

Curs 2

Cristian Niculescu

1 Probabilitatea condiționată, independența și teorema lui Bayes

1.1 Scopurile învățării

1. Să știe definițiile probabilității condiționate și independenței evenimentelor.
2. Să poată calcula probabilitatea condiționată direct din definiție.
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” probabilitățile condiționate.
6. Să poată organiza calculul probabilităților condiționate folosind arbori și tabele.
7. Să înțeleagă complet eroarea ratei de bază.

1.2 Probabilitatea condiționată

Probabilitatea condiționată răspunde la întrebarea ”cum se schimbă probabilitatea unui eveniment dacă avem informație suplimentară”.

Exemplul 1. Aruncăm o monedă corectă de 3 ori.

(a) Care este probabilitatea a 3 aversuri?

Răspuns: Spațiul 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 informație, cum ar trebui să calculăm probabilitatea a 3 aversuri?

Răspuns: Avem un nou spațiu al probelor (reduc): $\Omega' = \{HHH, HHT, HTH, HTT\}$.

Toate cazurile sunt egal probabile, deci

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

Aceasta este numită **probabilitate condiționată**, deoarece ia în considerare condiții suplimentare. Pentru a dezvolta notația, reformulăm (b) în termeni

de *evenimente*.

(b) reformulat. Fie A evenimentul "toate 3 aruncările sunt aversuri" = $\{HHH\}$.

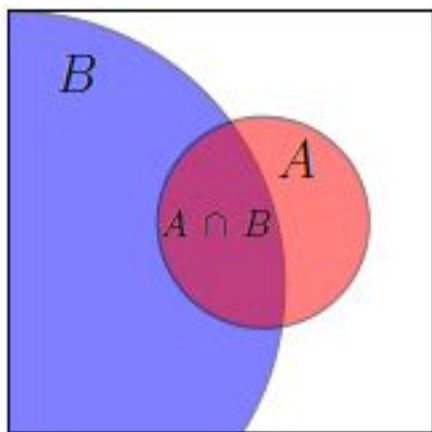
Fie B evenimentul "prima aruncare este avers" = $\{HHH, HHT, HTH, HTT\}$.

Probabilitatea condiționată a lui A cunoscând că B a avut loc este scrisă

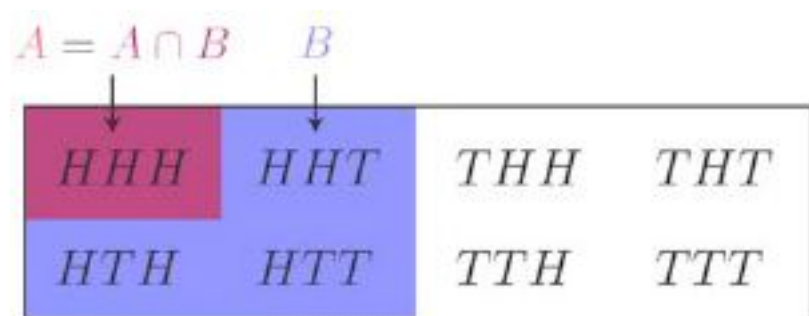
$$P(A|B).$$

Aceasta se citește "probabilitatea condiționată a lui A dat fiind B " sau "probabilitatea lui A condiționată de B " sau, mai simplu, "probabilitatea lui A dat fiind B ".

Putem vizualiza probabilitatea condiționată după cum urmează. Gândiți $P(A)$ ca proporția ariei ocupată de A din *tot* spațiul probelor. Pentru $P(A|B)$ ne restrângem atenția la B . Adică, $P(A|B)$ este proporția ocupată de A din aria lui B , i.e. $P(A \cap B)/P(B)$.



Probabilitatea condiționată: vizualizare abstractă



Probabilitatea condiționată: exemplul cu moneda

Observăm că $A \subset B$ în ultima figură, deci apar doar 2 culori.

Definiția formală a probabilității condiționate prinde esența exemplului și

vizualizărilor de mai sus.

Definiția formală a probabilității condiționate

Fie A și B evenimente. Definim **probabilitatea condiționată** a lui A dat fiind B ca

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

În continuare, când scriem o probabilitate condiționată, presupunem implicit că probabilitatea evenimentului care condiționează este nenulă, fără a mai scrie aceasta.

Să refacem exemplul cu aruncarea monezii folosind definiția din relația (1). Reamintim că $A =$ "3 aversuri" și $B =$ "prima aruncare este avers". Avem $P(A) = 1/8$ și $P(B) = 1/2$. 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**.

$$P(A \cap B) = P(A|B) \cdot P(B). \quad (2)$$

Aceasta este doar o rescriere a definiției din relația (1) a probabilității condiționate. Regula multiplicării este doar o versiune a regulii produsului.

Începem cu un exemplu simplu unde putem verifica direct toate probabilitățile numărând.

Exemplul 2. Tragem 2 cărți fără revenire dintr-un pachet de cărți de joc fără jokeri. Definim evenimentele: $S_1 =$ "prima carte este o pică" și $S_2 =$ "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 cărți rămase, 12 sunt pici.

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

Acum, să recalculăm asta folosind formula (1). Trebuie să calculăm $P(S_1)$, $P(S_2)$ și $P(S_1 \cap S_2)$. Știm că $P(S_1) = 1/4$ deoarece sunt 52 de moduri egal posibile de a trage prima carte și 13 din ele sunt pici. Aceeași 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 cărți afectează cu siguranță probabilitățile pentru a 2-a carte. Dar, dacă ne uităm la *toate* secvențele posibile de 2 cărți, vom vedea că fiecare carte din pachet are probabilitate egală de a fi a 2-a carte. Deoarece 13 din

cele 52 de cărți sunt pici obținem $P(S_2) = 13/52 = 1/4$. Un alt fel de a spune asta este: dacă nu avem dată valoarea pentru prima carte, atunci trebuie să considerăm toate posibilitățile 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 și împărțind prin numărul de moduri de a trage orice carte urmată de orice altă carte.) Acum, folosind (1) obținem

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

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

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

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

1.4 Legea probabilității totale

Legea probabilității totale ne permite să folosim regula multiplicării pentru a afla probabilități în exemple mai interesante. Necesită multă notație, dar ideea este destul de simplă. Enunțăm legea când spațiul probelor este împărțit în 3 părți. Este o chestiune simplă a extinde regula când sunt mai mult de 3 părți.

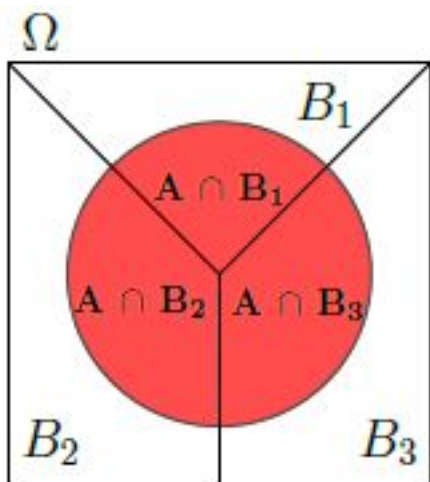
Legea probabilității totale

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

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

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

Prima relație spune "dacă A este împărțit în 3 părți disjuncte 2 câte 2, atunci $P(A)$ este suma probabilităților părților". Relația (3) se numește **legea probabilității totale**. Ea este doar o rescriere a relației de deasupra folosind regula multiplicării.



Spațiul probelor Ω și evenimentul A sunt fiecare împărțite în 3 părți disjuncte 2 câte 2.

Legea este valabilă dacă împărțim Ω în orice număr de evenimente, atât timp cât ele sunt *disjuncte 2 câte 2* și *acoperă* tot Ω (adică reuniunea lor este Ω). O astfel de împărțire este numită adesea o *partiție* a lui Ω .

Primul nostru exemplu va fi unul unde știm deja răspunsul și putem verifica legea.

Exemplul 3. O urnă conține 5 bile roșii și 2 bile verzi. 2 bile sunt scoase una după alta, fără revenire. Care este probabilitatea ca a 2-a bilă să fie roșie?

Răspuns: Spațiul probelor este $\Omega = \{rr, rg, gr, gg\}$. (Am pus "r" pentru roșu și "g" pentru verde.)

Fie R_1 evenimentul "prima bilă este roșie", $G_1 =$ "prima bilă este verde", $R_2 =$ "a 2-a bilă este roșie", $G_2 =$ "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 cărți de joc de mai sus. Fiecare bilă este egal probabilă să fie a 2-a bilă. Deoarece 5 din 7 bile sunt roșii, $P(R_2) = 5/7$.

Să calculăm această valoare folosind legea probabilității totale (3). Întâi, aflăm probabilitățile condiționate. Acesta este un exercițiu simplu de numărare.

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

Deoarece R_1 și G_1 partiționează Ω , legea probabilității totale spune

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

Urne de probabilitate

Exemplul precedent a folosit urne de probabilitate.

În probabilități și statistică, o problemă a urnei este un exercițiu mental idealizat în care unele obiecte de real interes (ca atomi, oameni, mașini, 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 proprietăți. 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 puțin mai complicate.

Exemplul 4. O urnă conține 5 bile roșii și 2 bile verzi. O bilă este extrasă. Dacă este verde, o bilă roșie este adăugată în urnă, iar dacă este roșie, 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 roșie?

Răspuns. Cu notațiile de la exemplul 3, legea probabilității totale spune că $P(R_2)$ poate fi calculată folosind relația (4). Doar valorile probabilităților condiționate se schimbă. Avem

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

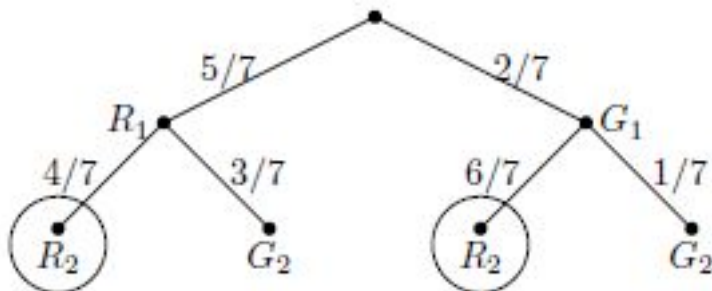
De aceea,

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

1.5 Utilizarea arborilor pentru a organiza calculul

Arborii sunt o metodă de a organiza calcule cu probabilități condiționate și legea probabilității totale. Ca la regula produsului, organizăm procesul de bază într-o secvență de acțiuni.

Începem refăcând exemplul 4. Secvența de acțiuni este: întâi extragem bila 1 (și adăugăm bila corespunzătoare în urnă) și 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 și așa 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.

Probabilitățile sunt scrise de-a lungul ramurilor. Probabilitatea lui R_1 (bilă roșie 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 probabilitățile **condiționate**. 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 probabilităților 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 înmulțirea probabilităților aflate de-a lungul drumului de la nodul rădăcină la nodul R_2 din extrema stângă.

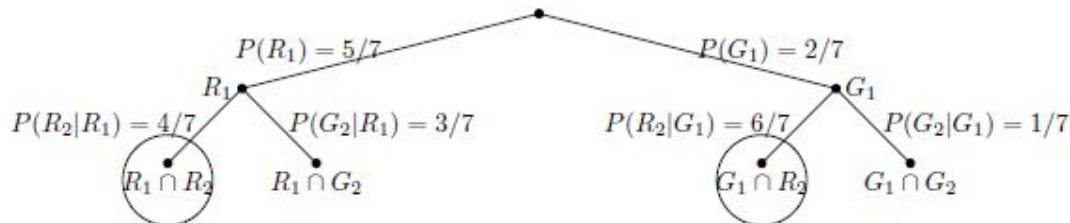
Legea probabilității totale este chiar afirmația că $P(R_2)$ este suma probabilităților tuturor drumurilor conducând de la nodul rădăcină la R_2 (cele 2 noduri încercuite din figură). În acest caz,

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

exact ca în exemplul precedent.

1.5.1 Arbori prescurtați vs. preciși

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 același arbore etichetat precis.



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

1.6 Independență

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

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

Independența experimentelor poate fi subminată de eșecul curățării sau recalibrării echipamentului între experimente, sau al izolării observatorilor presupuși independenți unul față de altul sau de o influență comună. Toți am experimentat aflarea aceluiași ”fapt” de la oameni diferiți. Aflându-l din surse diferite tindem să-i dăm crezare până aflăm că toți l-au aflat de la o sursă comună. Adică, sursele noastre nu erau independente.

Traducerea descrierii verbale în simboluri dă

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

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

Dacă A și 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 definiție tehnică a independenței.

Definiția formală a independenței: 2 evenimente A și B sunt **independente** \iff

$$P(A \cap B) = P(A) \cdot P(B). \quad (6)$$

Această definiție simetrică clarifică faptul că A este independent de $B \iff B$ este independent de A . Spre deosebire de relația (5), această definiție are sens și când $P(B) = 0$. În termeni de probabilități condiționate, avem:

1. Dacă $P(B) \neq 0$, atunci A și B sunt independente $\iff P(A|B) = P(A)$.
 2. Dacă $P(A) \neq 0$, atunci A și 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” și $H_2 =$ ”avers la a 2-a aruncare”. Sunt H_1 și 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 și H_2 sunt independente.

Putem întreba despre independența oricăror 2 evenimente, ca în următoarele 2 exemple.

Exemplul 7. Aruncăm o monedă corectă de 3 ori. Fie $H_1 =$ ”avers la prima aruncare” și $A =$ ”2 aversuri în total”. Sunt H_1 și A independente?

Răspuns. Știm că $P(A) = 3/8$. Deoarece $P(H_1) = 1/2 \neq 0$, putem verifica dacă formula din relația (5) este îndeplinită. $H_1 = \{HHH, HHT, HTH, HTT\}$ conține 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 și A nu sunt independente.

Exemplul 8. Tragem o carte dintr-un pachet de cărți de joc standard (adică fără jokeri). Să examinăm independența 2 câte 2 a 3 evenimente: ”cartea este as”, ”cartea este cupă” și ”cartea este roșie”.

Notăm evenimentele $A =$ ”as”, $H =$ ”cupă”, $R =$ ”roșie”.

a) Știm 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 sfârșit, ce putem spune despre H și R ? Deoarece $P(H) = 13/52 = 1/4$ și $P(H|R) = 13/26 = 1/2$, avem $P(H) \neq P(H|R)$, deci H și R nu sunt independente. Puteam de asemenea vedea asta invers: $P(R) = 26/52 = 1/2$ și $P(R|H) = 13/13 = 1$, deci $P(R) \neq P(R|H)$, de aceea H și R nu sunt independente.

1.6.1 Paradoxurile independenței

Un eveniment A cu probabilitatea 0 este independent de el însuși, deoarece în acest caz ambii membri ai relației (6) sunt 0. Aceasta apare paradoxal, deoarece cunoașterea faptului că A a avut loc sigur dă informație despre faptul dacă A a avut loc. Rezolvăm paradoxul observând că deoarece $P(A) = 0$, afirmația ” A a avut loc” este fără sens.

Gândiți: Pentru ce altă valoare a lui $P(A)$ este A independent de el însuși?

1.7 Teorema lui Bayes

Teorema lui Bayes este foarte importantă pentru probabilități și statistică. Pentru 2 evenimente A și B , [teorema lui Bayes](#) (de asemenea numită [regula](#)

lui Bayes și formula lui Bayes) spune

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

Comentarii: 1. Regula lui Bayes ne spune cum să ”inversăm” probabilitățile condiționate, i.e. să aflăm $P(B|A)$ din $P(A|B)$.

2. În practică, $P(A)$ este adesea calculată folosind legea probabilității totale.

Demonstrația regulii lui Bayes

Punctul cheie este că $A \cap B$ este simetrică în A și B . Deci regula multiplicării spune

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

Acum împărțim cu $P(A)$ pentru a obține regula lui Bayes.

O greșeală obișnuită este a confunda $P(A|B)$ cu $P(B|A)$. Ele pot fi foarte diferite. Acest fapt este ilustrat în următorul exemplu.

Exemplul 9. Aruncăm o monedă corectă de 5 ori. Fie H_1 = ”prima aruncare este avers” și 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 ușor de confundat semnificația lui $P(B|A)$ cu $P(A|B)$ când o situație este descrisă în cuvinte.

Exemplul 10. Eroarea ratei de bază

Considerăm un test de rutină pentru o boală. Presupunem că frecvența bolii în populație (rata de bază) este 0.5%. Testul are o rată ”fals pozitiv” de 5% și o rată ”fals negativ” de 10%.

Faceți testul și ieșiți pozitiv. Care este probabilitatea să aveți boala?

Răspuns. Vom face calculul de 3 ori: folosind arbori, tablele și simboluri. Vom utiliza următoarele notații pentru evenimentele relevante:

$D+$ = ”aveți boala”,

$D-$ = ”nu aveți boala”,

$T+$ = ”ați fost testat pozitiv”,

$T- =$ "ați fost testat negativ".

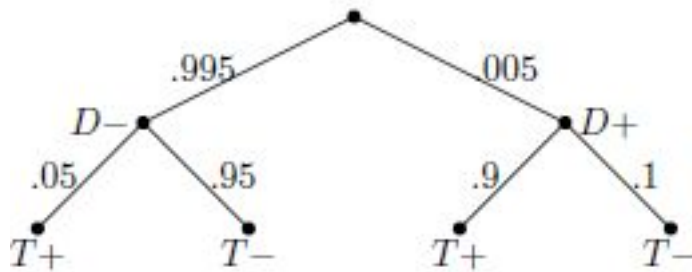
Avem dat $P(D+) = 0.005$ și de aceea $P(D-) = 0.995$. Ratele "fals pozitiv" și "fals negativ" sunt (prin definiție) probabilități condiționate.

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

Probabilitățile complementare sunt cunoscute ca ratele "adevărat negativ" și "adevărat pozitiv":

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

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



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

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

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

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

Astfel,

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

Observații. Aceasta este numită eroarea ratei de bază deoarece rata de bază a bolii în populație este atât de mică încât marea majoritate a oamenilor care fac testul sunt sănătoși și chiar cu un test precis majoritatea celor testați pozitiv vor fi oameni sănătoși.

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 probabilități este a face un tabel. Să refacem exemplul precedent folosind un tabel construit pentru un total de 10000 de oameni împărțiți conform probabilităților din acest exemplu.

Construim tabelul după cum urmează. Alegem un număr, să zicem 10000 de oameni, și 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 sănătoși. În acest moment tabelul arată așa:

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

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

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

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

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

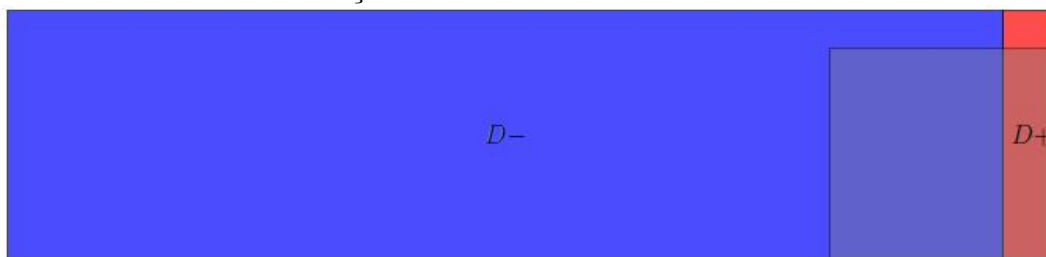
Folosind tabelul complet putem calcula

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

Simboluri: Pentru completitudine, arătăm soluția 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 toți oamenii sănătoși. Aria roșie mult mai mică reprezintă oamenii bolnavi. Dreptunghiul mai închis reprezintă oamenii care au fost testați pozitiv. Aria mai închisă acoperă cea mai mare parte a ariei roșii și doar o mică parte din aria albastră. Chiar și așa, cea mai mare parte a ariei mai închise este peste albastru. Adică, cele mai multe teste pozitive sunt ale oamenilor sănătoși.



2 Variabile aleatoare discrete

2.1 Scopurile învățării

1. Să știe definiția unei variabile aleatoare discrete.
2. Să știe distribuțiile (repartițiile) Bernoulli, binomială și geometrică și exemple de ce modelează ele.
3. Să poată să descrie funcția masă de probabilitate și funcția de distribuție cumulativă folosind tabele și formule.
4. Să poată să construiască noi variabile aleatoare din cele vechi.
5. Să știe cum să calculeze media.

2.2 Variabile aleatoare

Cuvintele cheie sunt

1. Variabilă aleatoare

2. Funcție masă de probabilitate (pmf)
3. Funcție de distribuție cumulativă (cdf), numită și funcție de repartiție.

2.2.1 Recapitulare

Un spațiu al probelor discret Ω este o mulțime finită sau numărabilă de cazuri (rezultate) $\{\omega_1, \omega_2, \dots\}$. Probabilitatea cazului ω este notată $P(\omega)$.

Un eveniment E este o submulțime a lui Ω . Probabilitatea evenimentului E este $P(E) = \sum_{\omega \in E} P(\omega)$.

2.2.2 Variabilele aleatoare ca funcții de plată

Exemplul 1. Un joc cu 2 zaruri corecte.

Aruncăm 2 zaruri corecte și înregistrăm cazurile ca (i, j) , unde i este rezultatul primului zar și j rezultatul celui de-al 2-lea. Putem lua spațiul probelor

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

Funcția de probabilitate este $P(i, j) = 1/36$.

În acest joc, câștigăm 500 de dolari dacă suma este 7 și pierdem 100 altfel. Dăm acestei funcții de plată numele X și 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 funcție de plată diferită. De exemplu

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

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

Aceste funcții de plată sunt exemple de variabile aleatoare. O variabilă aleatoare atribuie un număr fiecărui caz dintr-un spațiu al probelor. Mai formal:

Definiție. Fie Ω un spațiu al probelor. O variabilă aleatoare discretă este o funcție

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

care ia o mulțime discretă de valori. (Reamintim că \mathbb{R} este mulțimea numerelor reale.)

De ce este X numită variabilă aleatoare? Este "aleatoare" deoarece valoarea ei depinde de un caz aleator al unui experiment. Și tratăm pe X ca pe o variabilă uzuală: putem s-o adunăm la alte variabile aleatoare, s-o ridicăm la pătrat, ș.a.m.d.

2.2.3 Evenimente și 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 și 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 mulțimea $\{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}$, i.e. mulțimea tuturor cazurilor cu suma componentelor 7. $P(X = 500) = 6/36 = 1/6$.

Permitem lui a să ia orice valoare, chiar și 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 mulțimea vidă).

" $X = 1000$ " = \emptyset .

$P(X = 1000) = 0$.

2.2.4 Funcția masă de probabilitate

Devine obositor și greu să citim și să scriem $P(X = a)$ pentru probabilitatea lui $X = a$. Când știm că vorbim despre X vom scrie simplu $p(a)$. Dacă vrem să-l facem explicit pe X vom scrie $p_X(a)$.

Definiție. **Funcția masă de probabilitate (pmf)** a unei variabile aleatoare discrete este funcția $p : \mathbb{R} \rightarrow \mathbb{R}, p(a) = P(X = a)$.

Observăm:

1. Avem totdeauna $0 \leq p(a) \leq 1$.
2. Permite 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 spațiul 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 și probabilitățile 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: $p(8) = 0$.

Gândiți: Care este pmf pentru $Z(i, j) = i + j$? Arată familiar?

2.2.5 Evenimente și inegalități

Inegalitățile cu variabile aleatoare descriu evenimente. De exemplu, $X \leq a$ este mulțimea tuturor cazurilor ω a.î. $X(\omega) \leq a$.

Exemplul 5. Dacă spațiul probelor este mulțimea tuturor perechilor (i, j) provenind din aruncarea a 2 zaruri corecte și $Z(i, j) = i + j$ este suma zarurilor, atunci

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

2.2.6 Funcția de distribuție cumulativă (cdf)

Definiție. **Funcția de distribuție cumulativă (cdf)** a unei variabile aleatoare X este funcția $F : \mathbb{R} \rightarrow \mathbb{R}$ dată de $F(a) = P(X \leq a)$. Adesea vom prescurta aceasta **funcția de repartiție**.

Observăm că definiția 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 funcție de distribuție **cumulativă** deoarece $F(a)$ dă totalul probabilității care se acumulează adunând probabilitățile $p(b)$ 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:

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

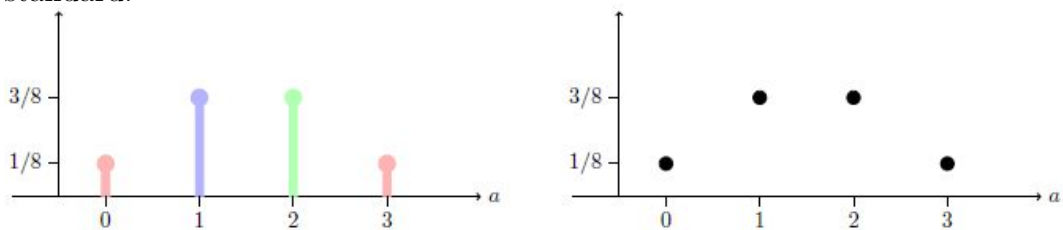
La fel ca funcția masă de probabilitate, $F(a)$ este definită pentru toate valorile reale a . În exemplul de mai sus, $F(8) = 1$, $F(-2) = 0$, $F(2.5) = 1/9$ și $F(\pi) = 1/4$.

2.2.7 Graficele lui p și F

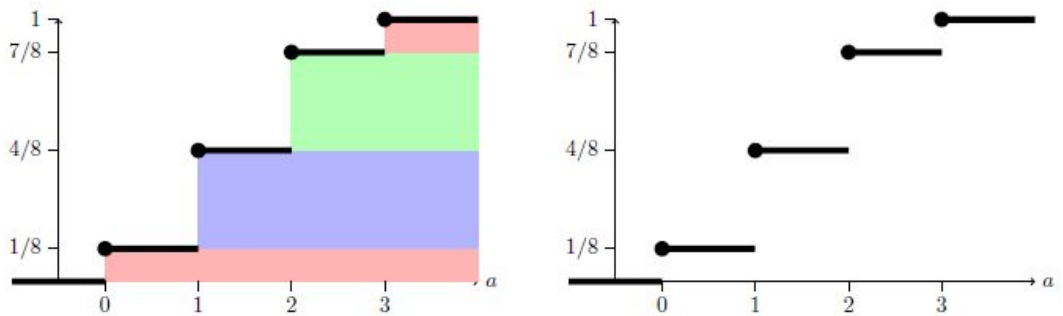
Putem vizualiza pmf și cdf cu grafice. De exemplu, fie X numărul de aversuri din 3 aruncări ale unei monede corecte:

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

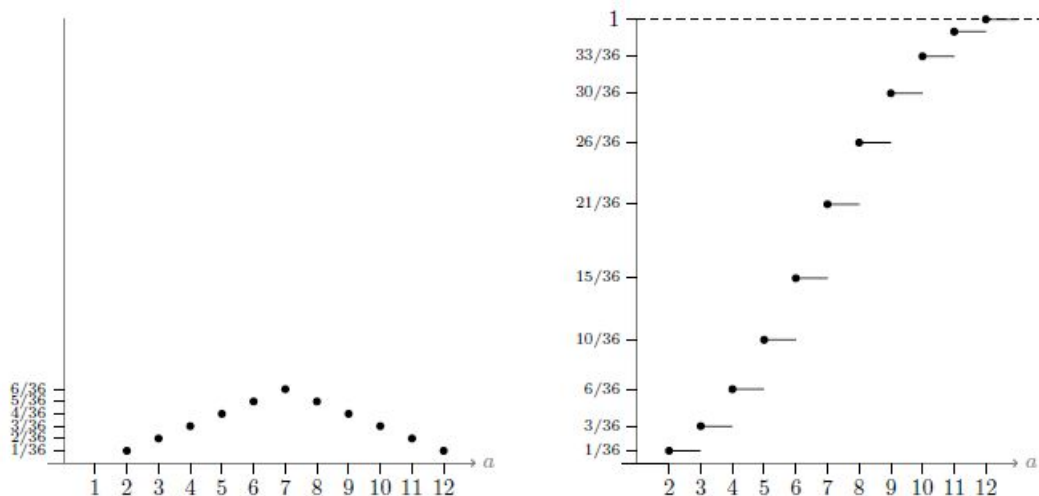
Graficele colorate arată cum funcția de distribuție cumulativă este construită prin **acumularea** probabilității când a crește. Graficele alb-negru sunt prezentări standard.



Funcția masă de probabilitate pentru X



Funcția de distribuție cumulativă a lui X



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

2.2.8 Proprietăți ale cdf F

Cdf F a unei variabile aleatoare satisface câteva proprietăți:

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. $\lim_{a \rightarrow -\infty} F(a) = 0, \lim_{a \rightarrow \infty} F(a) = 1$.

În cuvinte, (1) spune că $F(a)$ crește sau rămâne constantă când a crește, dar niciodată nu descrește; (2) spune că probabilitatea acumulată este totdeauna între 0 și 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 $X \leq a$.

Gândiți: De ce o cdf satisface fiecare dintre aceste proprietăți?

2.3 Repartiții (distribuții) clasice

2.3.1 Repartiții Bernoulli

Modelul: Repartiția Bernoulli modelează o probă într-un experiment care poate avea ca rezultat ori **succes** ori **eșec**. O variabilă aleatoare X are o **repartiție Bernoulli** cu parametrul $p \iff$

1. X ia valorile 0 și 1.
2. $P(X = 1) = p$ și $P(X = 0) = 1 - p$.

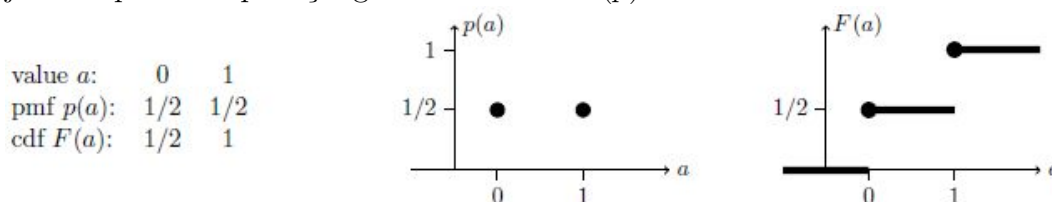
Vom scrie $X \sim \text{Bernoulli}(p)$ sau $\text{Ber}(p)$, ceea ce se citește " **X are o repartiție Bernoulli cu parametrul p** ".

Un model simplu pentru repartiția Bernoulli este aruncarea unei monede cu

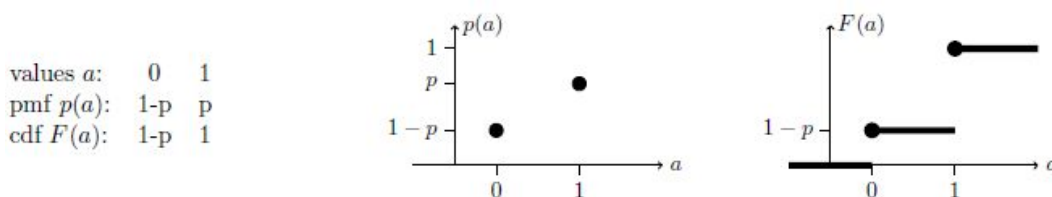
probabilitatea p a aversului, cu $X = 1$ pe avers și $X = 0$ pe revers. Terminologia generală este a spune că X este 1 la [succes](#) și 0 la [eșec](#), cu succesul și eșecul definite de context.

Multe decizii pot fi modelate ca o alegere binară, cum ar fi voturile pentru sau împotriva unei propuneri. Dacă p este proporția din populația care votează care favorizează propunerea, atunci votul unui individ aleator este modelat prin $\text{Bernoulli}(p)$.

Iată tabelul și graficele pentru pmf și cdf ale repartiției $\text{Bernoulli}(1/2)$ și mai jos cele pentru repartiția generală $\text{Bernoulli}(p)$.



Tabelul, pmf și cdf pentru repartiția $\text{Bernoulli}(1/2)$



Tabelul, pmf și cdf pentru repartiția $\text{Bernoulli}(p)$

2.3.2 Repartiții binomiale

[Repartiția binomială](#) $\text{Binomial}(n, p)$, sau $\text{Bin}(n, p)$ modelează numărul de succese din n probe independente $\text{Bernoulli}(p)$.

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, spațiul probelor Bernoulli este $\{H, T\}$. Spațiul probelor binomiale este format din toate [secvențele](#) de aversuri și reversuri de lungime n . O variabilă aleatoare Bernoulli ia valorile 0 și 1. O variabilă aleatoare binomială ia valorile 0, 1, 2, ..., n .

Exemplul 6. $\text{Binomial}(1, p)$ este la fel ca $\text{Bernoulli}(p)$.

Exemplul 7. Numărul de aversuri din n aruncări ale unei monede cu probabilitatea p a aversului are o repartiție $\text{Binomial}(n, p)$.

Descriem $X \sim \text{Binomial}(n, p)$ dând valorile și probabilitățile ei. Considerăm $k \in \{0, 1, \dots, n\}$.

Reamintim că numărul de moduri de a alege k obiecte dintr-o colecție de n

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

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

(Este numit și [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. Coeficienții binomiali asociați cu $n = 5$ sunt

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

și analog

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

Folosind aceste valori obținem următorul tabel pentru $X \sim \text{Binomial}(5, p)$.

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 probabilităților binomiale

Pentru concretețe, fie $n = 5$ și $k = 2$ (argumentul pentru n și k arbitrari este identic). Deci $X \sim \text{Binomial}(5, p)$ și vrem să calculăm $p(2)$. Calea lungă de a calcula $p(2)$ este a lista toate modurile de a obține exact 2 aversuri în 5 aruncări de monedă și a aduna probabilitățile lor. Lista are 10 elemente:

HHTTT, HTHTT, HTTHT, HTTTH, THHTT, THTHT, THTTH, TTHHT, TTHTH, TTTHH.

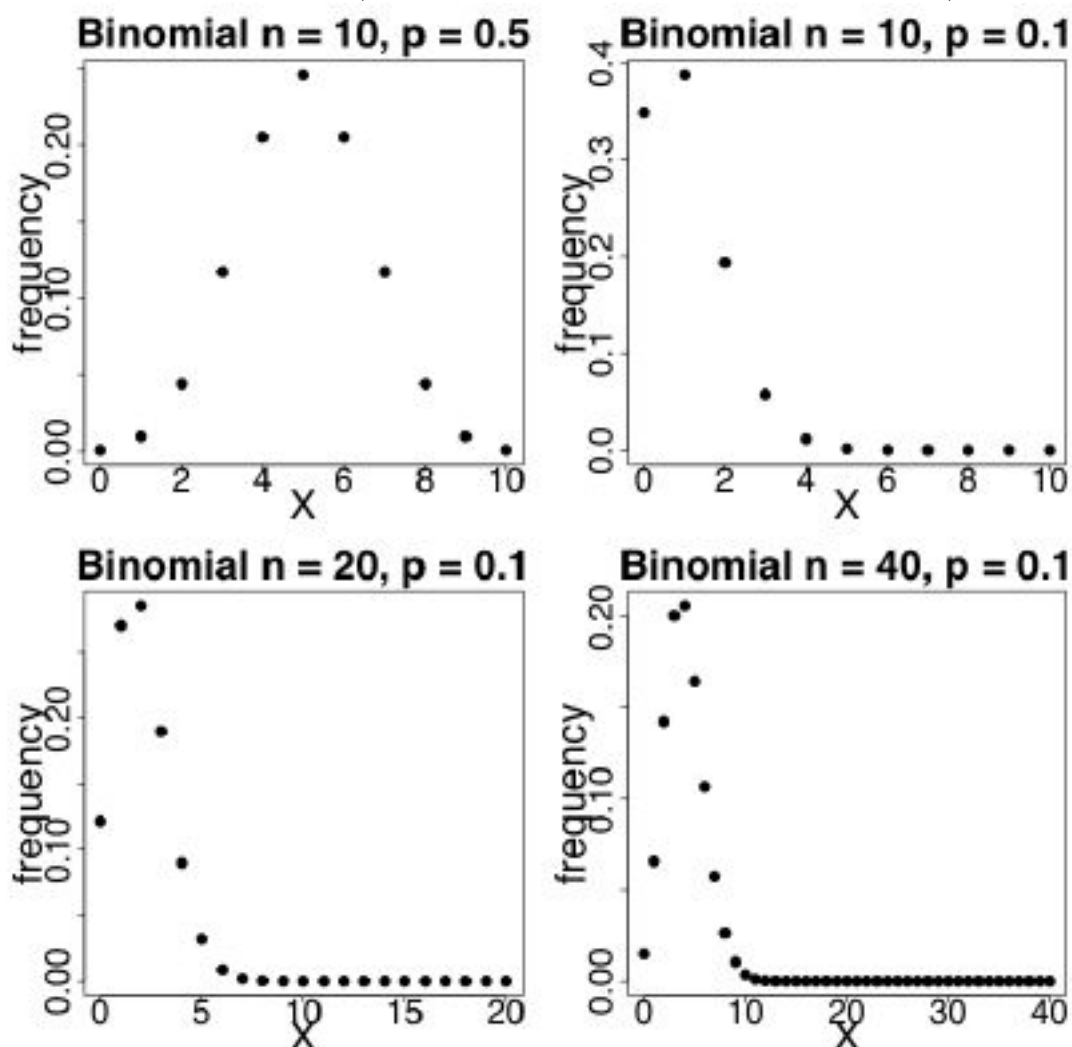
Fiecare element are aceeași probabilitate de a avea loc, și anume

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

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

individuale sunt independente, putem înmulți probabilitățile. De aceea, probabilitatea totală a exact 2 aversuri este suma celor 10 probabilități identice, i.e. $p(2) = 10p^2(1 - p)^3$, așa cum se arată în tabel.

Aceasta ne ghidează la calea scurtă de a face calculul. Trebuie să calculăm numărul de secvențe cu exact 2 aversuri. Pentru a face aceasta avem nevoie să alegem 2 dintre aruncări să fie aversuri și cele 3 rămase vor fi reversuri. Numărul de astfel de secvențe este numărul de moduri de a alege 2 din 5 obiecte, i.e. C_5^2 . Deoarece fiecare astfel de secvență are aceeași probabilitate, $p^2(1 - p)^3$, obținem 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 Repartiții geometrice

O **repartiție geometrică** modelează numărul de reversuri dinaintea primului avers într-o secvență 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 și unui avers valoarea 1. În acest caz, X este numărul de 0-uri dinainte de primul 1.

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

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

(e) Numim reversul eșec și aversul succes. Astfel, X este numărul de eșecuri dinaintea primului succes.

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

Definiția formală. Variabila aleatoare X are o **repartiție geometrică cu parametrul p** \iff

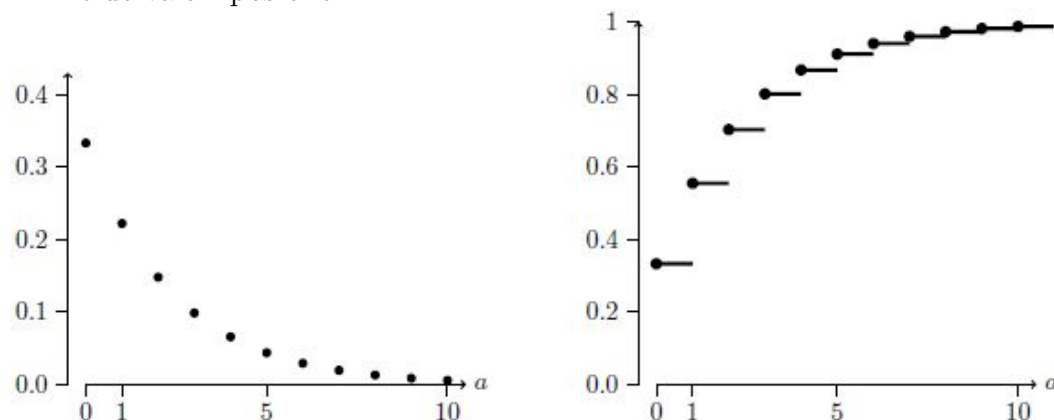
- X ia valorile 0,1,2,3,...
- are pmf $p(k) = P(X = k) = p(1 - p)^k$.

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.

Repartiția geometrică este un exemplu de repartiție discretă care ia un număr infinit de valori posibile.



Pmf și cdf pentru repartiția $\text{geometric}(1/3)$

Exemplul 10. Calculul probabilităților geometrice. Presupunem că locuitorii unei insule își 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 toți copiii supraviețuiesc, că orice familie poate face oricât de mulți copii și că nu sunt gemeni. Care este probabilitatea ca o familie să aibă k băieți?

Răspuns. Numărul de băieți dintr-o familie este numărul de băieți dinaintea primei fete.

Fie X numărul de băieți 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 băieți dacă secvența de copii din familie de la cel mai mare la cel mai mic este

$$BBB...BF,$$

primii k copii fiind băieți. Probabilitatea acestei secvențe este chiar produsul probabilităților pentru fiecare copil, i.e. $(1/2)^k \cdot (1/2) = (1/2)^{k+1}$. (Notă: Cele 4 presupuneri sunt simplificări ale realității.)

Mai multă confuzie geometrică. Altă definiție pentru repartiția 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 Repartiția uniformă

Repartiția uniformă modelează orice situație unde toate cazurile sunt egal probabile.

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

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

2.3.6 Un site cu graficele unor repartiții

Adresa <http://mathlets.org/mathlets/probability-distributions/> dă o vedere dinamică a 6 repartiții. Graficele se schimbă lin pe măsură ce mișcăm diferiți cursori.

Graficele au un cod de culori ales cu grijă. 2 lucruri cu aceeași culoare reprezintă noțiuni identice sau foarte strâns legate.

2.3.7 Alte repartiții

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

2.4 Aritmetica variabilelor aleatoare

Putem face operații aritmetice cu variabile aleatoare. De exemplu, le putem aduna, scădea, înmulți sau ridica la pătrat.

Există o idee simplă, dar **extrem de importantă** pentru numărare. Spune că dacă avem o secvență de numere care sunt ori 0 ori 1, atunci suma secvenței este numărul de 1-uri.

Exemplul 11. Considerăm secvența cu 5 1-uri

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

Este ușor de văzut că suma acestei secvențe 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 și 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, știm că $X \sim \text{binomial}(n, 1/2)$. Putem scrie de asemenea

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

Din nou, aceasta este deoarece termenii din suma din dreapta sunt toți 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 .

Gândiți: Presupunem că X și Y sunt independente, $X \sim \text{binomial}(n, 1/2)$ și $Y \sim \text{binomial}(m, 1/2)$. Ce fel de repartiție are $X + Y$?

(**Răspuns:** $\text{binomial}(n + m, 1/2)$. De ce?)

Exemplul 13. Presupunem că X și Y sunt variabile aleatoare independente.

$$X \sim \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1/10 & 2/10 & 3/10 & 4/10 \end{pmatrix}$$
$$Y \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1/15 & 2/15 & 3/15 & 4/15 & 5/15 \end{pmatrix}$$

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

Răspuns.

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

Analog pentru Y .

Facem un tabel 2-dimensional pentru spațiul probelor produs constând din perechile (x, y) , unde x este o valoare posibilă a lui X și y una a lui Y . Pentru a ajuta calculul, probabilitățile pentru valorile lui X sunt puse în coloana din extrema dreaptă și cele pentru Y sunt pe linia de jos. Deoarece X și Y sunt independente, probabilitatea pentru perechea (x, y) este chiar produsul probabilităților 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 mulțimile pătratelor unde $X + Y$ este aceeași. Pentru a calcula repartiția pentru $X + Y$ trebuie să adunăm probabilitățile pentru

fiecare bandă.

$$X+Y \sim \left(\begin{array}{cccccccc} 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1/150 & 4/150 & 10/150 & 20/150 & 30/150 & 34/150 & 31/150 & 20/150 \end{array} \right)$$

Simplificând fracțiile,

$$X + Y \sim \left(\begin{array}{cccccccc} 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1/150 & 2/75 & 1/15 & 2/15 & 1/5 & 17/75 & 31/150 & 2/15 \end{array} \right)$$

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

3 Variabile aleatoare discrete: Media

3.1 Media

Exemplul 1. Presupunem că avem un zar corect cu 6 fețe marcat cu 5 3-uri și un 6. Cât ne-am aștepta să fie media a 6000 de aruncări?

Răspuns. Dacă am fi știut valoarea fiecărei aruncări, am fi calculat media adunând cele 6000 de valori și împărțind la 6000. Fără să știm valorile, putem calcula [media](#) după cum urmează.

Deoarece sunt 5 3-uri și un 6 ne așteptăm ca aproximativ 5/6 dintre aruncări să fie 3 și 1/6 dintre ele să fie 6. Presupunând acest fapt exact adevărat, avem următorul tabel de valori și numărări:

valoarea:	3	6
numărări așteptate:	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 așteptată în sensul că "ne așteptăm" ca fiecare dintre valorile posibile să apară cu frecvențele date.

Exemplul 2. Aruncăm 2 zaruri corecte. Câștigăm 1000 de dolari dacă suma este 2 și pierdem 100 de dolari altfel. Cât ne așteptăm să câștigăm în medie pe aruncare?

Răspuns. Probabilitatea unui 2 este 1/36. Dacă jucăm de N ori, ne putem "aștepta" ca $\frac{1}{36} \cdot N$ dintre aruncări să dea 2 și $\frac{35}{36} \cdot N$ dintre aruncări să dea altceva. Astfel, câștigurile așteptate totale sunt

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

Pentru a obține media așteptată pe aruncare, împărțim totalul la N :

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

Gândiți: Ați vrea să jucați acest joc o dată? De mai multe ori?

Observăm că în ambele exemple suma pentru media așteptată constă din termeni care sunt o valoare a unei variabile aleatoare ori probabilitatea ei. Aceasta conduce la următoarea definiție.

Definiție. Presupunem că X este o variabilă aleatoare discretă care ia valorile x_1, x_2, \dots, x_n cu probabilitățile $p(x_1), p(x_2), \dots, p(x_n)$. **Media** lui X este notată $E(X)$ și definită prin

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

Observații:

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 **tendinței centrale** a unei variabile aleatoare.
4. Dacă toate valorile sunt egal probabile, media este chiar media aritmetică a valorilor.
5. Fie X cu o mulțime numărabilă de valori x_1, x_2, \dots luate cu probabilitățile $p(x_1), p(x_2), \dots$. Atunci, media lui X este

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

dacă seria este convergentă.

Exemplul 3. Aflați $E(X)$ 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 $X \sim \text{Bernoulli}(p)$. Aflați $E(X)$.

Răspuns. X ia valorile 0 și 1 cu probabilitățile $1 - p$ și p , deci

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

Important: Media unei variabile aleatoare Bernoulli(p) este p .

Gândiți: Care este media sumei a 2 zaruri corecte?

3.1.1 Media și centrul de masă

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

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

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



Centrul de masă este

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

Numim această formulă o medie "ponderată" a lui x_1 și 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 probabilitățile $p(x_j)$ în locul maselor. Putem spune că "media este punctul în care repartiția s-ar echilibra". Observați similaritatea dintre exemplul din fizică și exemplul 1.

3.1.2 Proprietăți algebrice ale lui $E(X)$

Când adunăm, scădam sau translatăm variabilele aleatoare, mediile fac același lucru. Modul matematic prescurtat de a spune aceasta este că $E(X)$ este [liniară](#).

1. Dacă X și Y sunt variabile aleatoare pe un spațiu al probelor Ω , atunci

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

2. Dacă a și b sunt constante, atunci

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

$aX + b$ se obține prin [scalarea](#) lui X cu a și apoi [translatarea](#) cu b .

Exemplul 6. Aruncăm 2 zaruri corecte și fie X suma lor. Aflați $E(X)$.

Răspuns: Fie X_1 numărul de pe primul zar și 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 $X \sim \text{binomial}(n, p)$. Aflați $E(X)$.

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

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

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

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

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 ușoară.

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

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

Încercați să calculați 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 Demonstrațiile proprietăților algebrice ale lui $E(X)$

Demonstrația proprietății (1) este simplă, dar are o subtilitate în chiar înțelegerea 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 și Y înseamnă a aduna valorile lui X și Y pentru același 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 $P(\omega)$:	$P(\omega_1)$	$P(\omega_2)$	$P(\omega_3)$...	$P(\omega_n)$

Demonstrația lui (1):

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

Demonstrația proprietății (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 $\sum_i p(x_i) = 1$.

Exemplul 9. Media unei repartiții geometrice

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

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

Demonstrația necesită un truc.

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

Suma seriei geometrice: $\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$.

Derivăm ambii membri: $\sum_{k=0}^{\infty} kx^{k-1} = \frac{1}{(1-x)^2}$.

Înmulțim cu x : $\sum_{k=0}^{\infty} kx^k = \frac{x}{(1-x)^2}$.

Înlocuim x cu $1-p$: $\sum_{k=0}^{\infty} k(1-p)^k = \frac{1-p}{p^2}$.

Înmulțim cu p : $\sum_{k=0}^{\infty} kp(1-p)^k = \frac{1-p}{p}$.

Ultima expresie este media.

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

Exemplul 10. Aruncăm o monedă corectă până obținem un avers pentru prima dată. Care este numărul așteptat 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 așteptat de aruncări libere pe care le-ar marca înaintea primei ratări?

Răspuns: Acesta este un exemplu unde vrem numărul de succese dinaintea primului eșec. Folosind limbajul de aversuri și reversuri: succesul este revers (probabilitate = $1 - p$) și eșecul este avers (probabilitate = p). De aceea $p = 0.2$ și 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 funcțiilor de o variabilă aleatoare

([Formula schimbării de variabile](#))

Dacă X este o variabilă aleatoare luând valorile x_1, x_2, \dots și h este o funcție, 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 și $Y = X^2$. Aflați $E(Y)$.

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 aceeași ca cea pentru valoarea corespunzătoare a lui X . Deci,

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

Exemplul 13. Aruncăm 2 zaruri corecte și fie X suma. Presupunem că funcția 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 ușor 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)
```

Obținem 13.83333, deci $E(Y) = 13.8(3)$.

Deoarece plata așteptată 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!

Gândiți: Este adevărat în exemplul precedent?

Întrebare: Dacă $Y = 3X + 77$, rezultă $E(Y) = 3E(X) + 77$?

Răspuns: Da. Din proprietatea (2), scalarea și translatarea se comportă astfel.