

Termodynamika i Fizyka Statystyczna R

Przewodnik dla ostrożnych

Wersja: 2 marca 2026

1 Wstęp: Z rozpaczy ku fizyce statystycznej

„Ludwig Boltzmann, who spent much of his life studying statistical mechanics, died in 1906, by his own hand. Paul Ehrenfest, carrying on the work, died similarly in 1933. Now it is our turn to study statistical mechanics. Perhaps it will be wise to approach the subject cautiously.”

— David Goodstein, *States of Matter*

Powyższy cytat, choć podszyty czarnym humorem, doskonale oddaje ciężar gatunkowy dziedziny, którą będziemy się zajmować. Dlaczego mechanika statystyczna doprowadzała wybitne umysły do takiej frustracji? Aby to zrozumieć, musimy spojrzeć na ograniczenia tradycyjnego opisu fizycznego.

Mechanika klasyczna i kwantowa odnoszą ogromne, wręcz spektakularne sukcesy w odkrywaniu i analizie praw rządzących ruchem pojedynczych ciał czy cząstek. Potrafimy z niezwykłą precyzją wyliczyć trajektorię rzuconego kamienia, ruch planet, czy stany energetyczne atomu wodoru. Jednak wraz ze wzrostem liczby cząstek w rozpatrywanym układzie, problem drastycznie się komplikuje, i to na tyle, że klasyczne metody analityczne po prostu zawodzą.

Warto przypomnieć, że już słynny problem trzech ciał oddziałujących grawitacyjnie nie posiada ogólnego rozwiązania analitycznego w postaci zwartej (tzw. *closed-form solution*). Owszem, w 1912 roku fiński matematyk Karl F. Sundman zdołał udowodnić istnienie ścisłego rozwiązania w postaci zbieżnego szeregu nieskończonego. Z praktycznego i fizycznego punktu widzenia jest ono jednak całkowicie bezużyteczne – zbiega ono tak wolno, że aby uzyskać sensowną precyzję dla krótkiego odcinka czasu, należałoby zsumować miliony wyrazów tego szeregu. Co więcej, pod koniec XIX wieku Henri Poincaré udowodnił, że układ trzech ciał jest fundamentalnie chaotyczny. Oznacza to ekstremalną wrażliwość na najdrobniejsze zmiany warunków początkowych. Trajektorie ciał splątują się ze sobą w skomplikowany, niepowtarzalny i niemożliwy do długoterminowego przewidzenia sposób.

A mówimy tu przecież zaledwie o trzech oddziałujących obiektach! A co, jeśli tych ciał jest nie trzy, lecz tryliony? Czy pozostaje nam tylko załamać ręce i pogrążyć się w rozpaczy?

1.1 Bariera obliczeniowa, czyli dlaczego komputery nas nie ocalą

Pozostają nam rzecz jasna jeszcze metody numeryczne. Naiwne podejście podpowiada: weźmy po prostu równania Newtona, wsadźmy je do potężnego komputera i pozwólmy maszynie rozwiązać je krok po kroku. Niestety, nawet one nie pozwalają na opis układów o rozmiarach makroskopowych.

Warto w tym miejscu spojrzeć prawdzie w oczy: w tej walce najpotężniejsze komputery świata są równie bezradne co kartka i ołówek. Dziś dysponujemy superkomputerami klasy eksa-skalowej (wykonującymi ponad 10^{18} operacji zmiennoprzecinkowych na sekundę, czyli eksaflopy). Obecnymi rekordzistami są giganty takie jak **Frontier** (znajdujący się w Oak Ridge National Laboratory w USA) czy **Aurora** (Argonne National Laboratory). Ich kolosalna moc obliczeniowa

nie opiera się na klasycznych procesorach, lecz na zaprzęgnięciu do pracy dziesiątek tysięcy potężnych akceleratorów graficznych (GPU) liczących wszystko równolegle.

Co to oznacza w praktyce dla fizyka? Największe z możliwych obecnie symulacji dynamiki molekularnej (MD) na takich maszynach potrafią śledzić ruch rzędu biliona (10^{12}) atomów. Brzmi to dumnie, ale czy rozwiązuje nasz problem? Absolutnie nie. Bilion atomów to wciąż zaledwie mikroskopijny pyłek, znikomy ułamek mikromola materii. Co gorsza, średni czas między oddziaływaniem cząstek, np. w gazie, jest rzędu nanosekund, a żeby symulacja numeryczna była w ogóle stabilna, krok czasowy musi być ułamkiem femtosekundy (10^{-15} s). To oznacza, że po tygodniach ciągłego działania największych klastrów obliczeniowych świata na pełnych obrotach, potrafimy prześledzić ewolucję takiego układu zaledwie przez ułamki mikrosekundy. Wynik jest przytłaczająco mizerny – ewolucja zaledwie *jednego mola* gazu (10^{24} cząstek) przez *jedną sekundę* wymagałaby mocy obliczeniowej o jakieś 25 rzędów wielkości większej niż ta, którą dysponuje dzisiaj cała ludzkość.

1.2 Zbój Gębon i demon Laplace’a

Jeżeli jednak przez moment wyobrazilibyśmy sobie, że taki nadnaturalny komputer istnieje, potrafi wyliczyć trajektorie cząstki po cząstce w szklance wody i wyrzuca z siebie w mgnieniu oka biliardy liczb opisujących konfiguracje układu w każdej kolejnej chwili czasu... moglibyśmy ze zdumieniem stwierdzić, że te potoki liczb **nikogo tak naprawdę nie interesują**.

Nikt nie zadałby pytania: „gdzie dokładnie znajduje się cząsteczka nr 156 893 211 w czwartej mikrosekundzie?”. Pytania, które byśmy zadawali, byłyby zupełnie innej natury: jaka jest gęstość cząstek tu czy tam? Gdzie średnio wędrują cząstki? Jakie wywierają ciśnienie? Czyli nie tyle interesowałoby nas zachowanie pojedynczych indywiduów, ile pewne własności kolektywne, uśrednione, makroskopowe.

Dramat zbója Gębona

Sytuacja ta do złudzenia przypomina los zbójcy Gębona z opowiadania Stanisława Lema „*Jak Trurl i Klapaucjusz demona drugiego rodzaju stworzyli, aby zbójcę Gębona pokonać*”. Gębon, łaknący potęgi płynącej z Wszelkiej Informacji, został ukarany przez genialnych konstruktorów maszyną (Demonem Drugiego Rodzaju), która z chaotycznego ruchu cieplnego atomów powietrza odczytywała absolutnie wszystkie prawdziwe fakty. Zbójca momentalnie został pogrzebany pod górami papierowych taśm zawierających nieskończoną ilość bezużytecznych dla niego danych:

„(...) wydawało się zbójowi, że zaraz już dowie się rzeczy niesłychanych, takich, które mu oczy na Istotę Bytu otworzą, więc wczytywał się we wszystko, co leciało spod brylancika (...) a tymczasem czytał o tym, jak się wiję arlebardzkie wiją i że córka króla Petrycego z Labaudii zwała się Garbunda, i co jadł na drugie śniadanie Fryderyk II, król bładawców, nim wojnę wypowiedział Gwendolinom, i ile powłok elektronowych liczyłby sobie atom termionolium, gdyby taki pierwiastek był możliwy, i jakie są wymiary dziurki tylnej małego ptaszka, zwanego kurkucielem, którego na swych rozamforach malują Marłajowie Wabędzcy, jak również o trzech smakach poliwonnych szlamu oceanicznego na Wodocji Przyzroczonej, i o kwiatku Łubuduku, który myśliwych staromalfandzkich wali siarczyście na odlew, świtem wzruszony, i jak wyprowadzić wzór na dostawę kąta podstawy wieloboku, ikoseadrem zwanego, i kto był jubilerem Fafucjusza, rzeźnika mańkuta Buwantów, i ile pism filatelistycznych będzie wychodziło w roku siedemdziesięcioletnim na Morkonaucji, i gdzie znajduje się trupek Cybrycji Kraśnopiętej, którą gwoździem przebił po pijanemu niejaki Malkonder, i czym się różni Maciąg od Naciągu, a także kto ma najmniejszą w Kosmosie pielownicę wzdłużną i dlaczego pchły smoczkotyłkie mchu jeść nie chcą, i na czym polega gra zwana Balansyer

Zadni Ściągany, i ile było ziarenek lwichwostu w tej kupce, co ją Abrukwian Po-listny nogą trącił, kiedy się pośliznął na ósmym kilometrze szosy albacjerskiej w Dolinie Wzduchów Szedziwych.”

Dokładnie to samo spotkałoby nas, gdybyśmy dysponowali komputerem śledzącym ruch i parametry każdego pojedynczego atomu. Utonęlibyśmy w oceanie absolutnie prawdziwych, lecz fizycznie bezużytecznych danych.

Innym krytycznym elementem w tych rozważaniach jest chaos deterministyczny. Najpierw musielibyśmy do naszego wysnionego komputera załadować gigantyczny plik z warunkami początkowymi wszystkich atomów. Musiałyby to być warunki znane z absolutnie nieskończoną precyzją. Rzeczywistość jest bowiem bezlitosna: w układach wielociałowych, już po czasach rzędu kilku czasów między zderzeniami, dwie trajektorie różniące się w warunkach początkowych choćby o niewyobrażalnie mały ułamek, pójdą dwiema zupełnie innymi drogami (tzw. efekt motyla w skali mikro).

Wszystko to prowadzi nas do nieuniknionego wniosku: **nie można po prostu „wyliczyć świata”**. Pełna informacja o mikro stanie układu w ogólności nie jest dla nas dostępna – to jest to, o czym mówiliśmy na samym początku wykładu. Taka absolutna wiedza deterministyczna jest dostępna wyłącznie dla hipotetycznych bytów, takich jak słynny demon pojawiający się w Esejach Laplace’a z roku 1814: „Wyobraźmy sobie istotę, która w danej chwili czasu znalazłaby wszystkie siły działające w świecie i położenia wszystkich jego składników. Jeśli istota ta byłaby wystarczająco potężna, by poddać te dane analizie, mogłaby zawrzeć w jednym równaniu ruch największych ciał Wszechświata i najmniejszego atomu, dla takiej istoty nic nie byłoby niepewne a przyszłość tak jak i przeszłość rozpościerałaby się przed jej oczami.” Daleko nam jednak do demona Laplace’a. Mamy drastycznie ograniczone możliwości pomiarowe oraz obliczeniowe i musimy sobie jakoś, skromnie, po ludzku z tym radzić.

1.3 Paradygmat statystyczny

Wszystko to sprawia, że musimy definitywnie zrezygnować z podejścia indywidualistycznego na rzecz opisu uśrednionego, statystycznego, wyrażonego w języku rachunku prawdopodobieństwa. Przyznajemy się do fundamentalnej niewiedzy: nie będziemy wiedzieli, dokąd lecą poszczególne atomy. Zadowolimy się wiedzą o tym, ile ich *średnio* poleciało tu i tam, oraz jakie są prawdopodobieństwa wystąpienia różnorodnych efektów.

Paradoksalnie, to zrzeczenie się precyzji na poziomie mikro pozwala nam na niesamowicie precyzyjny opis na poziomie makro. W ten sposób będziemy w stanie zrozumieć i przewidzieć najróżniejsze zjawiska – od tak prozaicznych (że kiedy ściskamy balon z gazem, to on się ogrzewa, albo że jeżeli ogrzewamy jakiś materiał, to on się rozszerza), aż po zjawiska kwantowe i egzotyczne, jak kondensacja Bosego-Einsteina (BEC).

Co ciekawe, takie statystyczne podejście zrodzone z fizyki może być (i z wielkim powodzeniem jest!) stosowane nie tylko do opisu martwej materii w laboratoriach. Odnajdujemy je wszędzie tam, gdzie duża liczba oddziałujących ze sobą jednostek tworzy zjawiska kolektywne.

- **W psychologii społecznej i socjologii** funkcjonuje zasada: „Nie wiem, jak zachowa się konkretna jednostka – ale znam psychologię tłumu”. Reakcja jednego człowieka na bodziec może być nieprzewidywalna, ale zachowanie stutysięcznego tłumu na stadionie poddaje się już dość rygorystycznym prawom statystyki.
- **W fizyce ruchu drogowego** mamy do czynienia z podobnym ewenementem. Chociaż nie potrafimy przewidzieć zachowania i ruchu poszczególnych kierowców na autostradzie (bo zawsze trafią się narwańcy, piraci drogowi, niedzielni „wolniacy” i osoby zamyślane), to jednak za pomocą zaledwie kilku założeń o ich *średnim zachowaniu* jesteśmy w stanie niezwykle skutecznie symulować ruch drogowy całych metropolii. Potrafimy przewidywać

powstawanie korków, czy też modelować powstawanie fal na przemian zagęszczonego i rozrzedzonego ruchu drogowego (tzw. *stop-and-go waves*). Zresztą samo powstawanie i propagacja korków to bardzo ciekawy przykład zjawiska emergentnego (kolektywnego), który pojawia się dopiero wtedy, kiedy mamy do czynienia z wieloma pojazdami – odpowiednikami „atomów” na drodze – i który jest absolutnie niemożliwy do przewidzenia na podstawie znajomości mechaniki i ruchu pojedynczych samochodów. No bo jak, na litość boską, wpaść na to analizując ruch samotnego samochodu na pustej drodze, że w gęstym ruchu ulicznym tzw. korki fantomowe (korki powstające bez wyraźnej przyczyny, np. wypadku) **propagują się zawsze do tyłu z prędkością około 15 km/h**? Co niesamowite, prędkość propagacji tej fali zagęszczenia w tył jest jak najwyraźniej stała i uniwersalna dla prawie wszystkich szos na świecie, niezależnie od tego, czy to amerykańska autostrada, czy niemiecki Autobahn!

- Zupełnie **podobnie zachowują się piesi** w tłumie. Zjawiska ucieczki, problematyka paniki, zatorów w wąskich przejściach ewakuacyjnych – to wszystko fizyka statystyczna w najczystszej postaci. Piesi traktowani są jako samobieżne cząstki oddziałujące na siebie siłami odpychania społecznego, co pozwala inżynierom i architektom projektować bezpieczniejsze stadiony i stacje metra.

To właśnie siła statystyki. Nie znając szczegółów, poznajemy całościowy obraz. W kolejnej części przejdziemy do formalnego aparatu matematycznego, który nam to umożliwi – do rachunku prawdopodobieństwa.

2 Rachunek prawdopodobieństwa: Intuicja, która wodzi na manowce

2.1 Pojęcia podstawowe

Większość z nas uważa, że intuicyjnie rozumie, czym jest prawdopodobieństwo. Przecież to proste – jaka jest szansa na to, że w rzucie monetą wypadnie reszka? Odpowiadamy bez wahania: $1/2$. Ale co właściwie przez to rozumiemy?

Mamy tu do dyspozycji dwie główne możliwości interpretacyjne:

1. Chodzi o wynik pojedynczego rzutu (Podejście subiektywne / informacyjne)

Mówi się wtedy, że prawdopodobieństwo P opisuje po prostu **stan naszej wiedzy** na temat wyniku jeszcze niewykonanego doświadczenia. Takie ujęcie natychmiast naraża się na mocny zarzut o subiektywizm („każdy obserwator wie coś innego”). Odpiera się go jednak tłumacząc, że wcale nie chodzi o subiektywne widzimisię badacza, lecz o to, co obiektywnie *można wiedzieć*, przeprowadzając dane doświadczenie dokładnie według zadanego przepisu.

2. Chodzi o częstość wystąpienia orła w długiej serii rzutów (Podejście częstościowe)

Jeśli nie podoba nam się abstrakcyjny opis „stanu wiedzy”, możemy oprzeć się na empirii. Rzucaliśmy monetą wystarczająco dużo razy i tak po prostu wyszło. To podejście prowadzi nas do **definicji częstościowej**, która mówi, że prawdopodobieństwo to granica stosunku liczby zdarzeń sprzyjających (N_r) do całkowitej liczby prób (N):

$$p_r = \lim_{N \rightarrow \infty} \frac{N_r}{N}$$

Niestety, i ta ścieżka kryje dwa poważne zarzuty:

- Każda fizyczna seria prób w naturze jest w rzeczywistości **skończona**. Nigdy nie osiągniemy granicy w nieskończoności.
- Skąd w ogóle wiemy, że zdarzenia w tej serii są ze sobą faktycznie niezależne?

Co gorsza, gdy zapytamy, co mówi teoria na temat częstości występowania danego zdarzenia w długiej, ale wciąż skończonej serii doświadczeń, okaże się, że teoria ta opisuje to... znowu używając języka prawdopodobieństwa. Wpadamy tu w zgrabne, błędne koło.

Co ciekawe, w historii nauki nie brakowało śmiazków (lub ludzi skrajnie znudzonych), którzy postanowili sprawdzić podejście częstościowe empirycznie:

- **John Kerrich** – angielski matematyk, który na początku II wojny światowej wykładał w Kopenhadze. Aresztowany przez Niemców, spędził wojnę w obozie jenieckim na Jutlandii. Z nudów przeprowadzał eksperymenty probabilistyczne. Rzucił monetą 10 tysięcy razy i otrzymał 5067 orłów.
- **Karl Pearson** – słynny statystyk, miał jeszcze więcej czasu i samozaparcia. Wykonał 24 000 rzutów, z czego 12 012 zakończyło się wyrzuceniem orła.

O monetach warto wiedzieć jeszcze kilka rzeczy. O ile rzuconej monety nie da się łatwo oszukać (w przeciwieństwie do kości do gry, które można obciążyć), o tyle sytuacja zmienia się dramatycznie, jeśli monetą **zakręcimy jak bączkiem**. Polscy matematycy z Siedlec, Tomasz Gliszczyński i Włodzimierz Zawadowski (o których w 2002 roku pisał nawet brytyjski *The Times*), kazali swoim studentom zakręcić 250 razy nową belgijską monetą o nominale 1 Euro. Wynik? Król Belgii (awers) pokazał się 140 razy, a reszka tylko 110. Strona z twarzą króla była po prostu cięższa.

Niezmiernie ważna jest też sama technika rzutu. Persi Diaconis – niegdyś profesjonalny magik, a dziś wybitny profesor matematyki na Uniwersytecie Stanforda – udowodnił, że potrafi rzucić monetą w taki sposób, aby za każdym razem wylądowała na żądanej stronie. Udowodnił, że rzut monetą to nie magia ani czysta losowość, lecz zagadnienie z zakresu mechaniki klasycznej.

Budowanie intuicji: Eksperyment Alicji i Boba

Aby uniknąć pułapek intuicji, w mechanice statystycznej musimy zbudować precyzyjny, matematyczny model zjawisk losowych. Zanim jednak podamy rygorystyczne aksjomaty, prześledźmy proces ich powstawania na bogatszym przykładzie.

Wyobraźmy sobie następującą sytuację: Alicja dzwoni do Boba i podaje mu wyniki swoich rzutów dwiema kostkami do gry. Ścisłej mówiąc, podaje mu tylko *sumy* wyrzuconych oczek:

4, 7, 5, 2, 6, 11, 7, 9, 9, 6...

Na razie jest to mało ciekawy ciąg liczb. Ale po dłuższym powtarzaniu eksperymentu Bob zauważa pewną prawidłowość – dostajemy ciąg, w którym najczęściej wychodzi 6, 7 i 8, a najrzadziej 2 i 12. Tym z Was, którzy grali w Catana, ten fakt jest na pewno dobrze znany. Jak to opisać i sformalizować? Najpierw doprecyzujmy, co Boba w ogóle interesuje. Niech $\mathcal{A}$ oznacza **zbiór interesujących nas zdarzeń**. To mogą być proste zdarzenia, np.:

- S_2 (suma wynosi 2),
- S_6 (suma wynosi 6),
- S_9 (suma wynosi 9).

Ale mogą to być również zdarzenia złożone, traktowane jako operacje na zbiorach, np. suma logiczna $A = S_2 \cup S_6 \cup S_9$, albo zdarzenie $B =$ „cokolwiek, byle nie 2 i 12”. Możemy mieć też w tym zbiorze istotne przypadki skrajne:

- **zdarzenie niemożliwe** (np. suma oczek wynosi 13). Z definicji chcemy, by $P(\text{niemożliwego}) = 0$.
- **zdarzenie pewne** (np. suma oczek to liczba całkowita). Z definicji $P(\text{pewnego}) = 1$.

Teraz chcielibyśmy każdemu zdarzeniu ze zbioru $\mathcal{A}$ przypisać prawdopodobieństwo. Będzie

to funkcja przyporządkowująca zdarzeniom liczby rzeczywiste:

$$P : \mathcal{A} \ni S \rightarrow P(S) \in \mathbb{R}$$

Rozsądnie jest też przyjąć addytywność prawdopodobieństw: alternatywa zdarzeń to w języku zbiorów ich suma ($A \cup B$), więc chcemy, aby prawdopodobieństwo alternatywy zdarzeń wykluczających się było równe sumie ich prawdopodobieństw.

Zauważmy fundamentalną rzecz: każde zdarzenie (np. S_5) analizowane przez Boba jest tak naprawdę uwarunkowane niewidocznymi dla niego, mikroskopowymi stanami kostek. Wprowadźmy zbiór **wszystkich możliwych par liczb** określających pojedyncze wyniki na obu kostkach. Niech Ω oznacza zbiór tych par:

$$\Omega = \{(i, j) \mid i, j \in \{1, 2, \dots, 6\}\}$$

Elementy tego bazowego zbioru nazywamy **zdarzeniami elementarnymi**.

Każde makroskopowe zdarzenie S to po prostu podzbiór zbioru Ω . Zauważmy jednak, że nie wszystkie zdarzenia elementarne są bezpośrednio identyfikowalne dla Boba! Np. konkretny rzut $(2, 3)$ należy do zdarzenia makroskopowego S_5 , ponieważ:

$$S_5 = \{(1, 4), (2, 3), (3, 2), (4, 1)\}$$

Gdybyśmy założyli (co dla idealnych kostek jest rozsądne), że wszystkie z 36 możliwych zdarzeń elementarnych (i, j) są równoprawdopodobne, to z zasady addytywności dostalibyśmy, że prawdopodobieństwo każdego z nich to $1/36$. Stąd łatwo wyznaczyć prawdopodobieństwa interesujących Boba sum:

$$\begin{aligned} P(S_5) &= \frac{4}{36} && \text{bo sprzyjają mu pary: } (1, 4), (2, 3), (3, 2), (4, 1) \\ P(S_7) &= \frac{6}{36} && \text{bo sprzyjają mu pary: } (1, 6), (6, 1), (2, 5), (5, 2), (4, 3), (3, 4) \end{aligned}$$

2.2 Formalizm matematyczny i aksjomatyka Kołmogorowa

Ogólny schemat tego, co intuicyjnie zrobiliśmy przed chwilą, stanowi fundament formalnego rachunku prawdopodobieństwa. Wprowadzamy w nim trzy filary, tworzące tzw. przestrzeń probabilistyczną $(\Omega, \mathcal{A}, P)$:

1. Przestrzeń próbek (Zbiór zdarzeń elementarnych) Ω

W każdym modelu występuje zbiór Ω , zawierający wszystkie możliwe wyniki danego eksperymentu losowego na poziomie najbardziej podstawowym. Dla dwóch rzutów monetą byłoby to $\Omega = \{(O, O), (O, R), (R, R), (R, O)\}$.

2. Zbiór zdarzeń losowych (σ -algebra) $\mathcal{A}$

Zdarzenia losowe (to, co nas interesuje) to podzbiory zbioru Ω . Wyodrębniamy zbiór $\mathcal{A}$, który zawiera te podzbiory. Aby matematyka działała poprawnie, od rodziny zbiorów $\mathcal{A}$ (zwanej σ -algebrą) wymagamy spełnienia trzech warunków:

1. $\Omega \in \mathcal{A}, \emptyset \in \mathcal{A}$ (zdarzenie pewne i niemożliwe zawsze należą do naszych rozważań).
2. $A, B \in \mathcal{A} \implies A \cup B, A \cap B, A \setminus B \in \mathcal{A}$ (zbiór jest zamknięty na podstawowe operacje logiczne: sumę, iloczyn, różnicę).
3. $A_n \in \mathcal{A}, n = 1, 2, \dots \implies \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$ (nieskończona suma zdarzeń również jest zdarzeniem).

Ten ostatni, nieco tajemniczy punkt (który dzięki prawom De Morgana gwarantuje również zamkniętość na nieskończone przecięcia zbiorów) zapewnia nam *ciągłość prawdopodobieństwa*. Możemy na to patrzeć na dwa równoważne sposoby:

- **Ciągłość z dołu:** dla rosnącego ciągu zdarzeń ($A_1 \subset A_2 \subset \dots$) chcemy, by $P(A_n) \rightarrow P(A)$, gdzie $A = \bigcup_{n=1}^{\infty} A_n$.
- **Ciągłość z góry:** dla malejącego ciągu zdarzeń ($A_1 \supset A_2 \supset \dots$) chcemy, by $P(A_n) \rightarrow P(A)$, gdzie $A = \bigcap_{n=1}^{\infty} A_n$.

Dlatego właśnie nieskończone sumy (jak i iloczyny) zdarzeń muszą należeć do naszej σ -algebry, aby granice te miały matematyczny sens.

3. Miara prawdopodobieństwa P

Prawdopodobieństwo to funkcja przypisująca poszczególnym zdarzeniom z algebry $\mathcal{A}$ liczby rzeczywiste, spełniająca aksjomaty Kołmogorowa:

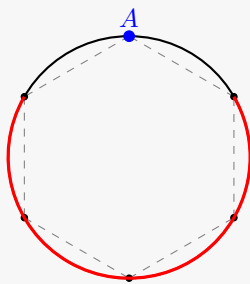
1. Nieujemność: Dla każdego $A \in \mathcal{A}$, zachodzi $P(A) \geq 0$.
2. Unormowanie: $P(\Omega) = 1$ oraz $P(\emptyset) = 0$.
3. Przeliczalna addytywność: Jeśli ciąg zdarzeń $A_1, A_2, \dots \in \mathcal{A}$ składa się ze zdarzeń parami rozłącznych ($A_i \cap A_j = \emptyset$ dla $i \neq j$), to:

$$P\left(\bigcup_n A_n\right) = \sum_n P(A_n)$$

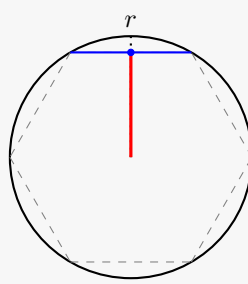
Ten ostatni aksjomat jest niezwykle ważny w praktyce. Gwarantuje on na przykład, że jeżeli podzielimy całą naszą przestrzeń zdarzeń elementarnych Ω na nieskończenie wiele (ale przeliczalnie wiele, np. ponumerowanych liczbami naturalnymi) rozłącznych kawałeczków A_n , to ich suma to po prostu Ω , a suma ich prawdopodobieństw sumuje się do jedności: $\sum_{n=1}^{\infty} P(A_n) = P(\Omega) = 1$. Bez tego nie moglibyśmy sensownie opisywać rozkładów na nieskończonych, ale dyskretnych zbiorach (np. gdy pytamy o liczbę atomów, które uległy rozpadowi – może to być $0, 1, 2, \dots$ w nieskończoność).

Przerywnik: Paradoks Bertranda

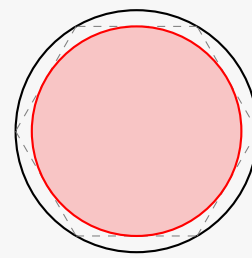
Aby pokazać, że matematyczny rygor nie jest sztuką dla sztuki, rozważmy klasyczny problem z 1889 roku. Losowo wybieramy cięciwę okręgu o promieniu r . Jakie jest prawdopodobieństwo, że jej długość jest **większa niż promień okręgu**?



Metoda 1: $p = \frac{4}{6} = \frac{2}{3}$



Metoda 2: $p = \frac{\sqrt{3}}{2}$



Metoda 3: $p = \frac{3}{4}$

Rysunek: Trzy różne rozwiązania paradoksu Bertranda (dla warunku, że długość cięciwy jest większa od promienia okręgu r , co geometrycznie odpowiada bokowi wpisanego sześciokąta). Kolorem czerwonym oznaczono zbiór zdarzeń sprzyjających dla każdego schematu losowania.

- **Odpowiedź nr 1:** Wybieramy losowo punkt na okręgu i z niego prowadzimy cięciwę do drugiego, losowo wybranego punktu na obwodzie. Rozważając sześciokąt foremny wpisany w okrąg (którego bok ma długość dokładnie r), zauważamy, że drugi punkt musi spaść na 4 z 6 łuków utworzonych przez wierzchołki sześciokąta. Prawdopodobieństwo wynosi $p = 4/6 = 2/3$.

- **Odpowiedź nr 2:** Załóżmy, że losujemy odległość środka cięciwy od środka okręgu – wartość z przedziału $(0, r)$. Aby cięciwa była dłuższa niż bok naszego sześciokąta (r), jej odległość od środka okręgu musi być mniejsza od wysokości trójkąta równobocznego budującego ten sześciokąt, czyli $r\sqrt{3}/2$. Ponieważ wybór jest jednorodny na odcinku r , daje nam to $p = \frac{r\sqrt{3}/2}{r} = \sqrt{3}/2 \approx 0.866$.
- **Odpowiedź nr 3:** Wybieramy losowo punkt wewnątrz całego okręgu, który staje się środkiem naszej cięciwy. Aby cięciwa była dłuższa od r , wybrany punkt musi leżeć wewnątrz mniejszego okręgu o promieniu $r\sqrt{3}/2$. Stosunek pól tych dwóch okręgów wynosi $3/4 = 0.75$.

Morał: Która odpowiedź jest poprawna? Wszystkie! Paradoks obnaża brutalną prawdę: zadanie probabilistyczne nie ma najmniejszego sensu bez jasnego sprecyzowania mechanizmu (schematu) losowania, a zatem wybory zbioru zdarzeń elementarnych. Pytanie „które zdarzenia są równoprawdopodobne?” jest ważniejsze niż same obliczenia.

2.3 Zdarzenia niezależne i prawdopodobieństwo warunkowe

Nasze przewidywania tego, co się wydarzy, zależą krytycznie od posiadanej informacji. I tak np. prawdopodobieństwo tego, że wracając wieczorem do domu zostaniemy pobici, mocno rośnie, gdy dowiemy się, że tego właśnie dnia Legia będzie grała mecz.

Jako mniej drastyczny przykład weźmy rzut dwiema monetami z przestrzenią próbek $\Omega = \{(O, O), (O, R), (R, O), (R, R)\}$. Niech zdarzenie $A = \{(O, O)\}$ (wypadły dwa orły), a zdarzenie B to „wypadła przynajmniej jedna reszka” czyli $\{(O, R), (R, O), (R, R)\}$. Jeśli wiemy, że zaszło zdarzenie B , to zdarzenie A staje się absolutnie niemożliwe. Stąd prawdopodobieństwo zdarzenia A pod warunkiem zajścia B wynosi zero.

Prowadzi to do definicji **prawdopodobieństwa warunkowego**:

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

Inny klasyczny przykład: w pokoju jest 100 osób. 20 z nich ma brodę, 25 to osoby czarnoskóre, a 5 to czarnoskóre osoby z brodą. Jakie jest prawdopodobieństwo, że wylosowana osoba ma brodę, *jeśli wiemy*, że jest czarnoskóra? $P(\text{brodaty}|\text{czarnoskóry}) = 5/25 = 1/5$.

Jeżeli jednak zdobyta informacja nie zmienia absolutnie nic, mówimy o **zdarzeniach niezależnych**. Zdarzenia A i B są niezależne, jeśli zajście B nie wpływa na szansę zajścia A :

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

Podstawiając to do definicji prawdopodobieństwa warunkowego, otrzymujemy kluczowy wzór na iloczyn zdarzeń niezależnych:

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

Np. szansa na wyrzucenie dwóch szóstek w dwóch rzutach niezależnymi kostkami to po prostu $1/6 \cdot 1/6 = 1/36$.

Brak zrozumienia pojęcia niezależności to źródło najpopularniejszych przesądów:

- „Po dwóch dziewczynkach teraz musi być chłopiec, żeby statystyka się wyrównała.” (Natura nie ma pamięci).
- Naiwna strategia gry w ruletkę: „Poczekam na długą serię czerwonych i wtedy na pewno wypadnie czarne.” (Kulka też nie ma pamięci!).
- Stary dowcip o matematyku, który po fali zamachów bombowych zaczął wносить na pokład samolotu własną bombę. Wyliczył bowiem, że szansa, iż na pokładzie znajdzie się jedna bomba, wynosi 1 na 10 tysięcy, ale szansa na to, że w jednym samolocie znajdą się *dwie*

bomby, wynosi już 1 na 100 milionów. (Błąd leży w niezależności – jego własna bomba nie zmniejsza prawdopodobieństwa cudzej bomby!).

Prawdopodobieństwo na sali sądowej

Z założeniem o niezależności trzeba niezmiernie uważać! Oto dwa mrozące krew w żyłach, autentyczne przykłady nadużycia rachunku prawdopodobieństwa przez oskarżycieli, które zrujnowały ludzkie życia.

Przypadek Josepha Trujillo (1946)

W USA zastrzelono oficera policji. O zabójstwo oskarżono Josepha Trujillo. Na jego ubraniu znaleziono 11 włókienek materiału odpowiadających materiałowi z ubrania ofiary. Prokurator stwierdził arbitralnie, że prawdopodobieństwo przypadkowej zgodności jednego włókna to 0.1. Mnożąc to przez siebie (zakładając niezależność!), przekonywał ławę przysięgłych, że dla 11 włókien szansa pomyłki to $0.1^{11} = 0.00000000001$. Co więcej, w pistolecie oskarżonego i w broni zbrodni znaleziono 5 cech wspólnych, m.in.:

- pistolety nieautomatyczne (prokurator założył $p = 0.5$),
- kaliber .38 ($p = 0.2$),
- lekko skrzywiona lufa ($p = 0.2$).

Pomnożenie tych „prawdopodobieństw” dało znów absurdalnie małą liczbę. Błąd polegał na tym, że te zdarzenia **nie były niezależne**. Włókna z tego samego płaszcza są ze sobą silnie skorelowane (jak spadnie jedno, to i drugie), a rewolwery kalibru .38 rzadziej bywały bronią automatyczną niż małe kalibry, więc cechy broni też były zależne. Niestety, Trujillo został skazany i stracony.

Sprawa małżeństwa Collinsów (1964)

W Kalifornii starsza kobieta została uduszona w alejce. Świadek zeznał, że z miejsca zbrodni wybiegła blondynka z włosami spiętymi w kucyki, wsiadła do żółtego samochodu prowadzonego przez czarnoskórego mężczyznę z wąsami i brodą. Krótco potem policja zatrzymała Janet i Malcolma Collinsów, którzy odpowiadali rysopisowi i posiadali żółtego Forda Lincolna. Prokurator, zachłyśnięty pseudomatematyką, zaprezentował ławie przysięgłych następujące prawdopodobieństwa:

- blondynka: $p = 1/3$
- kucyki: $p = 1/10$
- małżeństwo o różnych kolorach skóry: $p = 1/1000$
- żółty samochód: $p = 1/10$
- mężczyzna z wąsami: $p = 1/4$
- czarnoskóry z brodą: $p = 1/10$

Mnożąc te liczby, oskarżyciel otrzymał wynik 1 do 12000000, argumentując, że szansa, by winny był ktoś inny niż Collinsowie, jest bliska zeru. Ława przysięgłych skazała ich na wyroki wieloletniego więzienia (w pierwotnych doniesieniach groziła im kara śmierci).

Kilka lat później Sąd Najwyższy Kalifornii w głośnym wyroku *People v. Collins* zmasakrował te argumenty i anulował wyrok. Dlaczego? Po pierwsze, liczby były całkowicie wyssane z palca. Po drugie i najważniejsze – **cechy te absolutnie nie były niezależne!** Fakt, że mężczyzna nosi brodę, jest silnie skorelowany z posiadaniem przez niego wąsów. Fakt, że obserwujemy osobę o ciemnej skórze z brodą, w tamtym czasie nie był niezależny od faktu, że

jest to osoba o określonym statusie demograficznym. Co więcej, eksperci matematyczni powołani przez obronę pokazali, że stosując nawet te wadliwe wartości prokuratora do populacji kilku milionów mieszkańców aglomeracji Los Angeles, istniało około 41% szans, że w mieście znajduje się *jeszcze co najmniej jedna taka para*. Skazanie na podstawie samej statystyki było matematyczną zbrodnią.

3 Zmienne losowe i ich rozkłady

3.1 Zmienna losowa: pomost między abstrakcją a liczbą

Zanim przejdziemy dalej, zróbmy krótkie przypomnienie. Nasza matematyczna konstrukcja prawdopodobieństwa opiera się na trzech filarach: przestrzeni próbek (zbiór zdarzeń elementarnych) Ω , zbiorze zdarzeń losowych $\mathcal{A}$ i mierze prawdopodobieństwa P .

Jednak w fizyce rzadko interesuje nas to, czy moneta upadła orłem, czy reszką, albo jak dokładnie ułożyły się włókna w rzucie kością. Chcemy operować na konkretnych wielkościach fizycznych: energiach, prędkościach, położeniach. Potrzebujemy narzędzia, które przetłumaczy abstrakcyjne zdarzenia na język liczb rzeczywistych. Tym narzędziem jest **zmienna losowa**.

Zmienna losowa to funkcja określona na zbiorze zdarzeń elementarnych, przyjmująca wartości rzeczywiste:

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

Przykładowo: zdarzeniu „wypadła reszka” przyporządkowujemy liczbę 1, a „wypadł orzeł” liczbę 0. Albo inaczej: wylosowanej w loterii osobie (zdarzenie elementarne to konkretny człowiek) przyporządkowujemy jej wzrost w centymetrach.

Uwaga! Ważny wymóg formalny

Zmienna losowa nie może być zupełnie dowolną funkcją. Aby cała nasza misterna maszyna probabilistyczna działała, będziemy od niej wymagali, aby dla każdego przedziału $[a, b] \subset \mathbb{R}$ zbiór postaci:

$$A_{ab} = \{\omega \in \Omega : a \leq X(\omega) \leq b\}$$

należał do naszej algebry zdarzeń $\mathcal{A}$. Krótko mówiąc: przeciwobrazy odcinków w $\mathbb{R}$ muszą być elementami zbioru zdarzeń losowych.

Dlaczego to takie ważne? Ponieważ prawdopodobieństwo P potrafimy liczyć *tylko* dla elementów ze zbioru $\mathcal{A}$! Jeśli jakiś przeciwobraz nie trafi do $\mathcal{A}$, nie będziemy w stanie odpowiedzieć na pytanie „jakie jest prawdopodobieństwo, że zmienna X przyjmie wartość z tego przedziału?”.

Spójrzmy na **przykład patologiczny**: Niech nasza przestrzeń to $\Omega = \{1, 2, 3\}$, zaś algebra (bardzo uboga) to $\mathcal{A} = \{\{1\}, \{2, 3\}, \emptyset, \Omega\}$. Mamy też prawdopodobieństwa: $P(\{1\}) = 1/6$, $P(\{2, 3\}) = 5/6$. (Pamiętajmy, że zbiór pusty zawsze należy do $\mathcal{A}$, $P(\emptyset) = 0$).

Zdefiniujmy teraz funkcję $X : \Omega \rightarrow \mathbb{R}$ poprzez:

$$X(1) = 2, \quad X(2) = 3, \quad X(3) = 4.$$

Czy to jest zmienna losowa? **Nie!** Sprawdźmy przeciwobraz przedziału $(3.5, 4.5]$. Szukamy takich ω , dla których $3.5 < X(\omega) \leq 4.5$. Jedyńm takim argumentem jest $\omega = 3$. Zatem przeciwobraz to zbiór $\{3\}$. Niestety, zbioru $\{3\}$ nie ma w naszej algebrze $\mathcal{A}$! Nie potrafimy odpowiedzieć na pytanie, z jakim prawdopodobieństwem zmienna przyjmuje wartość 4.

Widząc powyższy problem patologiczny, można by zadać bardzo logiczne pytanie: „Czy to nie oznacza, że zbiór zdarzeń losowych $\mathcal{A}$ musi po prostu zawierać absolutnie wszystkie możliwe podzbiory Ω , żebyśmy mieli spokój?”.

Otóż **nie!** Ten formalny wymóg wcale nie zmusza nas do powiększania algebry $\mathcal{A}$. Zmusza nas do czegoś zupełnie innego: do **dopasowania zmiennej losowej do dostępnej informacji**

(w żargonie matematycznym mówimy, że zmienna losowa musi być *mierzalna* względem danej σ -algebry).

Wróćmy do naszego patologicznego przykładu. Zbiór $\mathcal{A} = \{\{1\}, \{2, 3\}, \emptyset, \Omega\}$ reprezentuje sytuację, w której nasz przyrząd pomiarowy potrafi odróżnić wynik 1 od reszty, ale **nie potrafi** rozróżnić 2 od 3 (wyniki te zawsze występują jako nierozdzielny pakiet $\{2, 3\}$).

Jeżeli spróbujemy zdefiniować funkcję X , która daje inny wynik dla 2 (np. $X(2) = 3$) i inny dla 3 (np. $X(3) = 4$), to funkcja ta po prostu wymaga większej precyzji, niż oferuje nasza algebra. Dlatego *nie jest* zmienną losową.

Gdybyśmy jednak zdefiniowali inną funkcję Y , np. reprezentującą wygraną w grze, w której za wyrzucenie 1 dostajemy 100 zł, a za wyrzucenie 2 lub 3 płacimy 50 zł kary:

$$Y(1) = 100, \quad Y(2) = -50, \quad Y(3) = -50$$

To taka funkcja **jest pełnoprawną zmienną losową** w tej bardzo ubogiej algebrze! Sprawdźmy: przeciwobraz wartości -50 to zbiór $\{2, 3\}$, który bezpiecznie leży w naszym $\mathcal{A}$. Zmienna losowa musi po prostu „szanować” ograniczenia informacyjne naszego zbioru zdarzeń i nie może rozróżniać stanów, których nasza algebra nie widzi.

Ta kluczowa własność przeciwobrazów pozwala nam na zdefiniowanie fundamentalnego pojęcia – **dystrybuanty zmiennej losowej**:

$$F_X(x) = P(\{\omega \in \Omega : -\infty < X(\omega) \leq x\}) \equiv P(X \leq x)$$

Zapis $P(X \leq x)$ to oczywiście skrót myślowy. Oznacza on „prawdopodobieństwo tych wszystkich zdarzeń elementarnych ω , dla których wartość zmiennej losowej $X(\omega)$ jest mniejsza bądź równa x ”.

Jeżeli dystrybuanta jest funkcją absolutnie ciągłą (i tym samym różniczkowalną prawie wszędzie), pozwala ona łatwo wyznaczyć tzw. **gęstość prawdopodobieństwa**, która jest po prostu jej pochodną:

$$\rho_X(x) = \frac{dF_X(x)}{dx}$$

3.2 Rozkład dyskretny i ciągły

Jeśli nasza zmienna losowa przyjmuje tylko pewien dyskretny ciąg wartości (np. $x_1, x_2, x_3 \dots$), to sytuacja staje się bardzo prosta. Aby w pełni opisać jej rozkład, wystarczy podać ciąg prawdopodobieństw:

$$p_k = P(X = x_k)$$

(Jeszcze raz przypominam: to zapis skrócony. Chodzi o prawdopodobieństwo tych wszystkich zdarzeń z Ω , dla których X przyjmuje konkretnie wartość x_k). Oczywiście prawdopodobieństwa te nie mogą być ujemne ($p_i \geq 0$) i muszą sumować się do jedności: $\sum_i p_i = 1$.

W języku analizy matematycznej możemy zapisać gęstość takiego dyskretnego rozkładu przy pomocy delty Diraca:

$$\rho_X(x) = \sum_i p_i \delta(x - x_i)$$

Z kolei dystrybuanta staje się funkcją „schodkową”, opisywaną za pomocą funkcji skokowej (thety) Heaviside’a (gdzie $\theta(x) = 0$ dla $x \leq 0$ oraz $\theta(x) = 1$ dla $x > 0$):

$$F_X(x) = \sum_i p_i \theta(x - x_i)$$

Przykład: Rzut kostką. Zmienna losowa to liczba wyrzuconych oczek ($X \in \{1, 2, 3, 4, 5, 6\}$). Gęstość rozkładu to sześć szpilek Diraca (każda o wadze $1/6$), a dystrybuanta to równe schodki pnące się w górę co każde oczko aż do wartości 1.

Kluczowe przykłady rozkładów dyskretnych spotykane w fizyce:

1. **Rozkład dwumianowy:** Mamy N niezależnych prób. W każdej próbie mamy tylko dwie możliwości (np. sukces/porażka, orzeł/reszka) z prawdopodobieństwami odpowiednio p oraz $1 - p$. Zmienną losową n jest łączna liczba odniesionych sukcesów (np. liczba wyrzucanych orłów w N rzutach). Prawdopodobieństwo wyrzucenia dokładnie n orłów to:

$$p_n = \binom{N}{n} p^n (1-p)^{N-n}$$

2. **Rozkład Poissona:**

$$p_n = e^{-\lambda} \frac{\lambda^n}{n!}$$

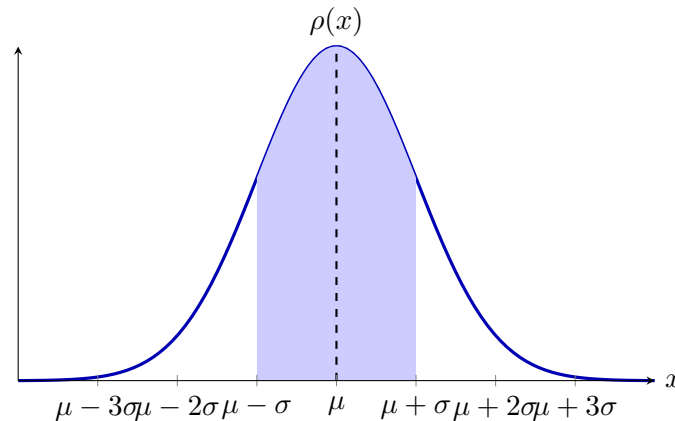
Jest to model opisujący prawdopodobieństwo wystąpienia dokładnie n rzadkich zdarzeń w określonym czasie t (np. liczba rozpadów promieniotwórczych w próbce w ciągu sekundy). Zakładamy, że kolejne wystąpienia są od siebie niezależne, a średnia spodziewana liczba zjawisk w tym czasie wynosi właśnie λ .

Zbiór wartości zmiennej losowej może być również **ciągły** – np. dzienny opad deszczu w Rabce, położenie cząstki w pudle, czy prędkość atomu w gazie.

Królem wszystkich rozkładów ciągłych jest oczywiście **Rozkład Gaussa (Normalny)**:

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

Gdzie μ to wartość oczekiwana, a σ to odchylenie standardowe determinujące szerokość „dzwonu”.



Rysunek 1: Gęstość ciągłego rozkładu normalnego (Gaussa). Zacieniowany obszar reprezentuje prawdopodobieństwo znalezienia wartości w przedziale jednego odchylenia standardowego od średniej.

3.3 Momenty rozkładu

Aby scharakteryzować rozkład w sposób syntetyczny, wprowadzamy pojęcie wartości oczekiwanej. Wartością oczekiwaną funkcji $f(X)$ zmiennej losowej X nazywamy liczbę:

$$\langle f(x) \rangle = \int_{-\infty}^{\infty} f(x) \rho(x) dx$$

(Dla rozkładów dyskretnych całka zamienia się oczywiście w sumę).

W szczególności interesują nas tzw. momenty rozkładu:

- $\langle x \rangle = \int x \rho(x) dx$ – średnia (wartość oczekiwana zmiennej), czyli **pierwszy moment**.

- $\langle x^k \rangle = \int x^k \rho(x) dx$ – **k -ty moment zwykły**.
- $\langle (x - \langle x \rangle)^k \rangle$ – **k -ty moment centralny** (mierzony względem średniej).
- $\langle (x - \langle x \rangle)^2 \rangle = \langle x^2 \rangle - \langle x \rangle^2 \equiv \sigma^2$ – **wariancja**, czyli drugi moment centralny. Opisuje, jak bardzo wyniki rozrzucane są wokół średniej.

Uwaga – nie każdy rozkład ma skończone momenty!

Istnieją rozkłady patologiczne, np. rozkład Cauchy’ego (opisujący m.in. kształt linii widmowych), dla którego średnia i wariancja są matematycznie nieokreślone (całki są rozbieżne), co sprawia sporo problemów nieostrożnym badaczom.

Paradoks Petersburski (Daniel Bernoulli, 1730)

Aby pokazać, że sama średnia (wartość oczekiwana) nie opisuje w pełni ludzkiego pojmowania szansy, rozważmy słynny paradoks, który spędzał sen z powiek osiemnastowiecznym matematykom.

Gramy przeciwko kasynu. Krupier rzuca monetą. Jeśli wypadnie orzeł – dostajemy 1 rubla, a gra się kończy. Jeśli reszka – gramy dalej. Jeśli orzeł wypadnie w drugim rzucie – dostajemy 2 ruble i kończymy. W trzecim – 4 ruble, w czwartym – 8, i tak dalej. Ogólnie: jeśli pierwszy orzeł wypadnie w n -tym rzucie, nasza wygrana wynosi 2^{n-1} rubli.

Zadajmy standardowe pytanie: jaka jest matematyczna wartość oczekiwana (średnia) naszej wygranej? Prawdopodobieństwo, że wygramy w 1. rzucie to $1/2$. Że w drugim (ciąg Reszka, Orzeł) to $1/4$. W n -tym rzucie to $1/2^n$.

Liczmy średnią wygraną:

$$\langle \text{wygrana} \rangle = \left(\frac{1}{2} \cdot 1 \right) + \left(\frac{1}{4} \cdot 2 \right) + \dots + \left(\frac{1}{2^n} \cdot 2^{n-1} \right) + \dots$$

$$\langle \text{wygrana} \rangle = \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \dots = \infty$$

Wartość oczekiwana naszej wygranej to nieskończoność! A teraz kluczowe pytanie z teorii gier: **Ile bylibyś w stanie zapłacić wpisowego, żeby wziąć udział w tej grze?**

Naiwnie, z punktu widzenia matematyki, gra jest dla nas opłacalna (spodziewamy się zysku), o ile wpisowe jest mniejsze niż oczekiwana wygrana. Skoro oczekiwana wygrana jest nieskończona, powinniśmy zastawić dom, sprzedać nerkę i wziąć gigantyczny kredyt, byle tylko móc zagrać. A jednak żaden rozsądny człowiek nie zapłaciłby za wejście do tej gry więcej niż 10, może 20 rubli.

Dlaczego? Pomocnicze pytanie: są dwie gry. W jednej masz prawdopodobieństwo jedna miliarda na wygranę miliarda złotych. W drugiej masz prawdopodobieństwo $1/2$ na wygranę 2 złotych. Wartość oczekiwana jest ta sama: około 1 złotego. W którą grę zagrasz chętniej? Ludzie instynktownie odrzucają nieskończenie małe prawdopodobieństwa potężnych wygranych, ponieważ szansa na odzyskanie wpisowego staje się nierealna. (Później doprowadziło to do powstania pojęcia „użyteczności krańcowej” w ekonomii).

Znajomość tego mechanizmu to podstawa **strategii d’Alemberta (martyngału)** w ruletce: stawiamy na czerwone. Przegrywamy? Podwajamy stawkę. Przegrywamy znowu? Znowu podwajamy. Teoretycznie, pierwsza wygrana pokryje nam wszystkie dotychczasowe straty z nawiązką. Problem w tym, że bardzo szybko docieramy do potężnych kwot (jak w Paradoksie Petersburskim), a kasyna mają limity maksymalnego zakładu, co brutalnie weryfikuje naiwne zastosowanie wartości oczekiwanej w praktyce.

3.4 Wielowymiarowe zmienne losowe

W fizyce rzadko mamy do czynienia z tylko jedną zmienną losową. Cząstka w gazie porusza się w trzech wymiarach, więc jej prędkość to wektor losowy składający się z trzech składowych. Załóżmy więc, że mamy więcej niż jedną zmienną losową: $X_1, X_2, X_3, \dots, X_n$, określone na tej samej przestrzeni zdarzeń Ω .

Ich **łącną dystrybuantę** definiujemy jako:

$$F_{X_1 \dots X_n}(x_1, \dots, x_n) = P(X_1 < x_1, \dots, X_n < x_n)$$

Pamiętajmy o interpretacji tego wyrażenia: jest to prawdopodobieństwo części wspólnej (przecięcia) zdarzeń. Innymi słowy, to prawdopodobieństwo, że zmienna X_1 przyjmie wartość mniejszą od x_1 , **i jednocześnie** zmienna X_2 będzie mniejsza od x_2 , itd. W przypadku dyskretnym przykładem może być rzut dwiema kostkami: interesuje nas wynik na pierwszej (X_1) i na drugiej (X_2). W przypadku ciągłym – np. łączny rozkład wysokości męża i żony w losowo wybranych małżeństwach w Polsce.

Analogicznie do przypadku jednowymiarowego, gęstość prawdopodobieństwa to po prostu wielokrotna pochodna mieszana dystrybuanty:

$$\rho_{\mathbf{X}}(x_1, \dots, x_n) = \frac{\partial^n F_{\mathbf{X}}(x_1, \dots, x_n)}{\partial x_1 \dots \partial x_n}$$

Często znając pełen rozkład, interesuje nas tylko jedna zmienna (np. prędkość tylko w osi x , ignorując resztę). Otrzymujemy wtedy tzw. **rozkłady marginalne (brzegowe)**, całkując (lub sumując) po wszystkich pozostałych zmiennych, których chcemy się „pozbyć”:

$$\rho_1(x_1) = \int_{-\infty}^{\infty} \rho(x_1, x_2) dx_2$$

Momenty wielowymiarowe definiujemy analogicznie, licząc wielokrotne całki: $\langle x_1 \rangle$, $\langle x_2 \rangle$, czy wreszcie momenty mieszane określające korelacje, np. $\langle x_1 x_2 \rangle$.

Korzystając z poznanego wcześniej pojęcia niezależności zdarzeń losowych, możemy zdefiniować **niezależność zmiennych losowych**. Mówimy, że zmienne X_1 i X_2 są niezależne, jeżeli dla dowolnych przedziałów $A_1, A_2 \subset \mathbb{R}$ ich przeciwobrazy są zdarzeniami niezależnymi. W praktyce (w języku gęstości prawdopodobieństwa) oznacza to, że łączna gęstość faktoryzuje się na iloczyn gęstości brzegowych:

$$X_1 \text{ i } X_2 \text{ są niezależne} \iff \rho(x_1, x_2) = \rho_1(x_1)\rho_2(x_2)$$

Dowód jest natychmiastowy:

$$\int_{A_1} dx_1 \int_{A_2} dx_2 \rho(x_1, x_2) = P(A_1 \cap A_2) = P(A_1)P(A_2) = \int_{A_1} dx_1 \rho_1(x_1) \int_{A_2} dx_2 \rho_2(x_2)$$

Dla zmiennych niezależnych wartość oczekiwana iloczynu jest iloczynem wartości oczekiwanych: $\langle x_1 x_2 \rangle = \langle x_1 \rangle \langle x_2 \rangle$ (ponieważ całka podwójna się po prostu rozkłada na dwie niezależne).

Korzystając z poznanego wcześniej pojęcia niezależności zdarzeń losowych, możemy zdefiniować **niezależność zmiennych losowych**. Z definicji mówimy, że zmienne X_1 i X_2 są niezależne, jeżeli dla dowolnych przedziałów $A_1, A_2 \subset \mathbb{R}$ prawdopodobieństwo tego, że obie zmienne jednocześnie trafią do swoich przedziałów, jest iloczynem prawdopodobieństw brzegowych:

$$P(X_1 \in A_1 \cap X_2 \in A_2) = P(X_1 \in A_1) \cdot P(X_2 \in A_2)$$

W praktyce (w języku gęstości prawdopodobieństwa) z definicji tej wynika niezwykle użyteczna własność: łączna gęstość faktoryzuje się na iloczyn gęstości brzegowych:

$$\rho(x_1, x_2) = \rho_1(x_1)\rho_2(x_2)$$

Dowód tego faktu jest bardzo pouczający. Wychodzimy od definicji niezależności (naszego założenia) i zapisujemy obie strony równania za pomocą całek:

- Lewa strona to z definicji całka z gęstości łącznej:

$$P(X_1 \in A_1 \cap X_2 \in A_2) = \int_{A_1} dx_1 \int_{A_2} dx_2 \rho(x_1, x_2)$$

- Prawa strona to iloczyn dwóch całek z gęstości brzegowych:

$$P(X_1 \in A_1) \cdot P(X_2 \in A_2) = \left(\int_{A_1} \rho_1(x_1) dx_1 \right) \left(\int_{A_2} \rho_2(x_2) dx_2 \right)$$

Z własności całek wielokrotnych wiemy, że iloczyn dwóch niezależnych całek możemy swobodnie zapisać jako jedną całkę podwójną z iloczynu funkcji. Zrównując zatem lewą i prawą stronę, otrzymujemy:

$$\int_{A_1} dx_1 \int_{A_2} dx_2 \rho(x_1, x_2) = \int_{A_1} dx_1 \int_{A_2} dx_2 (\rho_1(x_1) \rho_2(x_2))$$

Skoro ta równość całek musi zachodzić dla absolutnie **każdego, dowolnie wybranego** obszaru całkowania A_1 i A_2 , to na mocy twierdzeń analizy matematycznej same funkcje podcałkowe muszą być sobie równe prawie wszędzie. Stąd natychmiast wynika nasza teza: $\rho(x_1, x_2) = \rho_1(x_1) \rho_2(x_2)$.

Dla zmiennych niezależnych wartość oczekiwana iloczynu jest iloczynem wartości oczekiwanych: $\langle x_1 x_2 \rangle = \langle x_1 \rangle \langle x_2 \rangle$ (ponieważ całka podwójna w wartości oczekiwanej analogicznie rozbija się na dwie niezależne).

3.5 Przykład fizyczny: Rozkład Maxwella (1860) – oryginalne wyprowadzenie

To, co teraz zrobimy, to jedno z najpiękniejszych wyprowadzeń w historii fizyki. James Clerk Maxwell w 1860 roku postawił sobie za cel znalezienie rozkładu prędkości w gazie nieoddziałujących ze sobą cząsteczek w równowadze termicznej.

Maxwell przyjął tylko dwa, niesamowicie naturalne założenia:

1. **Niezależność składowych:** Prędkości w poszczególnych kierunkach kartezjańskich (v_x, v_y, v_z) są wzajemnie niezależnymi zmiennymi losowymi. Łączne prawdopodobieństwo to iloczyn prawdopodobieństw składowych:

$$p(v_x, v_y, v_z) = f(v_x) f(v_y) f(v_z)$$

2. **Izotropia przestrzeni:** W gazie w równowadze nie ma żadnego wyróżnionego kierunku. Jeśli obrócimy układ współrzędnych, rozkład nie powinien się zmienić. Gęstość musi zależeć wyłącznie od modułu prędkości (lub jej kwadratu $v^2 = v_x^2 + v_y^2 + v_z^2$), a nie od kierunku:

$$p(v_x, v_y, v_z) = \Psi(v_x^2 + v_y^2 + v_z^2)$$

Zestawiając oba założenia, dostajemy równanie funkcyjne:

$$f(v_x) f(v_y) f(v_z) = \Psi(v_x^2 + v_y^2 + v_z^2)$$

Rozpatrzmy teraz przypadek, gdy cząstka porusza się tylko wzdłuż osi X, czyli $v_y = v_z = 0$. Oznaczmy dla wygody $f(0) = c$. Mamy wtedy:

$$f(v_x) \cdot c \cdot c = \Psi(v_x^2 + 0 + 0) \implies c^2 f(v_x) = \Psi(v_x^2)$$

Zatem $f(v_x) = \Psi(v_x^2)/c^2$. Wstawiając to z powrotem do głównego równania, otrzymujemy:

$$\frac{\Psi(v_x^2)}{c^2} \frac{\Psi(v_y^2)}{c^2} \frac{\Psi(v_z^2)}{c^2} = \Psi(v_x^2 + v_y^2 + v_z^2)$$

$$\Psi(v_x^2)\Psi(v_y^2)\Psi(v_z^2) = c^6\Psi(v_x^2 + v_y^2 + v_z^2)$$

Dla ułatwienia zapisu zmienimy zmienne: niech $\alpha = v_x^2$, $\beta = v_y^2$, $\gamma = v_z^2$. Równanie przyjmuje postać:

$$c^6\Psi(\alpha)\Psi(\beta)\Psi(\gamma) = c^6\Psi(\alpha + \beta + \gamma)$$

Zróźniczkujemy teraz to wyrażenie stronami po zmiennej α :

$$c^6\Psi'(\alpha)\Psi(\beta)\Psi(\gamma) = c^6\Psi'(\alpha + \beta + \gamma)$$

Podzielimy to nowe równanie przez równanie wyjściowe:

$$\frac{\Psi'(\alpha)}{\Psi(\alpha)} = \frac{\Psi'(\alpha + \beta + \gamma)}{\Psi(\alpha + \beta + \gamma)}$$

Zauważmy, że lewa strona zależy *tylko* od α , a prawa od sumy α, β, γ . Taka równość dla dowolnych wartości zmiennych może zajść tylko wtedy, gdy obie strony są równe stałej! Podstawmy $\alpha = \gamma = 0$:

$$\frac{\Psi'(\beta)}{\Psi(\beta)} = \frac{\Psi'(0)}{\Psi(0)} = \text{const.} \equiv -\lambda$$

Gdzie ujemny znak wybieramy przezornie, wiedząc, że prawdopodobieństwo dla nieskończonych prędkości musi spadać do zera. Otrzymaliśmy proste równanie różniczkowe: $\frac{d\Psi}{\Psi} = -\lambda d\beta$, które całkujemy:

$$\ln \Psi(\beta) = -\lambda\beta + C \implies \Psi(\beta) = Ae^{-\lambda\beta}$$

Stąd:

$$\Psi(\alpha + \beta + \gamma) = \frac{A^3}{c^6}e^{-\lambda(\alpha+\beta+\gamma)}$$

Ostatecznie nasza łączna gęstość prawdopodobieństwa prędkości wynosi:

$$p(v_x, v_y, v_z) = Be^{-\lambda(v_x^2+v_y^2+v_z^2)}$$

Pozostaje nam jedynie wyznaczyć stałą normalizacyjną B (całka po całej przestrzeni pędów musi wynosić 1) oraz parametr fizyczny λ . Wiemy z analizy, że całka Gaussa to:

$$\int_{-\infty}^{\infty} e^{-\lambda v_x^2} dv_x = \sqrt{\frac{\pi}{\lambda}}$$

Co dla trzech wymiarów daje $B = \left(\frac{\lambda}{\pi}\right)^{3/2}$.

Parametr λ możemy powiązać ze średnią energią kinetyczną cząsteczek. Obliczając wartość oczekiwaną kwadratu prędkości $\langle v^2 \rangle$ (szczegóły na ćwiczeniach), otrzymujemy:

$$\frac{m\langle v^2 \rangle}{2} = \frac{3m}{4\lambda}$$

Wybiegając nieco w przód do termodynamiki – z zasady ekwipartycji energii wiemy, że średnia energia kinetyczna zależy od temperatury T :

$$\frac{m\langle v^2 \rangle}{2} = \frac{3}{2}kT$$

Zestawiając te dwa wzory, otrzymujemy ostatecznie parametr $\lambda = \frac{m}{2kT}$. W ten sposób abstrakcyjna matematyka połączyła się z mikroskopowym ruchem gazu i makroskopową temperaturą.

3.6 Funkcje zmiennych losowych

Często znamy rozkład zmiennej losowej X (o gęstości $\rho_X(x)$), ale w eksperymencie mierzymy wielkość Y , która jest pewną jej funkcją: $Y = f(X)$. Pytanie brzmi: jaka jest gęstość prawdopodobieństwa $\rho_Y(y)$ nowej zmiennej?

Mechanizm jest niezwykle elegancki i opiera się na delcie Diraca. Obliczamy prawdopodobieństwo, że nowa zmienna przyjmie wartość ze zbioru A :

$$P(Y \in A) = \int_A \rho_Y(y) dy = \int_{f^{-1}(A)} \rho_X(x) dx$$

Zamiast całkować po skomplikowanym przeciwobrazie $f^{-1}(A)$, możemy całkować po całej przestrzeni używając dystrybucji Diraca jako matematycznego „sita”, które wyłapie tylko te wartości x , dla których $f(x) = y$:

$$P(Y \in A) = \int_A dy \int_{-\infty}^{\infty} dx \rho_X(x) \delta(y - f(x))$$

Porównując lewą i prawą stronę, natychmiast odczytujemy wzór na nową gęstość:

$$\rho_Y(y) = \int_{-\infty}^{\infty} \rho_X(x) \delta(y - f(x)) dx = \langle \delta(y - f(X)) \rangle_{\rho_X}$$

Ten potężny wzór można łatwo uogólnić na przypadek wielowymiarowych zmiennych losowych $\mathbf{X} = (X_1, X_2, \dots, X_n)$ oraz wielowymiarowych funkcji $\mathbf{Y} = \mathbf{f}(\mathbf{X})$:

$$\rho_Y(\mathbf{y}) = \int d\mathbf{x} \rho_X(\mathbf{x}) \delta(\mathbf{y} - \mathbf{f}(\mathbf{x})) = \langle \delta(\mathbf{y} - \mathbf{f}(\mathbf{X})) \rangle_{\rho_X}$$

Przykład: Suma dwóch niezależnych zmiennych losowych

Mamy dwie rzutki, dwie kostki lub dwie cząstki gazu. Mamy zmienne niezależne X_1 i X_2 o zadanych rozkładach $\rho_1(x_1)$ oraz $\rho_2(x_2)$. Szukamy rozkładu gęstości ich **sumy**: $Y = X_1 + X_2$.

Używamy naszego wzoru z deltą Diraca (pamiętając, że dla zmiennych niezależnych łączna gęstość to iloczyn $\rho_1 \cdot \rho_2$):

$$\rho_Y(y) = \langle \delta(y - X_1 - X_2) \rangle_{\rho_1 \times \rho_2} = \int \int dx_1 dx_2 \delta(y - x_1 - x_2) \rho_1(x_1) \rho_2(x_2)$$

Całkujemy po zmiennej x_2 . Delta Diraca „wybiera” za nas wartość $x_2 = y - x_1$:

$$\rho_Y(y) = \int_{-\infty}^{\infty} dx_1 \rho_1(x_1) \rho_2(y - x_1)$$

Wniosek jest fundamentalny: gęstość prawdopodobieństwa sumy niezależnych zmiennych losowych jest splotem ich indywidualnych gęstości. To właśnie ten fakt leży u podstaw słynnego Centralnego Twierdzenia Granicznego.

4 Zdarzenia przypadkowe i ruchy Browna

4.1 Kropla mleka i siła życiowa

Na wykładzie mieliśmy okazję spojrzeć przez mikroskop na kroplę mleka rozcieńczoną w wodzie. Drobne kuleczki tłuszczu zawieszone w płynie nieustannie, chaotycznie przemieszczały się we wszystkich kierunkach. Ten nieustanny taniec mikroskopijnych cząstek to właśnie **ruchy Browna**. Zjawisko to było odkryte w 1827 roku przez szkockiego botanika Roberta Browna, który badał pod mikroskopem małe cząstki pyłku kwiatowego zawieszone w wodzie. Ponieważ

pyłek to element rozrodczy roślin, Brown sądził, że ten nieustanny ruch to przejaw mitycznej „siły życiowej” (*vis vitalis*) – legendarnego tchnienia, które miało odróżniać materię ożywioną od nieożywioną. Brown przeprowadził z pyłkiem imponującą liczbę eksperymentów. Zauważył, że ruchy te stają się szybsze wraz ze wzrostem temperatury płynu. Co więcej, pokazał, że ten ruch nigdy nie ustaje – próbki zamknięte w szafie nawet na kilka lat po wyjęciu wciąż wykazywały tę samą, nieprzerwaną aktywność.

Prawdziwe zrozumienie tego zjawiska przyszło dopiero na początku XX wieku (w latach 1905–1906), głównie za sprawą prac Alberta Einsteina oraz Mariana Smoluchowskiego. Uodowodnili oni, że ruchy Browna są efektem wynikającym bezpośrednio z faktu, że woda składa się z niewidzialnych atomów. Niezliczone miliardy maleńkich cząsteczek wody nieustannie i ze wszystkich stron zderzają się z dużą, zawieszoną w niej cząstką pyłku.

W tym miejscu pojawia się jednak bardzo naturalne pytanie: skoro cząstka pyłku jest bombardowana ze wszystkich stron przez miliardy cząsteczek wody, to czy średnio rzecz biorąc nie uderza w nią tyle samo cząsteczek z lewej, co z prawej strony? Dlaczego więc cząstka w ogóle się przemieszcza, a nie stoi w miejscu, wibrując jedynie wokół środka?

Marian Smoluchowski odpierał on te zarzuty, porównując ruch cząstki Browna do grającego w kasynie hazardzisty:

„Jest to taki sam błąd rozumowania, jak gdyby człowiek uprawiający grę hazardową (np. rzucanie kostki) sądził, że nigdy większej straty ani też większego zysku mieć nie będzie, niż wynosi stawka na jeden rzut. Wiemy dobrze, że szczęście i nieszczęście zwykle niezupełnie się równoważą; że im dłużej gra trwa, tym większa jest przeciętna suma albo wygrana, albo stracona”.

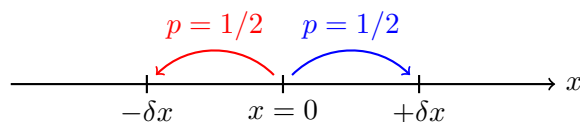
Załóżmy, że w każdej rundzie rzucamy monetą: jeśli wypadnie orzeł, gracz zyskuje 1 jednostkę kapitału ($+1$ z $p = 1/2$), a jeśli reszka – traci 1 jednostkę (-1 z $p = 1/2$). Średnio rzecz biorąc, wygrana z takiej gry to zero. Ale jeśli spojrzymy na rozkład kapitału w czasie, zobaczymy, że im dłużej nasz hazardzista gra, tym bardziej jego kapitał odchyła się od wartości początkowej. Kwadrat odchylenia stanu konta naszego gracza od stanu początkowego $(N_t - N_0)^2$ rośnie proporcjonalnie do liczby rozegranych partii n . Oznacza to, że typowe odchylenie od średniej (czyli statystyczna nadwyżka wygranych nad przegranych) rośnie jak pierwiastek z liczby prób: $\sqrt{n}$.

Aby przenieść tę hazardową analogię na grunt fizyki, musimy dokonać prostego tłumaczenia pojęć. Rola „rzutu monetą” przypada pojedynczym uderzeniom niewidzialnych molekuł cieczy, a „kapitałem” gracza staje się aktualne położenie cząstki pyłku względem punktu startu. Uderzenie z jednej strony to nasza wygrana, uderzenie z przeciwnej – strata. Zgodnie z prawami statystyki, nawet przy idealnej symetrii, z upływem czasu cząstka będzie doświadczać coraz większych fluktuacji. Oznacza to, że co rusz pojawiać się będą statystyczne **nadwyżki** uderzeń z jednego kierunku, które zaczną chaotycznie przesuwac pyłek w przestrzeni.

Sprawdźmy, czy te statystyczne nadwyżki mają w ogóle znaczenie w skali fizycznej. Ziarenko zawieszony jest uderzany przez molekuły cieczy średnio $n = 10^{20}$ razy w ciągu zaledwie jednej sekundy. Zgodnie z regułą opisaną wyżej, przeciętna nadwyżka uderzeń z jednej strony nad drugą wyniesie w tym czasie pierwiastek z tej gigantycznej liczby zderzeń, czyli około $\sqrt{10^{20}} = 10^{10}$ (dziesięć miliardów)! Nawet jeżeli pęd przekazany przez pojedynczą molekułę jest znikomo mały, to sumaryczny efekt dziesięciu miliardów „nadmiarowych” uderzeń z jednego kierunku jest wystarczająco potężny, by popychać cząstkę pyłku w widoczny pod mikroskopem sposób.

4.2 Matematyczny model błędzenia przypadkowego w 1D

Zbudujmy prosty, jednowymiarowy model tego zjawiska. Rozważmy cząstkę na osi x , która w chwili $t = 0$ startuje z punktu $x = 0$.



Wprowadzamy trzy proste założenia:

1. Każda cząstka w pojedynczym kroku czasowym δt wykonuje skok w prawo lub w lewo o ustaloną odległość δx . Dokładne wartości δt oraz δx zależą od fizycznych własności cieczy: jej lepkości, temperatury i rozmiaru samej cząstki.
2. Prawdopodobieństwo skoku w lewo lub w prawo jest równe i wynosi $1/2$. Co kluczowe, kierunek skoku **nie zależy od tego, co działo się wcześniej** – brakuje tu jakiegokolwiek pamięci, a kolejne kroki są zdarzeniami w pełni niezależnymi.
3. Różne cząstki w cieczy przemieszczają się niezależnie od siebie.

Jakie są konsekwencje takiego modelu dla ruchu po n krokach czasowych? Położenie po n -tym kroku opisuje równanie:

$$X(n) = X(n-1) + \Delta X$$

gdzie zmienna losowa (skok) ΔX przyjmuje wartość $+\delta x$ z $p = 1/2$ oraz $-\delta x$ również z $p = 1/2$.

1. Średnie przesunięcie

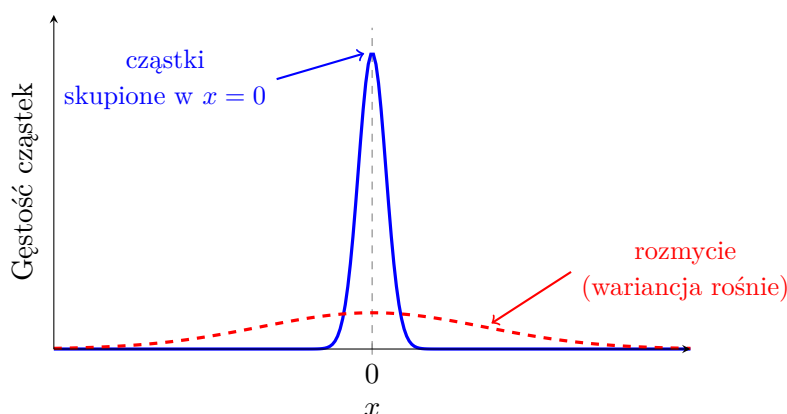
Wyśrednimy nasze równanie stronami:

$$\langle X(n) \rangle = \langle X(n-1) \rangle + \langle \Delta X \rangle$$

Ponieważ $\langle \Delta X \rangle = 0$, otrzymujemy:

$$\langle X(n) \rangle = \langle X(n-1) \rangle = \langle X(n-2) \rangle = \dots = \langle X(0) \rangle = 0$$

Zgodnie z przewidywaniami, średnie przesunięcie cząstki jest równe 0. Gdybyśmy zamiast jednej cząstki wypuszczali całą ich gromadę z $x = 0$ i przyjęli założenie że ich ruchy wzajemnie na siebie nie wpływają, to powyższy wynik oznaczałby, że choć to skupisko cząstek z upływem czasu rozmywa się na boki, środek paczki cały czas pozostaje w punkcie $x = 0$.



2. Średnie przesunięcie kwadratowe

Podnieśmy równanie ruchu do kwadratu:

$$X^2(n) = (X(n-1) + \Delta X)^2 = X^2(n-1) + 2\Delta X \cdot X(n-1) + (\Delta X)^2$$

i znów wyśredniujemy stronami:

$$\langle X^2(n) \rangle = \langle X^2(n-1) \rangle + 2\langle \Delta X \cdot X(n-1) \rangle + \langle (\Delta X)^2 \rangle$$

W tym miejscu korzystamy z założenia o braku pamięci. Skoro kierunek nowego skoku ΔX jest niezależny od przeszłego położenia $X(n-1)$, to wartość oczekiwana ich iloczynu jest iloczynem wartości oczekiwanych. A skoro $\langle \Delta X \rangle = 0$, to cały człon mieszany znika:

$$2\langle \Delta X \cdot X(n-1) \rangle = 2 \cdot \langle \Delta X \rangle \cdot \langle X(n-1) \rangle = 0$$

Wartość oczekiwana kwadratu pojedynczego skoku wynosi natomiast $\langle (\Delta X)^2 \rangle = (\delta x)^2$. Zostaje nam prosta rekurencja:

$$\langle X^2(n) \rangle = \langle X^2(n-1) \rangle + (\delta x)^2$$

Zwijając to równanie dla n kroków aż do zera (zakładając $\langle X^2(0) \rangle = 0$), dostajemy:

$$\langle X^2(n) \rangle = n(\delta x)^2$$

Ostatni krok to zamiana liczby kroków n na fizyczny czas. Skoro każdy krok trwa δt , to całkowity czas $t = n\delta t$, skąd $n = t/\delta t$. Podstawiając to do równania:

$$\langle X^2(t) \rangle = \left[\frac{(\delta x)^2}{\delta t} \right] t$$

Dla wygody wprowadzamy w tym miejscu wielkość zwaną **stałą dyfuzji** D , zdefiniowaną jako $D = \frac{(\delta x)^2}{2\delta t}$. W ten sposób otrzymujemy najważniejszy wzór na dyfuzję w jednym wymiarze:

$$\langle x^2 \rangle = 2Dt$$

W przypadku przestrzeni trójwymiarowej wariancje z trzech wzajemnie niezależnych osi po prostu się dodają, co daje wynik 3 razy większy:

$$\langle r^2 \rangle = \langle x^2 \rangle + \langle y^2 \rangle + \langle z^2 \rangle = 6Dt$$

4.3 Czy dyfuzja jest efektywna?

Aby określić rozmiar „rozmywania” się paczki cząstek po upływie zadanego czasu, patrzymy na odchylenie standardowe położenia, będące pierwiastkiem z wariancji:

$$\sqrt{\langle x^2 \rangle} = \sqrt{2Dt}$$

Sprawdźmy teraz, jak ten statystyczny mechanizm transportu radzi sobie w różnych rzędach wielkości, przyjmując wartość stałej dyfuzji dla małych cząsteczek w wodzie w temperaturze pokojowej wynoszącą $D \sim 10^{-5} \text{ cm}^2/\text{s}$. Z przekształcenia wzoru wiemy, że czas dyfuzji to mniej więcej $t \approx x^2/2D$.

- **Skala mikro (wymiary bakterii):** Dla odległości rzędu $x = 10^{-4} \text{ cm}$, czas wynosi $t \approx (10^{-4})^2/(2 \cdot 10^{-5}) \approx 5 \cdot 10^{-4} \text{ s}$. Zaledwie 0.5 milisekundy! W mikroskopijnym świecie pojedynczych komórek transport dyfuzyjny jest niesamowicie efektywny.
- **Skala próbki chemicznej:** Jeśli chcemy, aby nasza substancja przebyła $x = 1 \text{ cm}$, czas wzrasta lawinowo do $t \approx 1^2/(2 \cdot 10^{-5}) = 50\,000 \text{ sekund}$. To oznacza aż ~ 14 godzin. Dyfuzja w tej skali staje się procesem niezwykle mało efektywnym.
- **Skala makroskopowa (pokój):** Powszechnie uważa się, że zapachy w pokoju rozchodzą się dzięki mechanizmowi dyfuzji. Policzmy to! Dla dystansu 5 m (500 cm), dyfuzja wylicza nam czas na przebycie tego dystansu równy dobremu kilku miesiącom. W rzeczywistości, jeżeli po chwili czujemy na drugim końcu pokoju zapach perfum, nie jest to absolutnie zasługa powolnej i ślepej dyfuzji, lecz mechanizmu **konwekcji** – makroskopowych, fizycznych prądów mieszających powietrze w pomieszczeniu.