reversul succes şi aversul eşec. Astfel, X este numărul de succese dinaintea primului eşec.

(e) Numim reversul eşec şi aversul succes. Astfel, X este numărul de eşecuri dinaintea primului succes.

Putem vedea că aceasta modelează multe scenarii diferite de acest tip.

Definiţia formală. Variabila aleatoare X are o repartiţie geometrică cu parametrul p $\Longleftrightarrow$

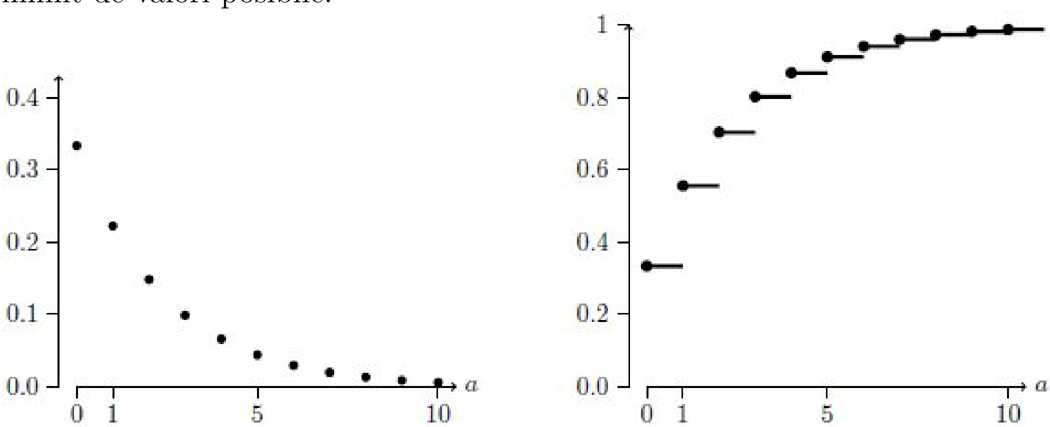
- X ia valorile 0,1,2,3,...
- are pmf $p(k) = P(X = k) = p(1-p)^k$.

Notăm $X \sim \text{geometric}(p)$ sau $\text{geo}(p)$. În formă de tabel avem:

valoarea a	0	1	2	3	...	k	...
pmf $p(a)$	p	$p(1-p)$	$p(1-p)^2$	$p(1-p)^3$	...	$p(1-p)^k$	...

Tabel: $X \sim \text{geometric}(p)$: X = numărul de 0-uri dinaintea primului 1.

Repartiţia geometrică este un exemplu de repartiţie discretă care ia un număr infinit de valori posibile.



Pmf şi cdf pentru repartiţia geometric(1/3)

Exemplul 10. Calculul probabilităţilor geometrice. Presupunem că locuitorii unei insule îşi planifică familiile având copii până este născută prima fată. Presupunem că probabilitatea de a avea o fată este de 0.5 la fiecare sarcină, independent de alte sarcini, că toţi copiii supravieţuiesc, că orice familie poate face oricât de mulţi copii şi că nu sunt gemeni. Care este probabilitatea ca o familie să aibă k băieţi?
Răspuns. Numărul de băieţi dintr-o familie este numărul de băieţi dinaintea primei fete.
Fie X numărul de băieţi dintr-o familie (aleasă aleator). X este o variabilă aleatoare geometrică. Ni se cere să aflăm $p(k) = P(X = k)$. O familie are k băieţi dacă secvenţa de copii din familie de la cel mai mare la cel mai mic este

$$BBB...BF,$$

primii k copii fiind băieţi. Probabilitatea acestei secvenţe este chiar produsul probabilităţilor pentru fiecare copil, i.e. $(1/2)^k \cdot (1/2) = (1/2)^{k+1}$. (Notă: Cele 4 presupuneri sunt simplificări ale realităţii.)
Mai multă confuzie geometrică. Altă definiţie pentru repartiţia geometrică este numărul de aruncări până la primul avers. În acest caz X poate lua valorile 1, i.e. prima aruncare este avers, 2, 3, ... Aceasta este chiar variabila aleatoare geometrică a noastră plus 1. Metodele de calcul cu ea sunt analoage celor folosite mai sus.

2.3.5 Repartiţia uniformă

Repartiţia uniformă modelează orice situaţie unde toate cazurile sunt egal probabile.

$$X \sim \text{uniform}(N).$$

X ia valorile $1, 2, 3, ..., N$, fiecare cu probabilitatea $1/N$. Această repartiţie apare, de exemplu, când modelăm aruncarea unei monede corecte ($N = 2$), aruncarea unui zar corect ($N = 6$), zile de naştere ($N = 365$) şi mâini de poker ($N = C_{52}^5$).

2.3.6 Un site cu graficele unor repartiţii

Adresa `http://mathlets.org/mathlets/probability-distributions/` dă o vedere dinamică a 6 repartiţii. Graficele se schimbă lin pe măsură ce mişcăm diferiţi cursori.
Graficele au un cod de culori ales cu grijă. 2 lucruri cu aceeaşi culoare reprezintă noţiuni identice sau foarte strâns legate.

2.3.7 Alte repartiţii

Sunt multe alte repartiţii cu nume apărute în diferite contexte. De exemplu, `http://en.wikipedia.org/wiki/Hypergeometric_distribution`.

2.4 Aritmetica variabilelor aleatoare

Putem face operaţii aritmetice cu variabile aleatoare. De exemplu, le putem aduna, scădea, înmulţi sau ridica la pătrat.
Există o idee simplă, dar extrem de importantă pentru numărare. Spune că dacă avem o secvenţă de numere care sunt ori 0 ori 1, atunci suma secvenţei este numărul de 1-uri.
Exemplul 11. Considerăm secvenţa cu 5 1-uri

$$1, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0.$$

Este uşor de văzut că suma acestei secvenţe este 5, numărul de 1-uri.

Ilustrăm această idee calculând numărul de aversuri în n aruncări ale unei monede.
Exemplul 12. Aruncăm o monedă corectă de n ori. Fie X_j 1, dacă a j-a aruncare este avers şi 0, dacă este revers. Deci X_j este o variabilă aleatoare Bernoulli(1/2). Fie X numărul total de aversuri din n aruncări. Presupunând că aruncările sunt independente, ştim că $X \sim$ binomial$(n, 1/2)$. Putem scrie de asemenea

$$X = X_1 + X_2 + X_3 + ... + X_n.$$

Din nou, aceasta este deoarece termenii din suma din dreapta sunt toţi ori 0 ori 1. Deci, suma este exact numărul de X_j-uri care sunt 1, i.e. numărul de aversuri.
Lucrul important de văzut în exemplul de mai sus este că am scris variabila aleatoare mai complicată X ca suma variabilelor aleatoare extrem de simple X_j. Aceasta ne permite să manipulăm algebric pe X.
Gândiţi: Presupunem că X şi Y sunt independente, $X \sim$ binomial$(n, 1/2)$ şi $Y \sim$ binomial$(m, 1/2)$. Ce fel de repartiţie are $X + Y$?
(**Răspuns:** binomial$(n + m, 1/2)$. De ce?)
Exemplul 13. Presupunem că X şi Y sunt variabile aleatoare independente.

$$X \sim \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1/10 & 2/10 & 3/10 & 4/10 \end{pmatrix}$$

$$Y \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1/15 & 2/15 & 3/15 & 4/15 & 5/15 \end{pmatrix}$$

Verificaţi că probabilitatea totală pentru fiecare variabilă aleatoare este 1. Ce repartiţie are variabila aleatoare $X+Y$?

Răspuns.

$$(1/10) + (2/10) + (3/10) + (4/10) = 1.$$

Analog pentru Y.

Facem un tabel 2-dimensional pentru spaţiul probelor produs constând din perechile (x, y), unde x este o valoare posibilă a lui X şi y una a lui Y. Pentru a ajuta calculul, probabilităţile pentru valorile lui X sunt puse în coloana din extrema dreaptă şi cele pentru Y sunt pe linia de jos. Deoarece X şi Y sunt independente, probabilitatea pentru perechea (x, y) este chiar produsul probabilităţilor individuale. ("X values" = "valorile lui X", "Y values" = "valorile lui Y).

		Y values				
		1	2	3	4	5	
X values	1	1/150	2/150	3/150	4/150	5/150	1/10
	2	2/150	4/150	6/150	8/150	10/150	2/10
	3	3/150	6/150	9/150	12/150	15/150	3/10
	4	4/150	8/150	12/150	16/150	20/150	4/10
		1/15	2/15	3/15	4/15	5/15	

Benzile diagonale arată mulţimile pătratelor unde $X+Y$ este aceeaşi. Pentru a calcula repartiţia pentru $X+Y$ trebuie să adunăm probabilităţile pentru fiecare bandă.

$$X{+}Y \sim \begin{pmatrix} 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1/150 & 4/150 & 10/150 & 20/150 & 30/150 & 34/150 & 31/150 & 20/150 \end{pmatrix}$$

Simplificând fracţiile,

$$X + Y \sim \begin{pmatrix} 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1/150 & 2/75 & 1/15 & 2/15 & 1/5 & 17/75 & 31/150 & 2/15 \end{pmatrix}$$

Când tabelele sunt prea mari pentru a putea fi scrise avem nevoie să folosim tehnici pur algebrice pentru a calcula probabilităţile sumei.

3 Variabile aleatoare discrete: Media

3.1 Media

Exemplul 1. Presupunem că avem un zar corect cu 6 feţe marcat cu 5 3-uri şi un 6. Cât ne-am aştepta să fie media a 6000 de aruncări?
Răspuns. Dacă am fi ştiut valoarea fiecărei aruncări, am fi calculat media adunând cele 6000 de valori şi împărţind la 6000. Fără să ştim valorile, putem calcula media după cum urmează.
Deoarece sunt 5 3-uri şi un 6 ne aşteptăm ca aproximativ 5/6 dintre aruncări să fie 3 şi 1/6 dintre ele să fie 6. Presupunând acest fapt exact adevărat, avem următorul tabel de valori şi numărări:

valoarea:	3	6
numărări aşteptate:	5000	1000

Atunci media celor 6000 de valori este

$$\frac{5000 \cdot 3 + 1000 \cdot 6}{6000} = \frac{5}{6} \cdot 3 + \frac{1}{6} \cdot 6 = 3.5.$$

Considerăm aceasta media aşteptată în sensul că "ne aşteptăm" ca fiecare dintre valorile posibile să apară cu frecvenţele date.
Exemplul 2. Aruncăm 2 zaruri corecte. Câştigăm 1000 de dolari dacă suma este 2 şi pierdem 100 de dolari altfel. Cât ne aşteptăm să câştigăm în medie pe aruncare?
Răspuns. Probabilitatea unui 2 este 1/36. Dacă jucăm de N ori, ne putem "aştepta" ca $\frac{1}{36} \cdot N$ dintre aruncări să dea 2 şi $\frac{35}{36} \cdot N$ dintre aruncări să dea altceva. Astfel, câştigurile aşteptate totale sunt

$$1000 \cdot \frac{N}{36} - 100 \cdot \frac{35N}{36}.$$

Pentru a obţine media aşteptată pe aruncare, împărţim totalul la N:

$$\text{media aşteptată} = 1000 \cdot \frac{1}{36} - 100 \cdot \frac{35}{36} = -69.(4).$$

Gândiţi: Aţi vrea să jucaţi acest joc o dată? De mai multe ori?
Observăm că în ambele exemple suma pentru media aşteptată constă din termeni care sunt o valoare a unei variabile aleatoare ori probabilitatea ei. Aceata conduce la următoarea definiţie.
Definiţie. Presupunem că X este o variabilă aleatoare discretă care ia valorile $x_1, x_2, ..., x_n$ cu probabilităţile $p(x_1), p(x_2), ..., p(x_n)$. Media lui X este notată $E(X)$ şi definită prin

$$E(X) = \sum_{j=1}^{n} x_j p(x_j) = x_1 p(x_1) + x_2 p(x_2) + ... + x_n p(x_n).$$

Observaţii:
1. Media se mai notează adesea cu μ ("miu") sau cu m.
2. După cum am văzut în exemplele de mai sus, media nu trebuie să fie o valoare posibilă a variabilei aleatoare. Mai degrabă este o medie ponderată a valorilor posibile.
3. Media dă o măsură a tendinţei centrale a unei variabile aleatoare.
4. Dacă toate valorile sunt egal probabile, media este chiar media aritmetică a valorilor.
5. Fie X cu o mulţime numărabilă de valori $x_1, x_2, ...$ luate cu probabilităţile $p(x_1), p(x_2), ...$. Atunci, media lui X este

$$E(X) = \sum_{j=1}^{\infty} x_j p(x_j),$$

dacă seria este convergentă.
Exemplul 3. Aflaţi $E(X)$ pentru

$$X \sim \begin{pmatrix} 1 & 3 & 5 \\ 1/6 & 1/6 & 2/3 \end{pmatrix}.$$

Răspuns. $E(X) = 1 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 5 \cdot \frac{2}{3} = \frac{24}{6} = 4.$
Exemplul 4. Fie $X \sim \text{Bernoulli}(p)$. Aflaţi $E(X)$.
Răspuns. X ia valorile 0 şi 1 cu probabilităţile $1 - p$ şi p, deci

$$E(X) = 0 \cdot (1 - p) + 1 \cdot p = p.$$

Important: Media unei variabile aleatoare Bernoulli(p) este p.
Gândiţi: Care este media sumei a 2 zaruri corecte?

3.1.1 Media şi centrul de masă

Poate v-aţi întrebat de ce folosim numele "funcţie masă de probabilitate". Iată motivul: dacă plasăm câte un obiect de masă $p(x_j)$ la poziţia $x_j, \forall j$, atunci $E(X)$ este poziţia centrului de masă.

Exemplul 5. Fie 2 mase de-a lungul axei Ox, masa $m_1 = 500$ la poziţia $x_1 = 3$ şi masa $m_2 = 100$ la poziţia $x_2 = 6$. Unde este centrul de masă?

Răspuns: Intuitiv, centrul de masă este mai aproape de masa mai mare.

Centrul de masă este

$$\overline{x} = \frac{m_1 x_1 + m_2 x_2}{m_1 + m_2} = \frac{500 \cdot 3 + 100 \cdot 6}{600} = 3.5.$$

Numim această formulă o medie "ponderată" a lui x_1 şi x_2. Aici x_1 este ponderat mai greu deoarece are mai multă masă.

Media $E(X)$ este o medie ponderată a valorilor lui X cu ponderile fiind probabilităţile $p(x_j)$ în locul maselor. Putem spune că "media este punctul în care repartiţia s-ar echilibra". Observaţi similaritatea dintre exemplul din fizică şi exemplul 1.

3.1.2 Proprietăţi algebrice ale lui $E(X)$

Când adunăm, scalăm sau translatăm variabilele aleatoare, mediile fac acelaşi lucru. Modul matematic prescurtat de a spune aceasta este că $E(X)$ este liniară.

1. Dacă X şi Y sunt variabile aleatoare pe un spaţiu al probelor Ω, atunci

$$E(X+Y) = E(X) + E(Y).$$

2. Dacă a şi b sunt constante, atunci

$$E(aX+b) = aE(X) + b.$$

$aX + b$ se obţine prin scalarea lui X cu a şi apoi translatarea cu b.

Exemplul 6. Aruncăm 2 zaruri corecte şi fie X suma lor. Aflaţi $E(X)$.

Răspuns: Fie X_1 numărul de pe primul zar şi X_2 numărul de pe cel de-al 2-lea zar. Deoarece $X = X_1 + X_2$, avem $E(X) = E(X_1) + E(X_2)$. Avem $E(X_1) = E(X_2) = \frac{1+2+3+4+5+6}{6} = \frac{21}{6} = 3.5$, de aceea $E(X) = 7$.

Exemplul 7. Fie $X \sim \text{binomial}(n, p)$. Aflaţi $E(X)$.

Răspuns: Reamintim că X modelează numărul de succese din n variabile aleatoare Bernoulli(p), pe care le notăm $X_1, ..., X_n$. Faptul cheie este că

$$X = \sum_{j=1}^{n} X_j.$$

Acum putem utiliza proprietatea algebrică (1) pentru a face calculele simple.

$$X = \sum_{j=1}^{n} X_j \Longrightarrow E(X) = \sum_{j=1}^{n} E(X_j) = \sum_{j=1}^{n} p = np.$$

Puteam calcula $E(X)$ direct

$$E(X) = \sum_{k=0}^{n} kp(k) = \sum_{k=0}^{n} kC_n^k p^k (1-p)^{n-k}.$$

Este posibil să arătăm că ultima sumă este într-adevăr np, dar metoda folosind proprietatea (1) este mult mai uşoară.

Exemplul 8. (Pentru variabile aleatoare discrete cu un număr infinit de valori media nu există totdeauna.) Fie X cu o mulţime numărabilă de valori,

$$X \sim \begin{pmatrix} 2 & 2^2 & 2^3 & ... & 2^k & ... \\ 1/2 & 1/2^2 & 1/2^3 & ... & 1/2^k & ... \end{pmatrix}.$$

Încercaţi să calculaţi media.

Răspuns: Media ar fi

$$E(X) = \sum_{k=1}^{\infty} 2^k \frac{1}{2^k} = \sum_{k=1}^{\infty} 1 = \infty.$$

Media nu există, deoarece seria nu este convergentă.

3.1.3 Demonstraţiile proprietăţilor algebrice ale lui $E(X)$

Demonstraţia proprietăţii (1) este simplă, dar are o subtilitate în chiar înţelegerea a ceea ce înseamnă adunarea a 2 variabile aleatoare. Reamintim că o valoare a unei variabile aleatoare este un număr determinat de rezultatul unui experiment. A aduna X şi Y înseamnă a aduna valorile lui X şi Y pentru acelaşi rezultat. În formă de tabel, în cazul finit:

rezultatul ω :	ω_1	ω_2	ω_3	...	ω_n
valoarea lui X :	x_1	x_2	x_3	...	x_n
valoarea lui Y :	y_1	y_2	y_3	...	y_n
valoarea lui $X+Y$:	x_1+y_1	x_2+y_2	x_3+y_3	...	x_n+y_n
probabilitatea $P(\omega)$:	$P(\omega_1)$	$P(\omega_2)$	$P(\omega_3)$	...	$P(\omega_n)$

Demonstraţia lui (1):

$$E(X+Y) = \sum_i (x_i + y_i)P(\omega_i) = \sum_i x_i P(\omega_i) + \sum_i y_i P(\omega_i) = E(X) + E(Y).$$

Demonstraţia proprietăţii (2):

$$E(aX+b) = \sum_i (ax_i + b)p(x_i) = a\sum_i x_i p(x_i) + b\sum_i p(x_i) = aE(X) + b.$$

Termenul b din ultima expresie rezultă deoarece $\sum\limits_i p(x_i) = 1$.

Exemplul 9. Media unei repartiţii geometrice
Fie $X \sim \text{geo}(p)$. Reamintim că aceasta înseamnă că X ia valorile $k \in \mathbb{N}$ cu probabilităţile $p(k) = p(1-p)^k$. (X modelează numărul de reversuri dinaintea primului avers într-o secvenţă de probe Bernoulli.) Media este

$$E(X) = \frac{1-p}{p}.$$

Demonstraţia necesită un truc.

$$E(X) = \sum_{k=0}^{\infty} kp(1-p)^k.$$

Suma seriei geometrice: $\sum\limits_{k=0}^{\infty} x^k = \frac{1}{1-x}$.
Derivăm ambii membri: $\sum\limits_{k=0}^{\infty} kx^{k-1} = \frac{1}{(1-x)^2}$.
Înmulţim cu x: $\sum\limits_{k=0}^{\infty} kx^k = \frac{x}{(1-x)^2}$.
Înlocuim x cu $1-p$: $\sum\limits_{k=0}^{\infty} k(1-p)^k = \frac{1-p}{p^2}$.
Înmulţim cu p: $\sum\limits_{k=0}^{\infty} kp(1-p)^k = \frac{1-p}{p}$.
Ultima expresie este media.

$$E(X) = \frac{1-p}{p}.$$

Exemplul 10. Aruncăm o monedă corectă până obţinem un avers pentru prima dată. Care este numărul aşteptat de reversuri?
Răspuns: Numărul de reversuri dinaintea primului avers este modelat de $X \sim \text{geo}(1/2)$. Din exemplul precedent, $E(X) = \frac{1/2}{1/2} = 1$.

Exemplul 11. Michael Jordan, baschetbalist, a marcat 80% dintre aruncările lui libere. Într-un joc, care este numărul aşteptat de aruncări libere pe care le-ar marca înaintea primei ratări?
Răspuns: Acesta este un exemplu unde vrem numărul de succese dinaintea primului eşec. Folosind limbajul de aversuri şi reversuri: succesul este revers (probabilitate $= 1 - p$) şi eşecul este avers (probabilitate $= p$). De aceea $p = 0.2$ şi numărul de reversuri (aruncări libere marcate) dinaintea primului avers (aruncare liberă ratată) este modelat de $X \sim \text{geo}(0.2)$. Din exemplul 9,

$$E(X) = \frac{1-p}{p} = \frac{0.8}{0.2} = 4.$$

3.1.4 Medii ale funcţiilor de o variabilă aleatoare

(Formula schimbării de variabile)
Dacă X este o variabilă aleatoare luând valorile $x_1, x_2, ...$ şi h este o funcţie, atunci $h(X)$ este o nouă variabilă aleatoare. Media ei este

$$E(h(X)) = \sum_j h(x_j)p(x_j).$$

Exemplul 12. Fie X valoarea aruncării unui zar corect şi $Y = X^2$. Aflaţi $E(Y)$.
Răspuns: Deoarece există un număr mic de valori, putem face un tabel.

X	1	2	3	4	5	6
Y	1	4	9	16	25	36
probabilitatea	1/6	1/6	1/6	1/6	1/6	1/6

Observăm că probabilitatea pentru fiecare valoare a lui Y este aceeaşi ca cea pentru valoarea corespunzătoare a lui X. Deci,

$$E(Y) = E(X^2) = 1^2 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + ... + 6^2 \cdot \frac{1}{6} = 15.1(6).$$

Exemplul 13. Aruncăm 2 zaruri corecte şi fie X suma. Presupunem că funcţia de plată este dată de $Y = X^2 - 6X + 1$. Este acesta un pariu bun din punct de vedere financiar?
Răspuns: Avem $E(Y) = \sum_{j=2}^{12} (j^2 - 6j + 1)p(j)$, unde $p(j) = P(X = j)$.
Tabelul:

X	2	3	4	5	6	7	8	9	10	11	12
Y	-7	-8	-7	-4	1	8	17	28	41	56	73
probabilitatea	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

Este mai uşor să calculăm în R. Iată codul:

```
x=2:12
y=x^ 2-6*x+1
p=c(1,2,3,4,5,6,5,4,3,2,1)/36
sum(p*y)
```

Obţinem 13.83333, deci $E(Y) = 13.8(3)$.
Deoarece plata aşteptată este pozitivă, pariul este bun din punct de vedere financiar.
Întrebare: Dacă $Y = h(X)$, rezultă $E(Y) = h(E(X))$?
Răspuns: **NU!!!** Nu este adevărat în general!
Gândiţi: Este adevărat în exemplul precedent?
Întrebare: Dacă $Y = 3X + 77$, rezultă $E(Y) = 3E(X) + 77$?
Răspuns: Da. Din proprietatea (2), scalarea şi translatarea se comportă astfel.