

Tikimybių teorijos pagrindai

- Įvadas

- Kombinatorika

- Tikimybės apibrėžimas

- Stochastiniai kintamieji

 - Pasiskirstymo funkcijos

 - Momentai

 - Charakteristinės funkcijos

 - Daugiamačiai stochastiniai kintamieji

- Binominiai skirstiniai

 - Binominis skirstinys

 - Gauso (normalusis) skirstinys

 - Puasono skirstinys

- Atsitiktinis klaidžiojimas

- Centrinė ribinė teorema ir didelių skaičių dėsnis

Įvadas

- Įvairios fenomenologinės teorijos, pvz. termodinamika, nustato sąryšius tarp įvairių sistemos makroskopinių dydžių tokių kaip temperatūra, slėgis, dalelių skaičius ir pan. Tačiau jos nieko nesako apie mikroskopinius sistemos procesus, t.y. iš jų neaišku kaip mikroskopiniai procesai įtakoja makroskopinį elgesį.

Mikroskopinės teorijos, tiriančios sistemas su dideliu laisvės laipsnių skaičiumi, remiasi tikimybių teorija. Be to chaoso teorija, kuri paprastai nagrinėja sistemas su nedideliu laisvės laipsnių skaičiumi taip pat naudoja statistikos metodus. Todėl čia trumpai apžvelgsime pagrindinius tikimybės teorijos teiginius, kurie mums bus reikalingi tolimesniam dėstymui.

- Mes pradėsime nuo **kombinatorikos** dėstymo ir pateiksime intuityvų tikimybės apibrėžimą, bei pagrindines **elementariosios tikimybių teorijos formules**.
- **Stochastiniai kintamieji (atsitiktiniai dydžiai)**, pagal apibrėžimą yra kintamieji, kurių vertės nustatomos iš eksperimento. Daugiausia ką mes galime sužinoti apie stochastinį kintamąjį, yra tikimybė, kad tam tikra jo vertė realizuosis eksperimente. Mes apibrėšime **pasiskirstymo funkcijų** sąvokas **tiesk diskretingiems tiesk ir tolydiesiems** stochastiniams kintamiesiems ir parodysime kaip nustatyti jų momentus panaudojus **charakteristines funkcijas**. Mes detalai aptarsime **binominį skirstinį** ir jo du ribinius atvejus – **Gausinį ir Puasono skirstinius**.
- Pagaliau mes išnagrinėsime **atsitiktinio klaidžiojimo uždavinį ir aptarsime centrinę ribinę teoremą bei didelių skaičių dėsnį**.

Kombinatorika (I)

- Taikant tikimybių teoriją realiose situacijose dažnai susiduriama su sudėtingomis kombinacijų apskaičiavimo problemomis. Tokiais atvejais pravartu turėti omenyje du labai svarbius principus:

- **Sudėties principas:** Jeigu vieną elementą galima pasirinkti iš dviejų skirtingų grupių turinčių m_1 ir m_2 elementų ir abiejose grupėse nėra vienodų elementų, tai elemento pasirinkimo būdų yra $m_1 + m_2$. Panašus sudėties principas galioja, kai yra daugiau nei dvi elementų grupės.

Pavyzdys: Vieną vaisių iš 4 kriaušių, 5 obuolių ir 3 slyvų galima pasirinkti 12 būdų.

- **Daugybės principas:** Jeigu du elementus galima pasirinkti iš dviejų skirtingų grupių (vieną iš vienos, o kita iš kitos) turinčių m_1 ir m_2 elementų, tai tų elementų pasirinkimo būdų yra $m_1 \times m_2$. Panašus daugybės principas galioja, kai yra daugiau nei dvi elementų grupės.

Pavyzdys: Keliais būdais galima sudaryti užrakto kodą, kai pirmasis simbolis yra raidė iš penkių galimų, o kiti trys simboliai - skaitmenys.

Sprendimas: Pirmą simbolį galima pasirinkti 5 būdais, antrą - 10 būdų, trečią - 10 būdų ir ketvirtą - 10 būdų. Todėl yra $5 \times 10 \times 10 \times 10 = 5000$ būdų, kuriais galima sudaryti užrakto kodą.

- Kai operuojama dideliais objektų skaičiais, dažnai tenka skaičiuoti įvairių objektų **junginių** skaičius.

Junginiai (kombinacijos) gaunami įvairiai jungiant tam tikros baigtinės aibės elementus.

Pavyzdys: Dviem krepšinio komandom reikia pasirinkti sportinę aprangą, kai yra trijų spalvų marškinėliai: geltoni (G), balti (B) ir mėlyni (M). Tuomet iš elementų aibės {G,B,M} gauname tokius junginius (kombinacijas):
(G,B); (G,M); (B,G); (B,M); (M,G); (M,B).

Junginyje elementų išdėstymo tvarka gali būti svarbi arba nesvarbi. Pagal tai junginiai klasifikuojami į **gretinius** ir **derinius**.

Gretiniais vadinami junginiai, kuriuose **elementai nesikartoja ir elementų išdėstymo tvarka yra svarbi**, t.y. sukeitus elementus vietomis gaunamas naujas junginys. Skirtingų gretinių skaičius imant iš N elementų po k yra:

$$P_N^k = N(N-1)(N-2) \times \cdots \times (N-k+1) = \frac{N!}{(N-k)!}$$

Pavyzdys: Kiek skirtingų trispalvių vėliavų galima pasiūti iš 5 skirtingų spalvų audeklų?

Sprendimas:

$$P_5^3 = \frac{5!}{(5-3)!} = 60$$

Kartotiniai gretiniai vadinami junginiai iš m elementų, kai juos galima rinktis iš N elementų aibės, **nekreipiant dėmesio į tai, kad elementas jau buvo pasirinktas**. Kartotinių gretinių skaičius yra:

$$\overline{P}_N^m = N^m$$

Pavyzdys: Kiek galima sudaryti penkiaženklų skaičių iš skaitmenų 2, 5, 9?

Sprendimas: $\overline{P}_3^5 = 3^5 = 243$

Kėliniais (perstatiniais) vadinami gretiniai sudaryti iš **visų** elementų aibės. Kėlinių skaičius yra $P_N = N(N-1) \times \dots \times 1 = N!$

Pavyzdys: Kiek būdų galima išrikiuoti 10 knygų lentynoje?

Sprendimas: $P_{10} = 10! = 3628800$.

Kartotiniu kėliniu vadinamas kėlinys, kuris turi vienodų n_1 elementų pirmos rūšies, n_2 vienodų elementų antros rūšies, ..., n_k vienodų elementų k rūšies. Kartotinių kėlinių skaičius yra:

$$P(n_1, n_2, \dots, n_k) = \frac{N!}{n_1! n_2! \dots n_k!}, \quad n_1 + n_2 + \dots + n_k = N$$

Pavyzdys: Kiek skirtingų žodžių galima parašyti perstatant raides žodyje STATISTIKA?

Sprendimas: Raidė S kartojasi 2 kartus, raidė T – 3 kartus, raidė A – 2 kartus, raidė I – 2 kartus. Bendras raidžių skaičius $N = 10$. Todėl skirtingų junginių iš šių raidžių yra $10!/(2!3!2!2!) = 75600$.

Deriniu vadinamas junginys, kuriame elementai nesikartoja ir jų išdėstymo **tvarka nėra svarbi**, t.y. sukeitus elementus vietomis gaunamas tas pats junginys. Skirtingų derinių skaičius imant iš N elementų po k yra:

$$C_N^k \equiv \binom{N}{k} = \frac{N!}{(N-k)!k!}$$

Iš tikrųjų, k skirtingų objektų turi $k!$ perstatinių (kėlinių). Todėl $k!C_N^k = P_N^k$, t.y. sudauginę kėlinių skaičių $k!$ iš derinių skaičiaus C_N^k gausime gretinių skaičių P_N^k .

Pavyzdys: Kiek skirtingų startinių penketukų galima sudaryti iš 10 krepšininukų?

Sprendimas:

$$C_{10}^5 = \frac{10!}{(10-5)!5!} = 252$$

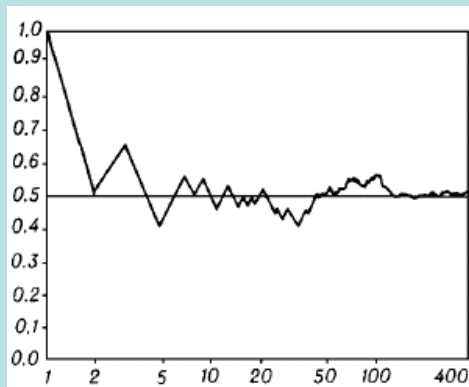
Tikimybės apibrėžimas (I)

Tikimybė yra kiekybinis laukiamo įvykio ar eksperimento įvertinimas. Tarkime, kad vienas iš galimų eksperimentinių rezultatų yra A . Tuomet tikimybė, kad eksperimente bus stebimas A yra $P(A)$, jeigu mes *tikimės*, kad atlikus N identiškų eksperimentų rezultatas A bus stebimas $N_A = NP(A)$ eksperimentuose. Mes *tikimės*, kad kai N artėja prie begalybės, tai santykinis eksperimentų skaičius (dažnis) N_A/N , kuriuose bus stebimas A artės prie $P(A)$.

- Labai svarbus atskiras atvejis, kai eksperimente gali būti stebimi n **vienodai tikimi rezultatai**. Jeigu m iš šių rezultatų atitinka įvykį A , tai $P(A) = m/n$.

Pavyzdys: Mėtant lošimo kauliuką natūralu manyti, kad kiekvienos sienelės (pažymėtos 1,2,3,4,5 ir 6 taškais) atsivertimo galimybės yra visiškai vienodos. Taigi turime $n=6$ vienodai tikimus rezultatus. Pažymėkime įvykiu A sienelės su nelyginiu taškų skaičiumi (1,3, arba 5) atsivertimą. Tuomet $m=3$ ir $P(A)=3/6=1/2$.

Pavyzdys:



Herbo atvirtimo dažnio kitimas (logaritminėje skalėje), metus monetą 400 kartu.

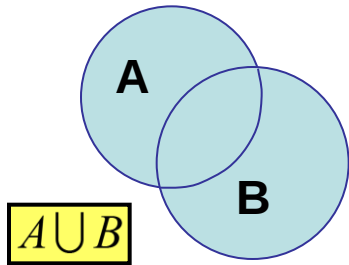
Analizuojant eksperimentinius rezultatus tenka nustatinėti sąryšius tarp įvairių tikimybių. Tokiu atveju patogiu apibrėžti **elementarių įvykių erdvės** Ω sąvoką. Ω – yra aibė visų elementarių įvykių $\{\omega_1, \omega_2, \dots, \omega_N\}$ susijusių su koku nors eksperimentu taip, kad bet kuris eksperimentinis rezultatas atitinka vieną ar daugiau Ω aibės elementų. **Elementarieji įvykiai** pagal apibrėžimą yra **neskaidytini ir negali įvykti kartu**. Bet kuris įvykis A atitiks tam tikrą Ω erdvės poerdvį.

Tuomet įvykio A tikimybę galima nustatyti atlikus tokius veiksmus:

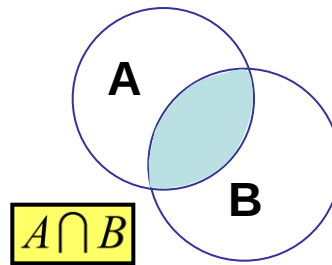
1. Sudaryti visų galimų eksperimentinių rezultatų įvykių erdvę Ω .
2. Priskirti įvykių erdvės elementams (erdvės taškams) tikimybių reikšmes. Jeigu įvykių erdvė sudaryta iš N vienodai tikimų įvykių tai kiekvienam taškui priskiriama tikimybė $1/N$.
3. Kad nustatytume įvykio A tikimybę turime susumuoti visų elementų tikimybes, kurie atitinka erdvės Ω įvykio A poerdvį.

Pavyzdys: Lošimo kauliuko elementarių įvykių erdvė Ω sudaryta iš 6 įvykių $\{\omega_1, \omega_2, \dots, \omega_6\}$. Čia ω_k reiškia k taškų atsivėrimą. Kiekvieno iš šių įvykių tikimybė yra $1/6$. Kitus įvykius galėsime sudaryti iš tų šešių įvykių. Sakysime, įvykis "atsivertė lyginis taškų skaičius" bus sudarytas iš trijų elementariųjų įvykių $\{\omega_2, \omega_4, \omega_6\}$. Jo tikimybė yra $P(\omega_2)+P(\omega_4)+P(\omega_6)=1/2$. Įvykis "atsivertė sveikas taškų skaičius" yra būtinas – jis sudarytas iš visų šešių elementariųjų įvykių. Jo tikimybė yra $P(\omega_1)+P(\omega_2)+\dots+P(\omega_6)=1$.

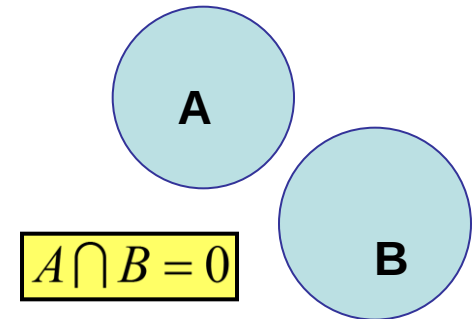
Tikimybių teorijoje dažnai naudojami aibių teorijos žymėjimai.



*A ir B įvykių **sajunga** (suma), kai įvyksta bent vienas iš įvykių A ir B.*



*A ir B įvykių **sankirta** (pjūvis), kai kartu įvyksta abu įvykiai A ir B.*



*A ir B **nesutaikomi** įvykiai, kai A ir B neturi bendrų elementų.*

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

$P(A \cup B) = P(A) + P(B)$ – tikimybė, kad įvyks bent vienas iš įvykių A ir B, kai A ir B **nesutaikomi** įvykiai.

Pavyzdys: Dėžėje yra 10 vienodo dydžio, tačiau skirtingų spalvų rutuliukai: trys - raudoni (R), du – balti (B) ir penki – geltoni (G). Akivaizdu, kad tikimybės ištraukti atitinkamų spalvų rutuliukus yra: $P(R)=3/10$, $P(B)=2/10=1/5$, $P(G)=5/10=1/2$. Kokią tikimybę ištraukti raudoną arba baltą rutuliuką?

Sprendimas: Įvykiai „ištrauktas raudonas rutuliukas“ ir „ištrauktas baltas rutuliukas“ yra **nesutaikomi**. Todėl tikimybė ištraukti raudoną arba baltą rutuliuką lygi šių dviejų įvykių tikimybių sumai:

$$P(R \cup B) = P(R) + P(B) = 3/10 + 1/5 = 1/2$$

$$P(A | B) = \frac{P(A \cap B)}{P(B)} \quad - \text{ **sąlyginė tikimybė (apibrėžimas)**, kad įvyks } A, \text{ jeigu įvyko } B$$

$$P(A | B)P(B) = P(B | A)P(A) \quad - \text{ nes } P(A \cap B) = P(B \cap A)$$

Sąlyginė įvykio A tikimybė tai faktiškai tikimybė, kurioje visa įvykių erdvė Ω pakeičiama įvykio B poerdviu.

Pavyzdys: Metamas lošimo kauliukas. Kokia tikimybė, kad atsivers 2 taškai, jeigu žinome jog atsivertė lyginis taškų skaičius?

Sprendimas: Įvykio „atsivertė 2 taškai (įvykis A) ir lyginis taškų skaičius (įvykis B)“ tikimybė $P(A \cap B) = P(A) = 1/6$. Įvykio „atsivertė lyginis taškų skaičius (įvykis B)“ tikimybė $P(B) = 1/2$. Todėl **sąlyginė tikimybė**, kad atsivers 2 taškai, kai žinome jog atsivertė lyginis taškų skaičius yra:

$$P(A | B) = P(A \cap B) / P(B) = (1/6) / (1/2) = 1/3$$

Kita vertus, sąlyginę įvykio A tikimybę galima apskaičiuoti kaip tikimybę, kurioje visa įvykių erdvė Ω (visi galimi kauliuko taškai): $\{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$ pakeičiama įvyko B poerdviu (lyginiai taškų skaičiai): $\{\omega_2, \omega_4, \omega_6\}$. Skaičiuojant sąlyginę tikimybę, palankių įvykių skaičių (toks įvykis tik vienas, ω_2) daliname iš poerdvio B įvykių skaičiaus (3) ir gauname tą patį rezultatą:

$$P(A | B) = 1/3$$

- Įvykiai A ir B vadinami **nepriklausomais**, jeigu kiekvienas iš jų nepriklauso nuo to, ar kitas įvykis realizavosi ar ne.

Jeigu įvykis A nepriklauso nuo B tai sąlyginė tikimybė:

$$P(A | B) = P(A)$$

Tuomet iš $P(A \cap B) = P(A | B)P(B)$ turime:

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

– jeigu A ir B **nepriklausomi** įvykiai, tai tikimybė, kad įvyks abu įvykiai A ir B , lygi tų įvykių **tikimybių sandaugai**.

Pavyzdys: Kartu metami du lošimo kauliukai. Kokia tikimybė, kad abiejuose kauliukuose kartu atsivers 5 taškai?

Sprendimas: Įvykiai „pirmame kauliuke atsivertė 5 taškai (įvykis A)“ ir „antrame kauliuke atsivertė 5 taškai (įvykis B)“ yra **nepriklausomi**. Todėl tikimybė, kad abiejuose kauliukuose kartu atsivers 5 taškai lygi šių dviejų įvykių tikimybių sandaugai:

$$P(A \cap B) = P(A)P(B) = (1/6)(1/6) = 1/36$$

Tarkime, kad A_i yra **nesutaikomi** įvykiai, kurie apima **visą** įvykių erdvę Ω , t.y.

$$A_i \cap A_j = \emptyset \ (i \neq j), \quad \bigcup_i A_i = \Omega \quad \longrightarrow \quad \sum_i P(A_i \cap B) = P(\bigcup_i A_i \cap B) = P(\Omega \cap B) = P(B)$$

$$P(A_i \cap B) = P(A_i | B)P(B) = P(B | A_i)P(A_i) \quad \longrightarrow \quad \sum_i P(B | A_i)P(A_i) = P(B)$$

$$P(A_i | B) = \frac{P(B | A_i)P(A_i)}{P(B)} = \frac{P(B | A_i)P(A_i)}{\sum_k P(B | A_k)P(A_k)} \quad - \text{Bayes formulė}$$

Pavyzdys: Televizijos žaidime pretendentai rodomos trys durys: už vieną iš jų yra brangi dovana – automobilis, už dviejų kitų pigios – kaklaraiščiai. Pretendentui reikia pasirinkti duris. Tarkime jis pasirinko 1-as. Tuomet vedantysis atidaro **kitas**, tarkime 2-as duris už kurių yra kaklaraištis. Toliau pretendentui vėl siūloma pasirinkti duris. Ar jam likti prie to paties pasirinkimo ar rinktis 3 duris? Kaip tokioje situacijoje pasielgtumėte Jūs?

Sprendimas: Pasinaudodami Bayes formule nustatysime optimalų pasirinkimą.

$$P(B|A_1) = P(2-os durys atidarytos po to kai pasirinktos 1-os | automobilis už 1-ą durų) = 1/2$$

$$P(B|A_2) = P(2-os durys atidarytos po to kai pasirinktos 1-os | automobilis už 2-ą durų) = 0$$

$$P(B|A_3) = P(2-os durys atidarytos po to kai pasirinktos 1-os | automobilis už 3-ą durų) = 1.$$

Tarsime, kad $P(A_i)=1/3$. Tuomet $P(B) = (1/2)(1/3)+0(1/3)+1(1/3) = 1/2$ ir

$$P(A_1|B) = P(\text{automob. už 1-ą durų} | 2-os d. atidarytos po to kai pasirinktos 1-os) = \frac{(1/2)(1/3)}{1/2} = \frac{1}{3}$$

$$P(A_3|B) = P(\text{automob. už 3-ą durų} | 2-os d. atidarytos po to kai pasirinktos 1-os) = \frac{(1)(1/3)}{1/2} = \frac{2}{3}$$

Taigi matome, kad **pasirinkimą reikia keisti**, t.y. rinktis 3-as duris.

Stochastiniai kintamieji (I)

Stochastinis kintamasis (arba atsitiktinis dydis) – tai dydis, kurio reikšmės priklauso nuo atsitiktinio eksperimento rezultatų. Tiksliau kalbant, stochastinis kintamasis yra elementariųjų įvykiu funkcija, t.y. kiekvienam elementariam įvykiui pagal tam tikrą taisyklę (funkciją) priskiriamas realusis dydis. Stochastinį dydį žymėsime didele raide X , o jo galimas vertes $\{x_k\}$ – mažomis.

Pavyzdys: Mėtome kartu tris monetas. Herbų atsivertimo skaičius kiekvieno metimo metu yra stochastinis kintamasis kurio vertės yra sveikieji skaičiai kintantys nuo 0 iki 3. Tai **diskretinio** stochastinio kintamojo pavyzdys.

Pavyzdys: Dujose molekulės koordinatė, grietis, energija ir pan. yra **tolydiniai** stochastiniai kintamieji.

Pasiskirstymo funkcijos

Diskretiniai stochastiniai kintamieji

• Tarkime, kad X yra diskretinis atsitiktinis dydis kuris gali įgyti vertes $\{x_1, x_2, \dots, x_n\}$. Tegul dydžių x_k tikimybės yra p_k taip, kad $\sum_{k=1}^n p_k = 1$. Tuomet kintamojo X **tikimybės tankio funkcija** apibrėžiama taip:

$$P_X(x) = \sum_{k=1}^n p_k \delta(x - x_k) \quad (\delta - \text{Dirako delta funkcija})$$

o kintamojo X **pasiskirstymo funkcijos** apibrėžimas yra toks:

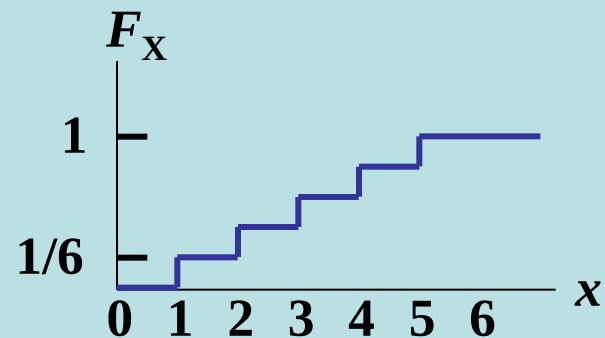
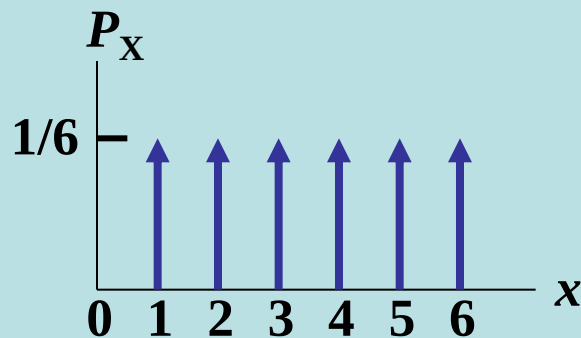
$$F_X(x) = \int_{-\infty}^x P_X(y) dy = \sum_{k=1}^n p_k \Theta(x - x_k)$$

($\Theta(x)$ – Heavisaido funkcija, t.y. $\Theta(x) = 0$, kai $x < 0$ ir $\Theta(x) = 1$, kai $x > 0$.)

$F_X(x)$ yra monotoniškai auganti funkcija, su kraštinėmis sąlygomis: $F_X(-\infty) = 0, F_X(\infty) = 1$. Skirstinys $P_X(x)$ yra pasiskirstymo funkcijos išvestinė, $P_X(x) = dF_X(x)/dx$. Pasiskirstymo funkcijos reikšmė $F_X(x)$ yra tikimybė, kad atsitiktinio dydžio X vertė yra intervale $(-\infty, x)$.

Pavyzdys: Panagrinėkime atsitiktinį dydį X , kurio vertės atitinka lošimo kauliuko taškų skaičius, $x_k = k, k = 1, 2, \dots, 6$. Mėtant kauliuką, kiekvienos vertės tikimybė $p_k = 1/6$. Nupiešime šios sistemos $P_X(x)$ ir $F_X(x)$ funkcijas.

Sprendimas



Tolydiniai stochastiniai kintamieji

- Tarkime, kad X yra tolydinis atsitiktinis dydis. Pagal stochastinio kintamojo apibrėžimą, intervalas $a \leq x \leq b$ atitiks tam tikrą įvykį. **Tikimybės tankio funkcijos** $P_X(x)$ apibrėžimas:

$$\Pr(a \leq x \leq b) = \int_a^b P_X(x) dx \quad - \text{tikimybė, kad } X \text{ yra intervale } a \leq x \leq b, \text{ arba}$$

$$P_X(x) dx \quad - \text{tikimybė, kad } X \text{ yra intervale } [x, x + dx].$$

Savybės: $P_X(x) \geq 0, \quad \int_{-\infty}^{\infty} P_X(x) dx = 1$ (normavimo sąlyga).

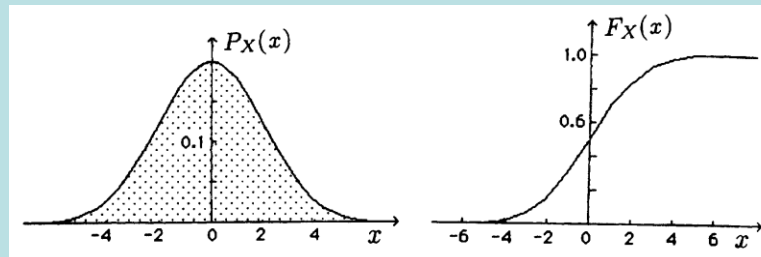
Kartais reikia pereiti prie **naujo kintamojo** $Y = H(X)$. Tuomet skirstinys $P_Y(y)$ surandamas taip:

$$P_Y(y) = \int_{-\infty}^{\infty} \delta(y - H(x)) P_X(x) dx \quad (\delta - \text{Dirako delta funkcija})$$

$$F_X(x) = \int_{-\infty}^x P_X(y) dy \quad - \text{pasiskirstymo funkcija (tikimybė, kad } X \text{ yra intervale } [-\infty, x]).$$

Savybės: $F_X(-\infty) = 0, F_X(\infty) = 1$

Pavyzdys: Gauso skirstiniui $P_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ pavaizduosime $P_X(x)$ ir $F_X(x)$ grafikus.
Sprendimas:



Momentai

- Jeigu mums pavyko surasti stochastinio kintamojo X skirstinį $P_X(x)$, tai reiškia, kad mes turime to kintamojo visą įmanomą informaciją. Praktiniuose uždaviniuose pilnai surasti skirstinį dažnai nepavyksta. Tačiau, kai skirstinys nėra pilnai žinomas, pavyksta nustatyti informaciją apie jo momentus. X kintamojo n **momentas** apibrėžiamas taip:

$$\langle x^n \rangle = \int_{-\infty}^{\infty} x^n P_X(x) dx$$

- Kai kurie momentai turi specialius pavadinimus. Momentas $\langle x \rangle$ vadinamas X **vidurkiu**. Kombinacija $\langle x^2 \rangle - \langle x \rangle^2 = \langle (x - \langle x \rangle)^2 \rangle$ vadinama **vidutiniu kvadratinu nuokrypiu** (variance), o **standartiniu nuokrypiu** vadinamas dydis:

$$\sigma_X = \sqrt{\langle x^2 \rangle - \langle x \rangle^2} = \sqrt{\langle (x - \langle x \rangle)^2 \rangle}$$

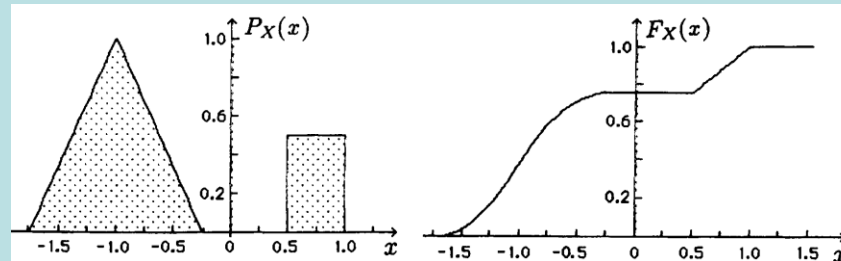
- Momentai suteikia informaciją apie skirstinio plotį ir formą. Svarbiausi yra žemų eilių momentai, nes jie apibūdina pagrindines skirstinio savybes, o aukštesni momentai, charakterizuoja skirstinio detales. Pademonstruosime tai pavyzdžiais.

Pirmasis momentas

- Pirmasis momentas $\langle x \rangle$ nusako skirstinio „masių centro“ padėtį. Jis kartais painiojimas su labiausiai tikėtina verte x_p ir mediana x_m . **Labiausiai tikėtina vertė** x_p atitinka tašką x kuriame $P_X(x)$ yra didžiausia. **Mediana** x_m yra x vertė, kuri dalina plotą po $P_X(x)$ kreive į dvi vienodas dalis, kitaip sakant $F_X(x_m) = 1/2$.

Praeitame pavyzdyje nagrinėjome Gauso skirstinį $P_X(x) = \frac{1}{2\sqrt{2\pi}} e^{-x^2/8}$. Dėl šio skirstinio simetrijos visos trys aukščiau apibrėžtos vertės sutampa, būtent $\langle x \rangle = x_p = x_m = 0$. Toliau demonstruojamas pavyzdys, kuriame visos šios vertės yra skirtingos.

Pavyzdys: Nustatyti $\langle x \rangle$, x_p ir x_m skirstiniui, kurio funkcijos $P_X(x)$ ir $F_X(x)$ pavaizduotos žemiau pateiktuose grafikuose.



Sprendimas: vidurkis $\langle x \rangle = -0.5625$, labiausiai tikėtina vertė $x_p = -1$, mediana $x_m = -0.862$.

Antrasis momentas

- Antrasis momentas $\langle x^2 \rangle$ nusako skirstinio „inercijos momentą“ koordinačių pradžios atžvilgiu. Standartinis nuokrypis σ_X rodo skirstinio plotį vidutinės vertės $\langle x \rangle$ atžvilgiu. Užpraeitam pavyzdyje, Gauso skirstiniui, $\langle x \rangle = 0$, $\langle x^2 \rangle = 4$ ir $\sigma_X = 2$. Praeitame pavyzdyje, $\langle x \rangle = -0.5625$, $\langle x^2 \rangle = 0.966$, $\sigma_X = 0.806$.

Trečiasis momentas

- Trečiasis momentas $\langle x^3 \rangle$ charakterizuoja skirstinio asimetriją koordinačių pradžios atžvilgiu. Gauso skirstinys yra simetrinis ir trečiasis momentas lygus nuliui, $\langle x^3 \rangle = 0$. Praeitame pavyzdyje, $\langle x^3 \rangle = -0.844$. Tai rodo, kad didesnė skirstinio dalis yra neigiamoje pusėje, $x < 0$.

Charakteristinės funkcijos

- Atsitiktinio kintamoji X **charakteristinė funkcija** apibrėžiama kaip **tikimybės tankio Furjė transformacija**:

$$f_X(k) = \langle e^{ikx} \rangle = \int_{-\infty}^{\infty} e^{ikx} P_X(x) dx = \sum_{n=0}^{\infty} \frac{(ik)^n \langle x^n \rangle}{n!}$$

- Pastarasis skleidimas eilute turi prasmę tik tuomet, kai aukštųjų eilių momentai yra maži, t.y. kai eilutė konverguoja. Iš šios išraiškos matome, kad norint vienareikšmiškai nustatyti skirstinį $P_X(x)$, reikia žinoti visus jo momentus.

Savybės: $f_X(0) = 1, \quad |f_X(k)| \leq 1, \quad f_X(-k) = f_X^*(k)$

- Jeigu charakteristinė funkcija yra žinoma, $P_X(x)$ galima surasti **atvirkštine transformacija**

$$P_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ikx} f_X(k) dk$$

Diferencijuojant charakteristinę funkciją galima nesunkiai surasti visus momentus:

$$\langle x^n \rangle = \lim_{k \rightarrow 0} (-i)^n \frac{d^n f_X(k)}{dk^n}$$

- Kartais charakteristinę funkciją patogiau skleisti **komuliantais** (comulant) $C_n(X)$, kurių apibrėžimas yra toks:

$$f_X(k) = \exp\left(\sum_{n=1}^{\infty} \frac{(ik)^n}{n!} C_n(X)\right) = \sum_{n=0}^{\infty} \frac{(ik)^n \langle x^n \rangle}{n!} \quad \left(\ln(f_X(k)) = \sum_{n=1}^{\infty} \frac{(ik)^n}{n!} C_n(X) \right)$$

- Išskleidę eksponentę k laipsniais ir sulyginę koeficientus prie vienodų k laipsnių gausime komuliantų sąryšius su momentais. Pirmųjų 4 komuliantų sąryšiai su momentais yra tokie:

$$C_1(X) = \langle x \rangle$$

$$C_2(X) = \langle x^2 \rangle - \langle x \rangle^2 = \langle (x - \langle x \rangle)^2 \rangle = \sigma_x^2$$

$$C_3(X) = \langle x^3 \rangle - 3\langle x \rangle \langle x^2 \rangle + 2\langle x \rangle^3 = \langle (x - \langle x \rangle)^3 \rangle$$

$$C_4(X) = \langle x^4 \rangle - 3\langle x^2 \rangle^2 - 4\langle x \rangle \langle x^3 \rangle + 12\langle x \rangle^2 \langle x^2 \rangle - 6\langle x \rangle^4 \neq \langle (x - \langle x \rangle)^4 \rangle$$

Iš čia matome, kad pirmas komuliantas yra X vidurkis, antras komuliantas yra vidutinis kvadratinis nuokrypis, trečias komuliantas yra vidutinis kubinis nuokrypis, tačiau ketvirtos eilės komuliantas nelygus ketvirtos eilės vidutiniam nuokrypiui.

Pavyzdys: Panagrinėsime skirstinį $P_X(x) = \frac{2}{\pi} \sqrt{1-x^2}$, kai $|x| \leq 1$ ir $P_X(x) = 0$, kai $|x| > 1$. Surasime pirmus keturis momentus ir komuliantus.

Sprendimas: $f_X(k) = \int_{-1}^1 e^{ikx} \sqrt{1-x^2} dx = \frac{2}{k} J_1(k)$

Beselio funkciją $J_1(k)$ skleidžiame Teiloro eilutę: $f_X(k) = \frac{2}{k} \left[\frac{k}{2} - \frac{k^3}{16} + \frac{k^5}{384} + \dots \right] = 1 - \frac{1}{4} \frac{k^2}{2!} + \frac{1}{8} \frac{k^4}{4!} - \dots$

Gauname: $\langle x \rangle = \langle x^3 \rangle = 0, \langle x^2 \rangle = 1/4, \langle x^4 \rangle = 1/8; \quad C_1 = C_3 = 0, C_2 = 1/4, C_4 = -1/16$

Daugiamačiai stochastiniai kintamieji

• Labai dažnai eksperimento rezultatams aprašyti neužtenka vieno atsitiktinio dydžio, bet reikia jų sistemos. Sakykime, iš patrankos šaudome į nejudantį plokščią taikinį. Pataikymo taškas nurodomas dviem dydžiais. Judančiai dujų molekulei apibudinti taip pat reikia kelių atsitiktinių dydžių. Nusakant kelių atsitiktinių dydžių sistemos savybes, nepakanka nurodyti atskirų atsitiktinių dydžių savybių – dar reikia žinoti ir ryšius tarp tų dydžių. Todėl tenka kurti kelių atsitiktinių, arba daugiamačių (vektorinių) atsitiktinių dydžių teoriją.

- Stochastinių kintamųjų $X_1, \dots, X_n$ ***daugiamatės pasiskirstymo funkcijos*** apibrėžiamas:

$$F_{X_1, \dots, X_n}(x_1, \dots, x_n) = \Pr(X_1 < x_1, \dots, X_n < x_n) - \text{tikimybė, kad visų kartu kintamųjų } X_k \text{ vertės yra intervaluose } -\infty < X_k < x_k$$

Daugiamatė tikimybės tankio funkcija (joint probability density) apibrėžiama taip:

$$P_{X_1, \dots, X_n}(x_1, \dots, x_n) = \frac{\partial F_{X_1, \dots, X_n}(x_1, \dots, x_n)}{\partial x_1 \cdots \partial x_n}$$

Integruojant tankio funkciją galima surasti pasiskirstymo funkciją:

$$F_{X_1, \dots, X_n}(x_1, \dots, x_n) = \int_{-\infty}^{x_1} dx_1 \cdots \int_{-\infty}^{x_n} dx_n P_{X_1, \dots, X_n}(x_1, \dots, x_n)$$

Panagrinėkime **dvimatį** tikimybės tankį $P_{X,Y}(x,y)$ ir pasiskirstymo funkciją $F_{X,Y}(x,y)$.

Savybės:

$$F_{X,Y}(-\infty, y) = F_{X,Y}(x, -\infty) = F_{X,Y}(-\infty, -\infty) = 0$$

$$0 \leq P_{X,Y} \leq 1$$

$$\int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy P_{X,Y}(x, y) = F_{X,Y}(\infty, \infty) = 1$$

$$F_X(x) = F_{X,Y}(x, \infty) = \int_{-\infty}^x dx' \int_{-\infty}^{\infty} dy P_{X,Y}(x', y)$$

$$P_X(x) = \int_{-\infty}^{\infty} dy P_{X,Y}(x, y)$$

Momentai:

$$\langle x^n \rangle = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy x^n P_{X,Y}(x, y)$$

$$\langle x^m y^n \rangle = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy x^m y^n P_{X,Y}(x, y)$$

Fizikoje dažniausiai vartojami momentai yra **kovariacija (covariance)**

$$\text{Cov}(x, y) = \langle (x - \langle x \rangle)(y - \langle y \rangle) \rangle = \langle xy \rangle - \langle x \rangle \langle y \rangle$$

ir **koreliacijos koeficientas**

$$\text{Cor}(x, y) = \frac{\langle (x - \langle x \rangle)(y - \langle y \rangle) \rangle}{\sigma_X \sigma_Y}$$

Koreliacijos koeficientas yra bedimensinis dydis, kuris **charakterizuoja atsitiktinių kintamųjų** X ir Y **tarpusavio priklausomybę**. Koreliacijos koeficiento **savybės**:

(a) $\text{Cor}(X, Y) = \text{Cor}(Y, X)$;	(b) $\text{Cor}(X, X) = 1, \text{Cor}(X, -X) = -1$;
(c) $-1 \leq \text{Cor}(X, Y) \leq 1$;	(d) $\text{Cor}(aX + b, cY + d) = \text{Cor}(X, Y)$.

Jeigu du stochastiniai kintamieji X ir Y **nepriklausomi**, tai galioja tokios lygybės:

$$\begin{aligned} (1) \quad & P_{X,Y}(x,y) = P_X(x)P_Y(y); & (2) \quad & \langle XY \rangle = \langle X \rangle \langle Y \rangle; \\ (3) \quad & \langle (X+Y)^2 \rangle - \langle (X+Y) \rangle^2 = \langle X^2 \rangle - \langle X \rangle^2 + \langle Y^2 \rangle - \langle Y \rangle^2; & (4) \quad & Cor(X,Y) = 0. \end{aligned}$$

Atkreipkime dėmesį į tai, kad iš (4) sąlygos neišplaukia atvirkštinis tvirtinimas, t.y. jeigu $Cor(X,Y) = 0$, tai stochastiniai kintamieji X ir Y nebūtinai nepriklausomi.

Pavyzdys: Pademonstruosime, kad koreliacijos koeficientas $Cor(X,Y)$ iš tikrųjų yra stochastinių kintamųjų X ir Y tarpusavio priklausomybės matas.

Sprendimas: Tarkime, kad tarp kintamųjų X ir Y yra tiesinis ryšys $X = aY + b$.

Panagrinėkime vidutinę kvadratinę paklaidą

$$e = \langle (x - ay - b)^2 \rangle$$

Parinkime koeficientus a ir b taip, kad ši paklaida būtų minimali:

$$\delta e = \frac{\partial e}{\partial a} \delta a + \frac{\partial e}{\partial b} \delta b = 0 \Rightarrow \begin{aligned} \frac{\partial e}{\partial a} = 0 & \Rightarrow -2\langle xy \rangle + 2a\langle y^2 \rangle + 2b\langle y \rangle = 0 \\ \frac{\partial e}{\partial b} = 0 & \Rightarrow -2\langle x \rangle + 2a\langle y \rangle + 2b = 0 \end{aligned}$$

Išsprendę lygčių sistema a ir b atžvilgiu gauname

$$a = \frac{\langle xy \rangle - \langle x \rangle \langle y \rangle}{\langle y^2 \rangle - \langle y \rangle^2} = \frac{Cov(X,Y)}{\sigma_Y^2} = \frac{\sigma_X}{\sigma_Y} Cor(X,Y)$$

Matome kad, kai $Cor(X,Y)=0$, tai $a=0$, t.y. kintamieji X ir Y yra tiesiškai nepriklausomi.

- Kai turime reikalą su daugiamačiais stochastiniais kintamaisiais dažnai prireikia **pakeisti kintamąjį**, t.y. surasti skirstinį naujiems kintamiesiems, kurie yra senų kintamųjų funkcijos. Tarkime mes žinome dviejų kintamųjų tankio funkciją $P_{X,Y}(x,y)$ ir norime surasti kintamojo Z skirstinį $P_Z(z)$, kai tas kintamasis yra tam tikra X ir Y funkcija, $Z=G(X,Y)$. Tai daroma pagal tokią formulę:

$$P_Z(z) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \delta(z - G(x, y)) P_{X,Y}(x, y)$$

Iš čia galime nesunkiai surasti kintamojo Z charakteristinę funkciją:

$$f_Z(k) = \int_{-\infty}^{\infty} dz \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \delta(z - G(x, y)) e^{ikz} P_{X,Y}(x, y) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy e^{ikG(x,y)} P_{X,Y}(x, y)$$

Pavyzdys: Tarkime X ir Y yra nepriklausomi kintamieji pasiskirstę pagal Gausą ($\langle x \rangle = \langle y \rangle = 0$, $\sigma_X = \sigma_Y = 1$). Surasime kintamųjų $V = X + Y$ ir $W = X - Y$ skirstinį. Ar V ir W nepriklausomi?

Sprendimas: $P_X(x) = (1/\sqrt{2\pi})e^{-x^2/2}$, $P_Y(y) = (1/\sqrt{2\pi})e^{-y^2/2}$. Kadangi X ir Y nepriklausomi, tai

$$P_{V,W}(v, w) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \delta(v - (x + y)) \delta(w - (x - y)) e^{-(x^2 + y^2)/2}$$

$$dx dy = J dv dw, \quad \begin{matrix} v = x + y & x = (v + w)/2 \\ w = x - y & y = (v - w)/2 \end{matrix}, \quad J = \left| \frac{\partial(x, y)}{\partial(v, w)} \right| = \frac{1}{2} \quad (\text{Jakobianas})$$

$$P_{V,W}(v, w) = (1/4\pi) e^{-(v^2 + w^2)/4} = (1/2\sqrt{\pi}) e^{-v^2/4} (1/2\sqrt{\pi}) e^{-w^2/4} = P_V(v) P_W(w)$$

Kintamieji V ir W **nepriklausomi**, nes $P_{V,W}(v, w) = P_V(v) P_W(w)$!

Binominiai skirstiniai (I)

- Tikimybių teorijos taikymas sąlyginai paprastas, kai eksperimente galimi **tik du rezultatai**. Pavyzdžiui, mėtant monetą gali išvirsti tik pinigas arba herbas. Simetrinės monetos atveju šių dviejų įvykių tikimybės yra vienodos. Mes nagrinėsime bendresnį atvejį (tarkim, nesimetrinę monetą, arba krepšininko baudų metimus), kai dviejų galimų įvykių (krepšininko atveju pataikymo ir nepataikymo) tikimybės p ir q gali būti skirtingos ($p \neq q$, $p + q = 1$).
- Skirstinys, kuris aprašo tikimybę vienos iš realizacijų (pavyzdžiui baudos pataikymo), kai atliekama N nepriklausomų eksperimentų vadinamas **binominiu skirstiniu**. Didelių N skaičių riboje binominis skirstinys pereina į **Gauso**, arba **Puasono** skirstinius. Tai priklauso nuo p ir q tikimybių santykio. Mes panagrinėsime visus šiuos tris skirstinius ir binominio skirstininio teoriją pritaikysime atsitiktinio klaidžiojimo uždavinyje.

Binominis skirstinys

- Tarkime, kad eksperimente stebimo atsitiktinio kintamojo vertės gali būti tik 0 ir 1.

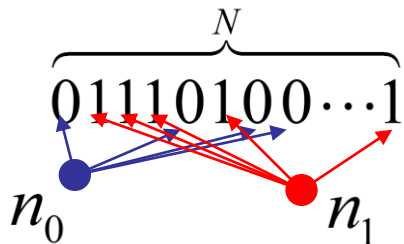
Pažymėkime:

q – nulio (0) pasirodymo tikimybė (pvz. baudos nepataikymo)

akivaizdu, kad $p+q=1$

p – vieneto (1) pasirodymo tikimybė (pvz. baudos pataikymo)

- Panagrinėkime N nepriklausomų bandymų seką, pavyzdžiui



n_0 – nulių skaičius sekoje

n_1 – vienetų skaičius sekoje

akivaizdu, kad $n_0+n_1=N$

$q^{n_0} p^{n_1}$ – tikimybė **vienos duotos sekos** iš N bandymų, kurioje realizavosi n_0 nulių ir n_1 vienetų (atskiri bandymai statistiškai nepriklausomi).

$\frac{N!}{n_0!n_1!}$ – sekų skaičius, duodančių tą patį rezultatą, t.y. n_0 nulių ir n_1 vienetų

Tikimybė, kad sekoje iš N bandymų bus n_0 nulių ir n_1 vienetų:

$$P_N(n_1) = \frac{N!}{n_0!n_1!} q^{n_0} p^{n_1}$$

– Tai yra **binominis skirstinys**
($p+q=1$, $n_0+n_1=N$)

Pavyzdys: Krepšininko baudų pataikymo tikimybė yra $1/3$. Kokia yra tikimybė, kad iš penkių metimų serijos bent trys metimai bus tikslūs?

Sprendimas: Pažymėkime n_1 pataikymų skaičių iš penkių metimų serijos. Tuomet $N = n_0 + n_1 = 5$, $p = 1/3$, $q = 2/3$. Tikimybė pataikyti n_1 krepšį n_1 kartų iš $N=5$ metimų yra

$$P_5(n_1) = \frac{5!}{n_1!(5-n_1)!} \left(\frac{1}{3}\right)^{n_1} \left(\frac{2}{3}\right)^{5-n_1}$$

Tikimybė pataikyti mažiausiai tris kartus $= P_5(3) + P_5(4) + P_5(5) = 51/243 \approx 0.21$

- Binominis skirstinys tenkina normavimo sąlygą:

$$\sum_{n_1=0}^N P(n_1) = \sum_{n_1=0}^N \frac{N!}{n_1!(N-n_1)!} p^{n_1} q^{N-n_1} = (p+q)^N = 1$$

- Binominiam skirstiniui galima užrašyti tikimybės tankio funkciją, panaudojus δ Dirako funkciją. Pažymėkime eksperimente stebimą stochastinį kintamąjį X_k , kurio vertės gali būti 0 arba 1. Apibrėžkime stochastinį kintamąjį $Y_N = X_1 + X_2 + \dots + X_N$, kuris atitinka N serijos eksperimentinės sekos sumą. Atkreipkime dėmesį į tai, kad **sumoje indėlį duoda tik tie eksperimentai, kuriuose matuojamas dydis $X_k=1$** . Kintamojo Y_N **tikimybės tankio funkcija** yra tokia:

$$P_{Y_N}(y) = \sum_{n_1=0}^N P_N(n_1) \delta(y - n_1)$$

$$P_{Y_N}(y) = \sum_{n_1=0}^N P_N(n_1) \delta(y - n_1)$$

Charakteristinė funkcija:

$$f_{Y_N}(k) = \int_{-\infty}^{\infty} dy e^{iky} P_{Y_N}(y) = \sum_{n_1=0}^N e^{ikn_1} \frac{N! p^{n_1} q^{N-n_1}}{n_1! (N-n_1)!} = (q + p e^{ik})^N$$

Momentai:

$$\langle y^m \rangle = \int_{-\infty}^{\infty} dy y^m \sum_{n_1=0}^N P_N(n_1) \delta(y - n_1) = \sum_{n_1=0}^N n_1^m P_N(n_1) \equiv \langle n_1^m \rangle$$

$$\langle y \rangle = \langle n_1 \rangle = \sum_{n_1=0}^N n_1 P_N(n_1) = \sum_{n_1=0}^N \frac{n_1 N! p^{n_1} q^{N-n_1}}{n_1! (N-n_1)!} = p \frac{\partial}{\partial p} (p + q)^N = pN$$

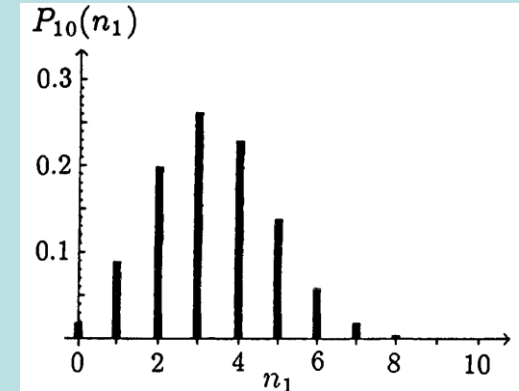
$$\langle y^2 \rangle = \langle n_1^2 \rangle = \sum_{n_1=0}^N n_1^2 P_N(n_1) = (Np)^2 + Npq$$

$$\langle n_1^2 \rangle - \langle n_1 \rangle^2 = Npq, \quad \sigma_N = \sqrt{Npq}$$

Santykinis nuokrypis:

$$\frac{\sigma_N}{\langle n_1 \rangle} = \sqrt{\left\langle \left(\frac{n_1/N}{p} - 1 \right)^2 \right\rangle} = \sqrt{\frac{q}{p}} \frac{1}{\sqrt{N}}$$

Pavyzdys



Binominis skirstinys:

$N=10, p=1/3$ ($\langle n_1 \rangle = 10/3$).

(Įmestų baidų n_1 pasiskirstymas serijose iš 10 metimų, kai vienos baidos pataikymo tikimybė yra 1/3)

Santykinis nuokrypis $\sigma_N / \langle n_1 \rangle$ rodo kiek dydis n_1/N , nusakantis vienetuko pasirodymo dažnį sekoje iš N bandymų, nukrypsta nuo laukiamos vertės p bet kurioje sekoje iš N bandymų. Mažos santykinio nuokrypio vertės reiškia, kad n_1/N yra artimas p ir, kai $N \rightarrow \infty$, tai $n_1/N \rightarrow p$.

Gauso (normalusis) skirstinys

Binominiame skirstinyje pakeiskime kintamąjį:

$$Z_N = \frac{Y_N - \langle y \rangle}{\sigma_Y} = \frac{Y_N - pN}{\sqrt{pqN}}$$

Tuomet $\langle z \rangle = 0$, $\sigma_Z = 1$ ir kintamojo Z tikimybės tankio funkcija yra:

$$P_{Z_N}(z) = \int_{-\infty}^{\infty} dy \delta\left(z - \frac{y - pN}{\sqrt{pqN}}\right) P_{Y_N}(y)$$

Surasime charakteristinę funkciją:

$$\begin{aligned} f_{Z_N}(k) &= \int_{-\infty}^{\infty} dz e^{ikz} P_{Z_N}(z) \\ &= \int_{-\infty}^{\infty} dz \int_{-\infty}^{\infty} dy e^{ikz} \delta\left(z - \frac{y - pN}{\sqrt{pqN}}\right) P_{Y_N}(y) \\ &= e^{-ik\sqrt{pN/q}} \int_{-\infty}^{\infty} dy e^{iky/\sqrt{pqN}} P_{Y_N}(y) \\ &= e^{-ik\sqrt{pN/q}} f_{Y_N}\left(\frac{k}{\sqrt{pqN}}\right) = e^{-ik\sqrt{pN/q}} \left(q + pe^{ik/\sqrt{pqN}}\right)^N \\ &= \left(qe^{-ik\sqrt{p/qN}} + pe^{ik\sqrt{q/pN}}\right)^N \end{aligned}$$

Skačiuosime $f_{Z_N}(k)$ riboje $N \rightarrow \infty$
Pradžioje eksponentes skliaustose išskleisime Teiloro eilute:

$$f_{Z_N}(k) = \left[1 - \frac{k^2}{2N} + O\left(\frac{1}{N^{3/2}}\right)\right]^N$$

Kai $N \rightarrow \infty$, tai $O(1/N^{3/2})$ narius galima išbraukti, ir tada

$$f_Z(k) \equiv \lim_{N \rightarrow \infty} f_{Z_N}(k) = \lim_{N \rightarrow \infty} \left[1 - \frac{k^2}{2N}\right]^N = e^{-k^2/2}$$

Čia pasinaudojome gerai žinoma riba:

$$\lim_{N \rightarrow \infty} (1 + s/N)^N = e^s$$

Taigi riboje $N \rightarrow \infty$ gauname

$$P_Z(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{-ikz} e^{-k^2/2} = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

Kai $N \gg 1$ tačiau baigtinis mes turime apytikslę lygybę:

$$P_{Z_N}(z) \approx \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

Grįžkime prie pradinių kintamųjų

$$y = \sigma_Y z + \langle y \rangle$$

Tuomet gauname

$$\begin{aligned} P_{Y_N}(y) &\approx \int_{-\infty}^{\infty} dz \delta(y - \sigma_Y z - \langle y \rangle) P_Z(z) \\ &= \frac{1}{\sigma_Y \sqrt{2\pi}} \exp\left(-\frac{(y - \langle y \rangle)^2}{2\sigma_Y^2}\right) \end{aligned}$$

Tai yra **Gauso** arba **normalusis** skirstinys. Jis gautas iš binominio skirstinio riboje $N \rightarrow \infty$ tariant, kad p yra baigtinis.

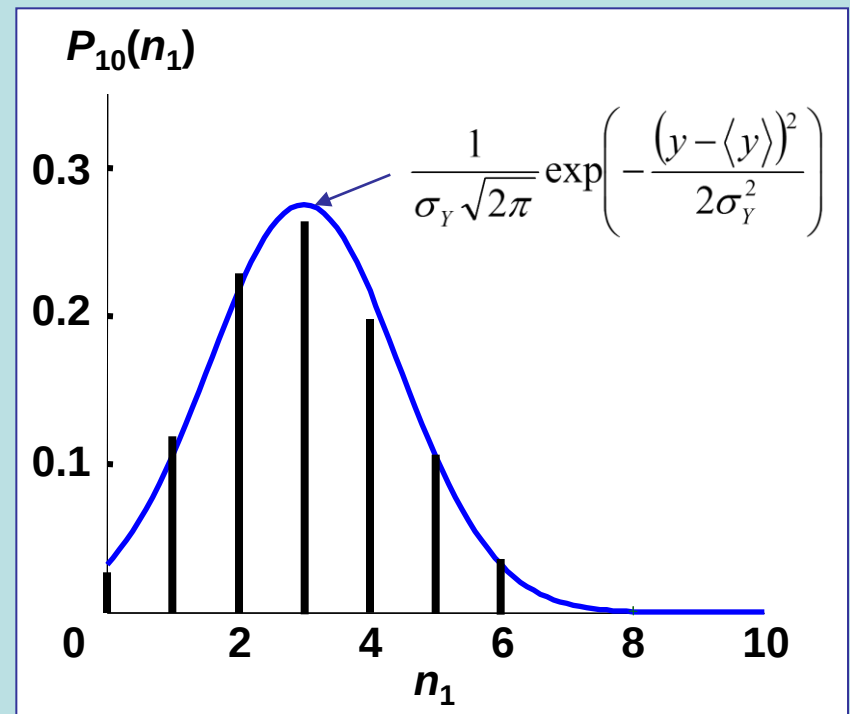
Gauso skirstinys priklauso tik nuo dviejų parametrų, momentų $\langle y \rangle$ ir $\langle y^2 \rangle$, arba nuo dviejų pirmų komuliantų:

$$C_1(Y) = \langle y \rangle, C_2(Y) = \sigma_Y^2 = \langle y^2 \rangle - \langle y \rangle^2$$

Galima nesunkiai įrodyti, kad **Gauso skirstinio visi aukštesnės eilės komuliantai lygus nuliui**.

Gauso skirstinys gerai aproksimuoja binominį skirstinį, netgi, kai N nėra labai didelis.

Pavyzdys:



Binominio ir Gauso skirstinių palyginimas.
Binominio sk. parametrai: $N=10, p=1/3$.
Gauso sk. parametrai: $\langle y \rangle = pN, \sigma_Y = \sqrt{Np(1-p)}$

Puasono skirstinys

Puasono skirstinys gaunamas iš binominio riboje:

$$N \rightarrow \infty, p \rightarrow 0, Np = a = \text{const.}$$

Įrašome į binominį skirstinį $p=a/N$ ir apskaičiuojame ribą $N \rightarrow \infty$

$$P_N(n_1) = \frac{N!}{n_1!(N-n_1)!} \left(\frac{a}{N}\right)^{n_1} \left(1 - \frac{a}{N}\right)^{N-n_1}$$

Nesunku apskaičiuoti tokias ribas:

$$\lim_{N \rightarrow \infty} \frac{N!}{(N-n_1)! N^{n_1}} = \lim_{N \rightarrow \infty} \frac{N(N-1)\cdots(N-n_1+1)}{N^{n_1}} = 1$$

$$\lim_{N \rightarrow \infty} \left(1 - \frac{a}{N}\right)^{N-n_1} = e^{-a}$$

Tuomet

$$\lim_{N \rightarrow \infty} P_N(n_1) = \frac{a^{n_1}}{n_1!} e^{-a} \equiv P(n_1)$$

Taigi **Puasono skirstinys** yra toks:

$$P(n_1) = \frac{a^{n_1}}{n_1!} e^{-a}$$

Šis skirstinys **priklauso tik nuo vieno parametro** a . Nesunku įsitikinti, kad tai yra pirmasis momentas (vidurkis).

Momentai:

$$\langle n_1 \rangle = \sum_{n_1=0}^{\infty} n_1 \frac{a^{n_1}}{n_1!} e^{-a} = e^{-a} a \frac{\partial}{\partial a} \sum_{n_1=0}^{\infty} \frac{a^{n_1}}{n_1!} = e^{-a} a e^a = a$$

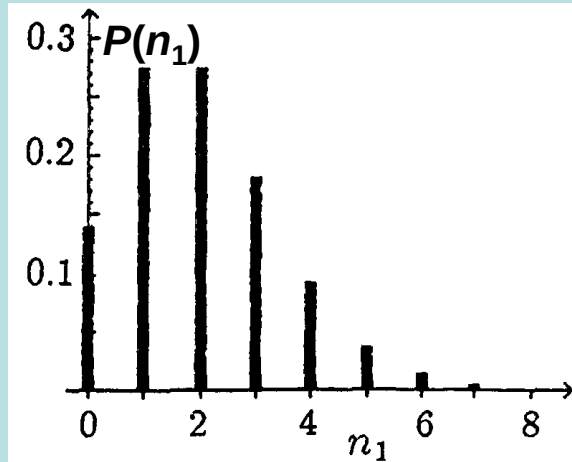
$$\langle n_1^2 \rangle = \sum_{n_1=0}^{\infty} n_1^2 \frac{a^{n_1}}{n_1!} e^{-a} = e^{-a} a^2 \frac{\partial^2}{\partial a^2} \sum_{n_1=0}^{\infty} \frac{a^{n_1}}{n_1!} + \langle n_1 \rangle = a^2 + a$$

$$\langle n_1^2 \rangle - \langle n_1 \rangle^2 = a = \langle n_1 \rangle, \sigma = \sqrt{\langle n_1 \rangle}$$

$$\frac{\sigma}{\langle n_1 \rangle} = \frac{1}{\sqrt{\langle n_1 \rangle}}$$

Puasono skirstinio santykinis nuokrypis atvirkščiai proporcingas šakniai iš vidurkio.

Pavyzdys



Puasono skirstinys, kai $a = \langle n_1 \rangle = 2$.

Pavyzdys: dalelių skaičiaus fluktuacijos mažame idealiųjų dujų turelyje

Tarkime, kad tūrio V inde yra N dalelių. Tuomet vidutinė dalelių koncentracija yra $n = N/V$. Išskirkime mažą tūrelį $V_1 \ll V$ ir nagrinėkime kaip, esant šiluminiam judėjimui, kinta dalelių skaičiaus šiame tūrelyje. Natūralu manyti, kad tikimybė surasti tam tikrą dalelę tūrelyje V_1 yra lygi tūrių santykiui, $p = V_1/V = nV_1/N$.

$$N, V$$

$$n = \frac{N}{V}$$

$$p = \frac{V_1}{V} = \frac{nV_1}{N}$$

$$N_1, V_1$$

Tikimybė surasti N_1 dalelių išskirtame tūrelyje tenkina binominį skirstinį:

$$P_N(N_1) = \frac{N!}{N_1!(N - N_1)!} p^{N_1} (1 - p)^{N - N_1}$$

Riboje $N \rightarrow \infty, nV_1 \equiv \bar{N}_1 = \text{const}$ tikimybė $p \rightarrow 0$ ir binominis skirstinys pereina į Puasoninį

$$P(N_1) = \frac{\bar{N}_1^{N_1}}{N_1!} e^{-\bar{N}_1}$$

$\bar{N}_1 \equiv \langle N_1 \rangle$ – vidutinis dalelių skaičius V_1 tūrelyje.

$$\langle N_1^2 \rangle - \langle N_1 \rangle^2 = \langle N_1 \rangle$$

Dalelių vidutinis kvadratinis nuokrypis mažame tūrelyje lygus dalelių skaičiui tame tūrelyje.

$$\frac{\sqrt{\langle N_1^2 \rangle - \langle N_1 \rangle^2}}{\langle N_1 \rangle} = \frac{1}{\sqrt{\langle N_1 \rangle}}$$

Santykinė fluktuacija tuo didesnė, kuo mažiau dalelių išskirtame tūrelyje.

Iš Puasono skirstinio istorijos

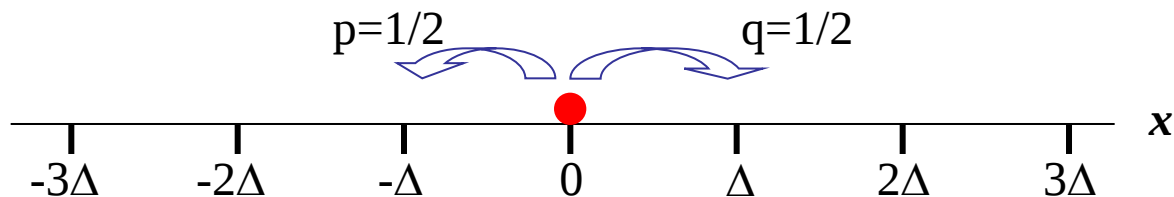
- Puasono skirstinys plačiai taikomas įvairiuose praktiniuose uždaviniuose. Jis galioja, kai yra didelė serija nepriklausomų įvykių, kurių tikimybė labai maža. Istoriskai šį skirstinį Puasonas taikė nagrinėdamas teisminių nuosprendžių statistiką. Apie tai 1837 m. jis parašė knygą. Puasono skirstinio svarba nebuvo iš karto derimai įvertinta ir 19 amžiaus viduryje šis darbas buvo pamirštas.



- Tačiau 19 amžiaus pabaigoje Puasono skirstinys buvo pritaikytas nagrinėjant labai keistą reiškinį. Vokietijoje, 14-oje skirtingų kavalerijos korpusų, buvo surinkta statistika apie tragiškas kareivių žūtis, kai jie buvo užmušti žirgo kanopos smūgiu. Statistiniai duomenys apėmė 20 metų laikotarpį nuo 1875m. iki 1894 m. Buvo atlikta 280 patikrinimų (280 patikrinimų = 20 metų x 14 korpusų) ir nustatyta, kad tokiu būdu žuvo 196 kareiviai. Taigi vienam patikrinimui atitiko vidutiniškai $\langle n \rangle = 196/280 = 0.7$ mirčių. Jeigu mirčių skaičius tenkintų Puasono skirstinį $P(n) = 0.7^n e^{-0.7} / n!$ su $\langle n \rangle = 0.7$, tai galima tikėtis, kad $280 \times P(0) \approx 139$ patikrinimuose neturėtų būti nė vienos (0) mirties, $280 \times P(1) \approx 97$ patikrinimuose – 1 mirtis, $280 \times P(2) \approx 34$ patikrinimuose – 2 mirtys, ir t.t. Tikri skaičiai buvo tokie: 140 patikrinimuose – 0 mirčių, 91 patikrinime – 1 mirtis, 32 patikrinimuose – 2 mirtys, ir t.t. Taigi buvo gautas puikus teorijos ir praktikos sutapimas. Po šio įspūdingo palyginimo Puasono skirstinį pradėjo plačiai taikyti.

Atsitiktinis klaidžiojimas (I)

- Atsitiktinio klaidžiojimo uždavinys – tai binominio skirstinio panaudojimo pavyzdys fizikoje. Tarkime, kad dalelė diskretiniu žingsniu Δ atsitiktinai juda x ašimi: su ta pačia tikimybe $p=q=1/2$ ji gali pajudėti tiek į kairę tiek į dešinę. Tegul pradžioje dalelė yra taške $x=0$. Mūsų tikslas surasti pasiskirstymo funkciją, kuri nusako dalelės koordinatę po N žingsnių. Tariame, kad kiekvienas žingsnis yra statistškai nepriklausomas nuo praeito.



- Pastebėsime, kad šis uždavinys yra ekvivalentus monetos mėtymo uždaviniui. Tarkime kad herbas atitinka dalelės judėjimą į kairę, o pinigas – į dešinę. Tuomet skirtumas tarp iškritusių pinigų ir herbų skaičiaus serijoje iš N metimų atitiks klaidžiojančios dalelės koordinatę po N žingsnių. Kartais šis uždavinys formuluojamas, kaip girto žmogaus klaidžiojimo uždavinys ir formuluojamas klausymo pavidalo: kaip toli gali nuklysti girtas žmogus?
- Šį uždavinį galima išspręsti pasitelkus kombinatorikos metodus. Tačiau čia mes pademonstruosime kitą sprendimo būdą, kuriame **nenaudojamas kombinatorikos mokslas**. Šis būdas **grindžiamas charakteristinės funkcijos** teorija. Toks sprendimo būdas yra gana universalus ir juo nesunkiai galima gauti aukščiau aprašytus rezultatus. Pavyzdžiui, juo nesunkiai galima išvesti binominio skirstinio formulę, nenaudojant kombinatorikos mokslo.

i -to žingsnio metu stochastinio kintamo X_i realizacijos yra $x+\Delta$, kai žingsniuojama į dešinę, ir $x-\Delta$, kai žingsniuojama į kairę. i -to žingsnio tikimybės tankio funkcija yra

$$P_{X_i} = \frac{1}{2}(\delta(x + \Delta) + \delta(x - \Delta))$$

Šios tikimybės charakteristinė funkcija

$$f_{X_i} = \int_{-\infty}^{\infty} dx e^{ikx} P_{X_i} = \cos(k\Delta)$$

Pažymėkime Y_N stochastinį kintamąjį, kuris nusako dalelės koordinatę po N žingsnių

$$Y_N = X_1 + \dots + X_N$$

Kadangi žingsniai nepriklausomi, tai tikimybės tankio funkcija kintamajam Y_N yra

$$P_{Y_N}(y) = \int \dots \int dx_1 \dots dx_N \delta(y - x_1 - \dots - x_N) P_{X_1}(x_1) \dots P_{X_N}(x_N)$$

Surasime kintamojo Y_N charakteristinę funkciją

$$\begin{aligned} f_{Y_N}(k) &= \int dy e^{iky} P_{Y_N}(y) \\ &= \int dx_1 \dots \int dx_N e^{ik(x_1 + \dots + x_N)} P_{X_1}(x_1) \dots P_{X_N}(x_N) \\ &= f_{X_1}(k) \dots f_{X_N}(k) = (\cos(k\Delta))^N \end{aligned}$$

Matome, kad **nepriklausomų įvykių sumos charakteristinė funkcija lygi tų įvykių charakteristinių funkcijų sandaugai.**

Išskleidžiame charakteristinę funkciją

Teiloro eilute pagal k

$$f_{Y_N}(k) = \left(1 - \frac{k^2 \Delta^2}{2!} + \dots \right)^N \approx 1 - \frac{Nk^2 \Delta^2}{2!} + \dots$$

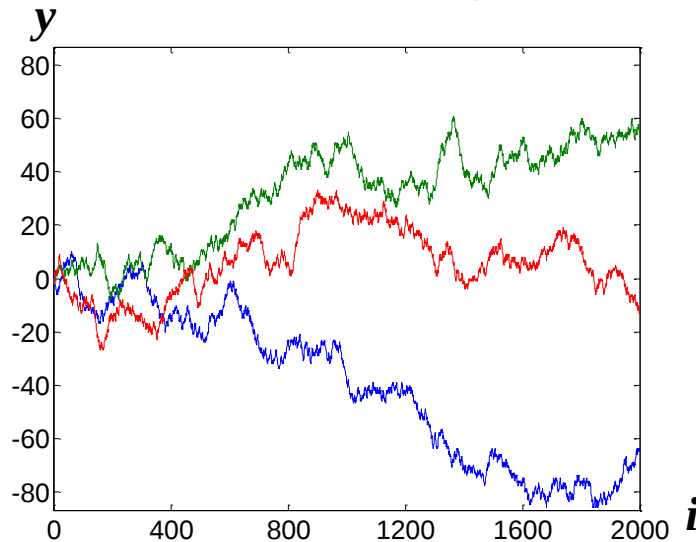
Dabar nesunkiai nustatome Y_N momentus

$$\langle y \rangle = \lim_{k \rightarrow 0} (-i) \frac{df_{Y_N}(k)}{dk} = 0 \quad \langle y^2 \rangle = \lim_{k \rightarrow 0} (-i)^2 \frac{d^2 f_{Y_N}(k)}{dk^2} = N\Delta^2$$

$$\sigma_{Y_N} = \Delta \sqrt{N}$$

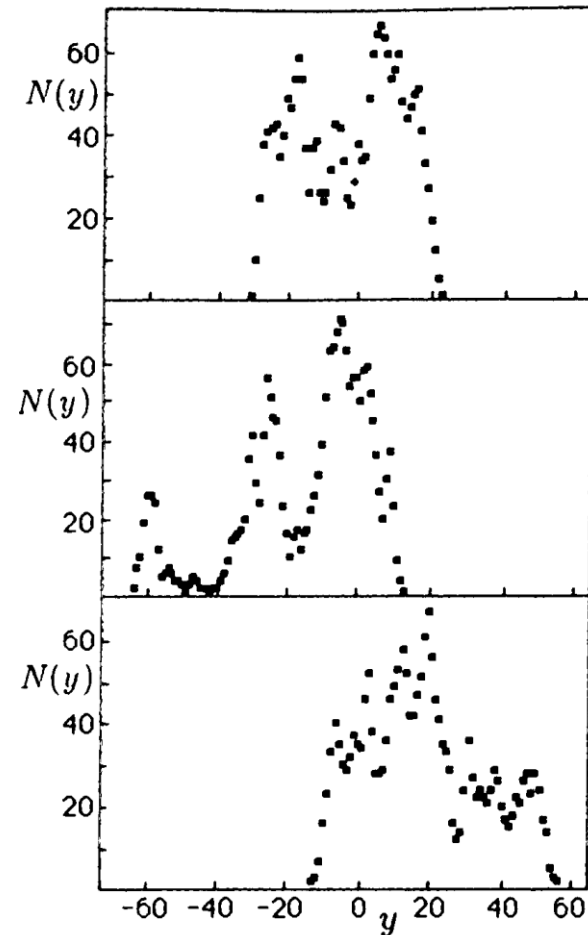
Taigi vidutinis dalelės nuokrypis nuo koordinatų pradžios lygus 0, o vidutinis kvadratinis nuokrypis proporcingas N .

Skaitmeniniai pavyzdžiai



Fiksuotai N vertei skirtingų galimų realizacijų skaičius yra 2^N . Čia pavaizduotos trys realizacijos ($N=2000$, $\Delta=1$). Pastebėkime, kad šį grafiką galima aiškinti dvejopai. Galima įsivaizduoti, kad tai **vienos dalelės trys trajektorijos**, atitinkančios tris nepriklausomus eksperimentus. Kita vertus, galima įsivaizduoti, kad tai **trijų vienu metu klaidžiojančių dalelių trajektorijos**. Pastaruoju atveju sakoma, kad stebimas vienodų dalelių **ansamblio** judėjimas. Ansamblio sąvoka plačiai taikoma fizikoje.

Atsitiktinis klaidžiojimas (III)



Čia yra alternatyvus atsitiktinio klaidžiojimo vaizdavimas. $N(y)$ yra skaičius, kuris rodo kiek kartų dalelė pasiekė tašką y . Čia taip pat pavaizduotos trys realizacijos, $N=2000$, $\Delta=1$.

Nykstamai mažo žingsnio riba

Riboje $\Delta \rightarrow 0, \tau \rightarrow 0, N \rightarrow \infty$ tikimybės tankio funkcijai ir charakteristinei funkcijai galima gauti diferencialines lygtis. Čia τ yra laikas, per kurį dalelė nueina vieną atsitiktinį žingsnį Δ . Per laiką $t = N\tau$ dalelė nueina N žingsnių. Pažymėkime

$$P_{Y_N}(y) \equiv P_Y(y, N\tau) = P_Y(y, t), \quad f_{Y_N}(k) \equiv f_Y(k, N\tau) = f_Y(k, t)$$

Tuomet

$$\frac{1}{\tau} \left| f_Y(k, (N+1)\tau) - f_Y(k, N\tau) \right| = (\cos(k\Delta) - 1) f_Y(k, N\tau) = \left(-\frac{k^2 \Delta^2}{2} + \dots \right) f_Y(k, N\tau)$$

$$\lim_{N \rightarrow \infty} \lim_{\tau \rightarrow 0} \frac{f_Y(k, (N+1)\tau) - f_Y(k, N\tau)}{\tau} = \frac{\partial f_Y(k, t)}{\partial t}, \quad \lim_{N \rightarrow \infty} \lim_{\tau \rightarrow 0} \lim_{\Delta \rightarrow 0} \left(-\frac{k^2 \Delta^2}{2\tau} + \dots \right) f_Y(k, N\tau) = -Dk^2 f_Y(k, t)$$

$$\frac{\partial f_Y(k, t)}{\partial t} = -Dk^2 f_Y(k, t)$$

$$D = \frac{\Delta^2}{2\tau} = \text{const.}$$

Pradinės sąlygos: $P_Y(y, 0) = \delta(y) \Rightarrow f_Y(k, 0) = 1$
(Kai $t=0$ dalelė yra taške $y=0$)

$$f_Y(k, t) = e^{-Dk^2 t}$$

$$P_Y(y, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{-iky} e^{-Dk^2 t} = \sqrt{\frac{1}{4\pi Dt}} \exp\left(\frac{-y^2}{4Dt}\right)$$

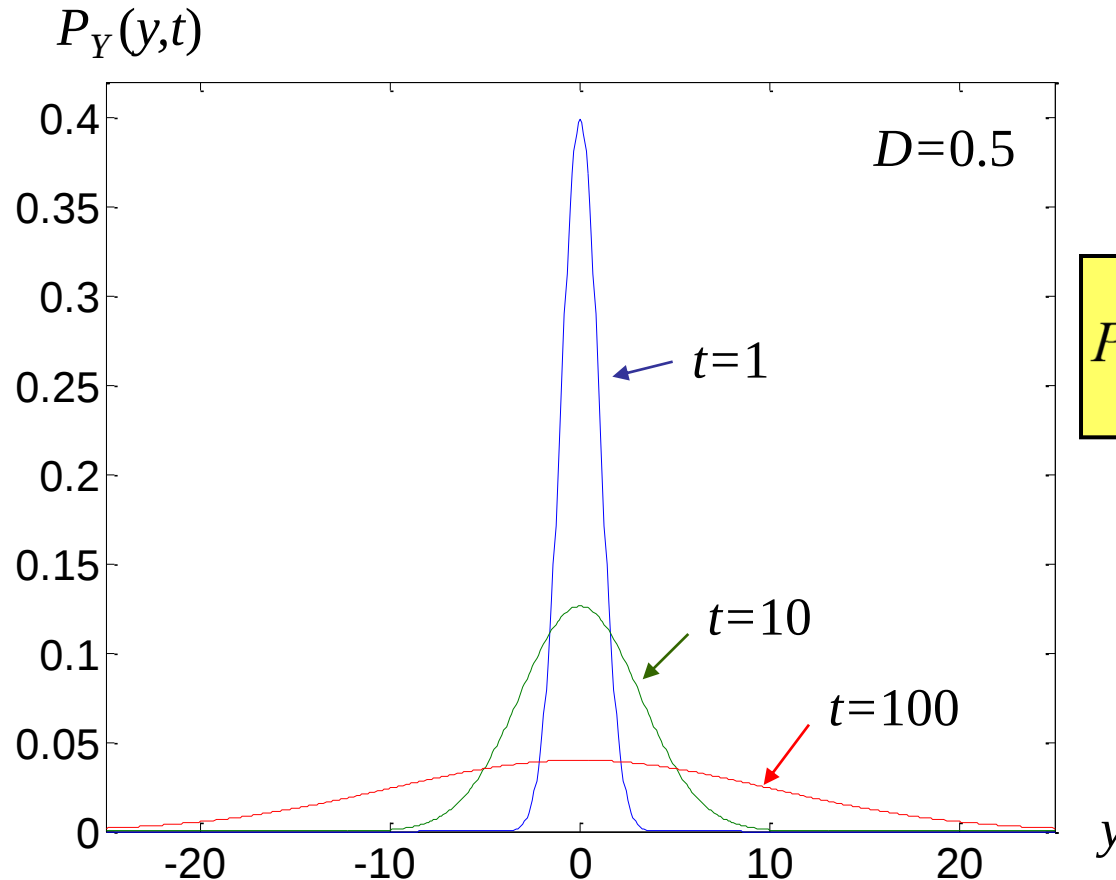
Atsitiktinis klaidžiojimas nusakomas
Gauso skirstiniu, kurio $\sigma = \sqrt{2Dt}$

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{-iky} \left| \frac{\partial f_Y(k, t)}{\partial t} \right| = -Dk^2 f_Y(k, t)$$

$$\frac{\partial P_Y(y, t)}{\partial t} = D \frac{\partial^2 P_Y(y, t)}{\partial y^2}$$

Tikimybės tankio funkcija tenkina difuzijos lygtį. Gauso funkcija yra jos sprendinys.

Skaitmeninės iliustracijos



$$P_Y(y, t) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(\frac{-y^2}{2\sigma^2}\right)$$

$$\sigma = \sqrt{2Dt}$$

*Laikui bėgant, tikimybės tankio funkcija plečiasi. Dalelių **ansamblio** interpretacijoje tai atitinka **difuzijos procesą**. Jeigu pradinio laiko momentu patalpiname daleles koordinatų pradžioje, tai jos difunduoja ir dalelių tankis kinta pagal Gauso dėsnį (animacija).*

Centrinė ribinė teorema

- Centrinė ribinė teorema teigia, kad bet kuris atsitiktinis dydis sudarytas iš daugelio atsitiktinių dydžių sumos turi universalią, Gauso pavidalo, tikimybės tankio funkciją.** Ši teorema paaiškina, kodėl Gauso skirstinys taip dažnai stebimas įvairiuose eksperimentuose. Panagrinėsime atskirą šios teoremos variantą, kuris aprašo eksperimentinių dydžių matavimo statistiką (paklaidų pasiskirstymą).

Tarkime, kad eksperimente N kartų matuojamas atsitiktinis dydis X ir skaičiuojamas vidurkis

$$Y_N = \frac{1}{N}(X_1 + \dots + X_N)$$

Tegul matavimo rezultatai yra statistiškai nepriklausomi. Nagrinėsime kintamojo Y_N nuokrypį nuo tikslaus kintamojo X vidurkio $\langle x \rangle$. Dėl to apibrėšime atsitiktinį dydį ΔY_N :

$$\Delta Y_N = Y_N - \langle x \rangle = Z_1 + \dots + Z_N$$

Čia

$$Z_i = \frac{1}{N}(X_i - \langle x \rangle)$$

Kintamojo Z tikimybės tankio ir charakteristinė funkcijos yra:

$$P_Z(z) \equiv \int_{-\infty}^{\infty} dx \delta\left(z - \frac{1}{N}(x - \langle x \rangle)\right) P_X(x)$$

$$\begin{aligned} f_Z(k) &= \int_{-\infty}^{\infty} dz e^{ikz} P_Z(z) \\ &= \int_{-\infty}^{\infty} dx e^{i(k/N)(x - \langle x \rangle)} P_X(x) = 1 - \frac{1}{2} \frac{k^2}{N^2} \sigma_X^2 + \dots \\ \sigma_X^2 &= \langle x^2 \rangle - \langle x \rangle^2 \end{aligned}$$

Kintamojo ΔY_N charakteristinė funkcija:

$$f_{\Delta Y_N}(k) = \left(1 - \frac{1}{2} \frac{k^2}{N^2} \sigma_X^2 + \dots\right)^N \xrightarrow{N \rightarrow \infty} \exp\left(-\frac{k^2 \sigma_X^2}{2N}\right)$$

Begalo didelių N riboje, kintamojo ΔY_N tikimybės tankio funkcija yra Gausinė:

$$P_{\Delta Y_N}(\Delta y) \xrightarrow{N \rightarrow \infty} \frac{1}{\sqrt{2\pi} \sigma_X} \int_{-\infty}^{\infty} dk e^{-ik\Delta y} \exp\left(-\frac{k^2 \sigma_X^2}{2N}\right) = \sqrt{\frac{N}{2\pi \sigma_X^2}} \exp\left(-\frac{N(\Delta y)^2}{2\sigma_X^2}\right)$$

Šis rezultatas nepriklauso nuo matuojamo dydžio skirstinio $P_X(x)$ formos (svarbu tik, kad jo momentai būtų baigtiniai)!



Didelių skaičių dėsnis

Didelių skaičių dėsnis teigia, kad jeigu mes turime N nepriklausomus atsitiktinio dydžio matavimo rezultatus X_i , $i=1, \dots, N$, tai tikimybė, kad matavimo vidurkis

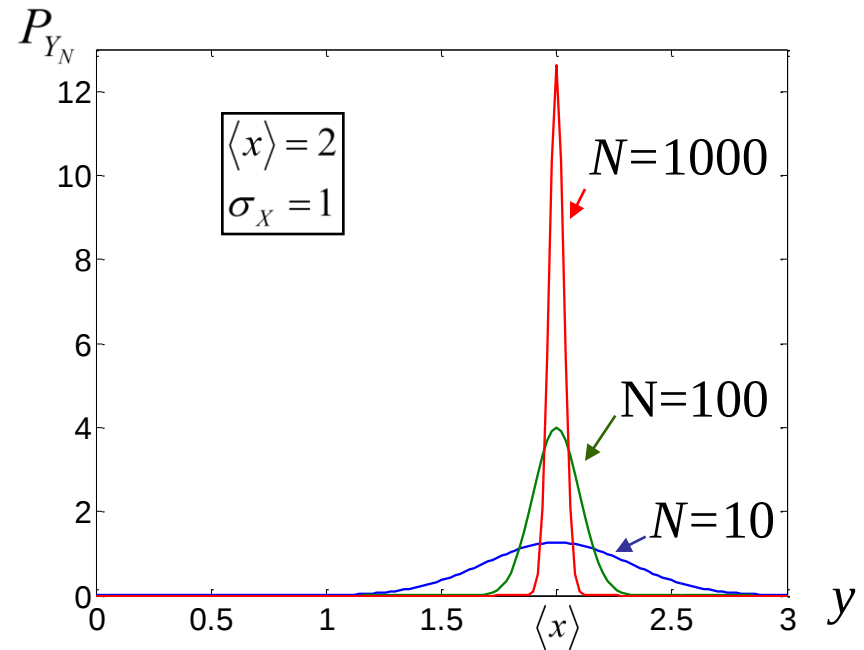
$$Y_N = \frac{1}{N} (X_1 + \dots + X_N)$$

nukryps nuo tikslios kintamojo X vidutinės vertės $\langle x \rangle$, artėja prie nulio, kai N artėja prie begalybės.

Šis teiginys iš esmės išplaukia iš centrinės ribinės teoremos. Mes nustatėme, kad didelių N riboje kintamojo Y_N tikimybės tankio funkcija artėja prie Gausinio skirstinio, kurį galima užrašyti taip:

$$P_{Y_N}(y) \xrightarrow{N \rightarrow \infty} \frac{1}{\sigma_Y \sqrt{2\pi}} \exp\left(-\frac{(y - \langle x \rangle)^2}{2\sigma_Y^2}\right)$$

$$\sigma_Y = \frac{\sigma_X}{\sqrt{N}} = \sqrt{\frac{\langle x^2 \rangle - \langle x \rangle^2}{N}} \xrightarrow{N \rightarrow \infty} 0$$



Kai N artėja prie begalybės, σ_Y artėja prie nulio. Tai reiškia kad skirstinys siaurėja ir virsta δ -Dirako funkcija lokalizuota ties $y = \langle x \rangle$:

$$P_{Y_N}(y) \xrightarrow{N \rightarrow \infty} \delta(y - \langle x \rangle)$$

Taigi didinant matavimų skaičių N galima patikimai nustatyti eksperimente matuojamo dydžio vidurkį $\langle x \rangle$.