Sveučilište J.J.Strossmayera u Osijeku
Odjel za matematiku

Petra Pandža

Nasumični hod u jednoj dimenziji

Sveučilište Josipa Jurja Strossmayera u Osijeku
Odjel za matematiku

Petra Pandža

Nasumični hod u jednoj dimenziji

Diplomski rad

Voditelj: Doc. dr. sc. Z. Glumac

Osijek, 2012

Sadržaj

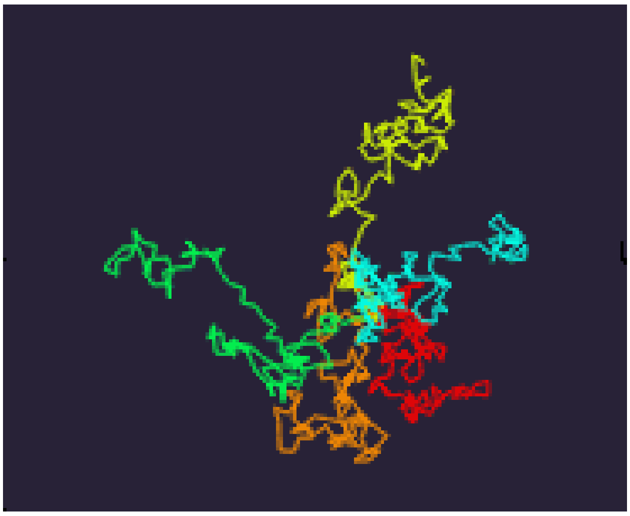
1 Uvod

Sustave građene od velikog broja čestica komplicirano je promatrati sa gledišta klasične mehanike jer je za tako veliki broj čestica gotovo nemoguće naći jednadžbu gibanja svake pojedine čestice. Zato se takve sustave promatra statistički, odnosno, procese u kojima oni sudjeluju smatramo nizom slučajnih pokusa. Takve procese nazivamo stohastičkim procesima jer njihov ishod ovisi o slučaju. Tema ovoga rada upravo je jedan od najjednostavnjijig stohastičkih procesa: nasumičan hod u jednoj dimenziji.
Jednostavna interpretacija problema nasumičnog hoda u jednoj dimenziji je ova: čestica se giba u jednoj dimenziji tako da izvodi pomake u desnu stranu s vjerojatnosti p i u lijevu stranu s vjeroatnosti $q = 1 - p$.
Želimo odrediti pomak čestice od njezina početnog položaja, vjerojatnost da se čestica nakon određenog broja koraka nađe u nekom položaju i slično.
U prvom poglavlju rada govori se o problemu nasumičnog hoda u jednoj dimenziji. Tu će se računati sljedeće: pomak čestice od ishodišta, srednji broj koraka čestice u lijevu i u desnu stranu te vjerojatnost pomaka, odnosno, broja koraka čestice.
U drugom poglavlju problem nasumičnog hoda poopćuje se na više dimenzija te za taj problem određuje vjerojatnost da se čestica nakon određenog broja koraka nađe u određenom položaju.

2 NASUMIČNI HOD I BINOMNA RASPODJELA

2.1 Osnovni statistički pojmovi i primjeri

Nasumičan hod je najjednostavniji primjer stohastičkog procesa. Stohastičkim procesom nazivamo niz pokusa gdje ishod svakog pojedinog pokusa ovisi o slučaju. Dakle, to je niz pokusa koji se odvijaju kontinuirano s vremenom, a rezultat svakog pokusa je slučajan.
Takvi su procesi npr. difuzija tekućina ili plinova, Brownovo gibanje, radioaktivni raspad i sl. Uzmimo za primjer difuziju dvije tekućine. Tekućine se sastoje od velikog broja molekula, miješanjem tekućina dolazi do sudara molekula; dakle, molekule se sudaraju i time mijenjaju svoju brzinu i smjer. Odnosno, možemo reći da je njihovo trenutno stanje rezultat slučajnog pokusa. Želimo, na primjer, odrediti vrijeme potrebno da se takav sustav uravnoteži.



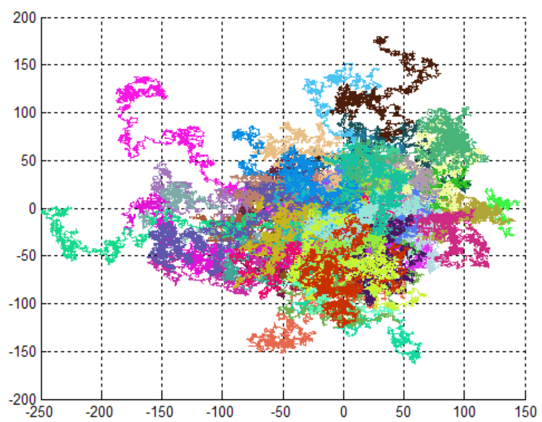
Slika 1: Nasumični hod u dvije dimenzije - slika prikazuje difuziju pet čestica u ravnini koje kreću od istog ishodišta

Na ovakva i slična pitanja, odgovor nam daje područje fizike koje zovemo statistička fizika. Ako bismo rješenje problema pokušali naći pomoću klasične mehanike, bilo bi to neizvedivo upravo zato jer se promatrani sustav sastoji od velikog broja čestica (molekula) te je stoga nemoguće naći jednadžbu gibanja svake pojedine čestice. Niti kvantnomehanički pristup ne bi dao zadovoljavajuće rješenje problema. Dakle, najjednostavnije ćemo naći rješenje tako da dani problem promatramo kao sustav sastavljen od velikog broja čestica i njegovo rješenje tražimo pomoću statističke fizike. Odnosno, tražimo vjerojatnost nekog stanja sustava (u prethodnom primjeru to je vjerojatnost da se ravnoteža sustava uspostavi nakon nekog vremena t).

Dakle, definirajmo sada problem nasumičnog hoda. Nasumičan se hod pojavljuje u različitim područjima ljudske djelatnosti, često ga zovemo i pijančev hod. Pojednostavljeno, problem nasumičnog hoda možemo opisati na sljedeći način: promatramo pijanog čovjeka koji započinje svoje gibanje od stupa koji se nalazi na ulici. Svaki čovjekov korak jednake je duljine, l. S obzirom da je čovjek pijan, svaki njegov korak neovisan je o prethodnom koraku. Neka je vjerojatnost da čovjek napravi korak u desno p, a vjerojatnost da napravi korak u lijevo $q = 1 - p$. Općenito,

$$p \neq q.$$

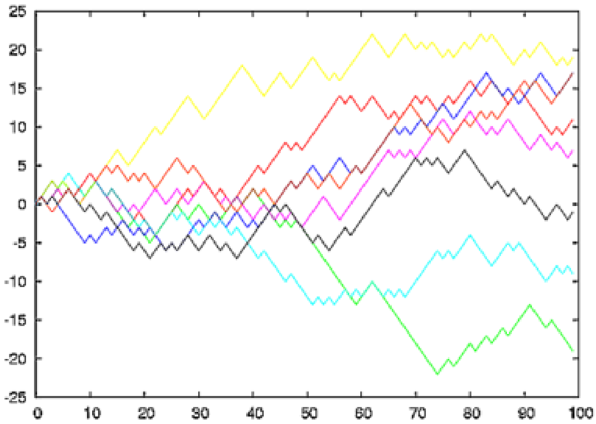
Neka ulica leži na x-osi i to tako je stup od kojeg čovjek započinje gibanje smješten u ishodištu $x = 0$. Položaj čovjeka nakon N koraka jednak je $x = ml$, gdje je m cijeli broj.
Pitamo se, kolika je vjerojatnost da se čovjek nakon N koraka nalazi u položaju $x = ml$?



Slika 2: Nasumični hod u dvije dimenzije - na slici je prikazano gibanje više pijanih ljudi

Ovaj je primjer jednostavna ilustracija nasumičnog hoda u jednoj dimenziji, ali se vrlo lako može primijeniti i u fizici: dodajemo N vektora jednake duljine, ali različitih smjerova (ili smjerova određenih nekom raspodjelom vjerojatnosti), te tražimo vjerojatnost da rezultanta ovih vektora ima određeni iznos i smjer.
Nasumičan hod u jednoj dimenziji, kojim ćemo se ovdje najviše baviti, ili njegova poopćenja na više dimenzija mogu se primijeniti u različitim područjima fizike. Na primjer
- pri difuziji molekula u plinu ili tekućini: što smo već ranije spomenuli, molekule se pri difuziji sudaraju pa možemo tražiti udaljenost molekule nakon N sudara ili nešto slično.
- u teoriji magnetizma: atom ima spin $\frac{1}{2}$ i magnetski moment μ. Njegov spin, s obzirom na smjer, može biti prema gore ili prema dolje. Ako pretpostavimo da su obje mogućnosti

jednako vjerojatne, pitamo se, koliki je ukupni magnetski moment N atoma.
- pri računanju intenziteta svjetlosti N nekoherentnih izvora: amplituda svakog izvora dana je dvodimenzionalnim vektorom čiji smjer određuju faze pomaka koje su nasumične. Tada je ukupan intenzitet svih izvora svjetlosti određen statističkom sredinom.



Slika 3: Nasumični hod u jednoj dimenziji za osam različitih šetač koji kreću iz istog ishodišta nula. Slika prikazuje ovisnost trenutnog položaja šetač o vremenu.

2.2 Problem nasumičnog hoda u jednoj dimenziji

Pojednostavljen problem nasumičnog hoda u jednoj dimenziji može se definirati na sljedeći način: čestica se giba u jednoj dimenziji (npr. po x-osi), tako da izvodi pomake jednake duljine l u lijevu i desnu stranu, vjerojatnost pomaka u desno jednaka je p, dok je vjerojatnost da je korak u lijevu stranu jednaka $q = 1 - p$. Potrebno je odrediti pomak čestice ili njezin položaj nakon N koraka ili nešto slično.
Kako čestica u svom gibanju napravi ukupno N koraka, njezin je pomak cijeli broj iz intervala

$$-N \leq m \leq N$$

Čestica će se nakon N koraka naći u položaju

$$x = ml$$

Treba naći vjerojatnost $P_N(m)$ da je pomak čestice nakon N koraka jednak m.
Neka je n_D broj koraka u desno, a n_L broj koraka koje čestica napravi u lijevo. Ukupan broj koraka N očito je tada

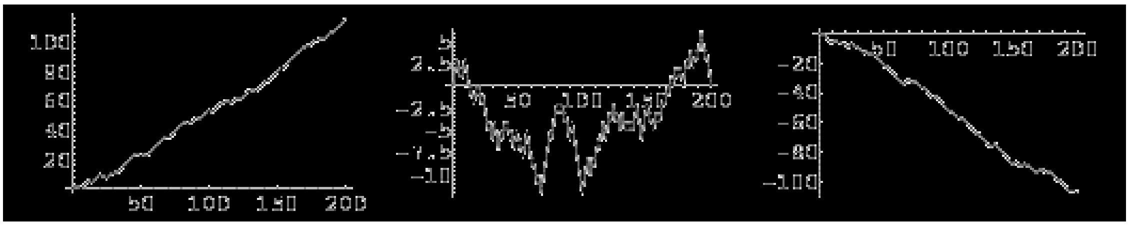
$$N = n_D + n_L \tag{2.2.1}$$

Pomak je

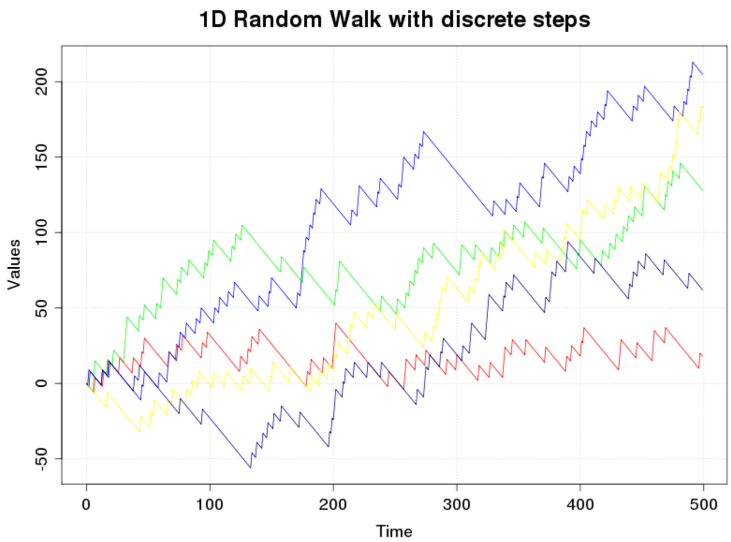
$$m = n_D - n_L \tag{2.2.2}$$

Odnosno

$$m = n_D - (N - n_D) = 2n_D - N \tag{2.2.3}$$



Slika 4: Nasumični hod u jednoj dimenziji. Slika prikazuje raspodjelu vjerojatnosti $W(n_D)$ da je n_D koraka u desno za česticu koja učini $N = 200$ koraka, pri čemu su vjerojatnosti koraka u desno redom $p = 0.1$, $p = 0.5$ i $p = 0.9$



Slika 5: Nasumični hod u jednoj dimenziji kada je pomak čestice diskretna varijabla

Dakle, pomak čestice od ishodišta određen je upravo brojem koraka koje čestica napravi u desno, n_D. S obzirom da se čestica giba nasumično, svaki sljedeći korak neovisan je o prethodnom. Odnosno, kažemo da su koraci statistički neovisni. To znači da, neovisno o prethodnim koracima, vjerojatnost svakog koraka može biti

$$\begin{aligned} p &= \textit{vjerojatnost da je korak u desno} \\ q &= \textit{vjerojatnost da je korak u lijevo} \end{aligned}$$

Tada je vjerojatnost da od ukupno N koraka koje čestica napravi n_D bude u desno, a n_L u lijevo dana izrazom

$$\underbrace{p \cdot p \cdot p \cdot \dots \cdot p}_{n_D} \cdot \underbrace{q \cdot q \cdot q \cdot \dots \cdot q}_{n_L} = p^{n_D} q^{n_L} \tag{2.2.4}$$

Dakle, n_D je broj koraka udesno, a p vjerojatnost svakog tog pojedinog koraka, stoga ćemo p pomnožiti upravo n_D puta. Isto tako, vjerojatnost q da je korak u lijevo, množimo n_L puta, odnosno toliko puta koliko je lijevih koraka.

Zanima nas koja je vjerojatnost da od ukupno N koraka koje čestica napravi n_D koraka bude u desno. Zato prvo treba odrediti broj načina na koje čestica može napraviti N koraka tako da njih n_D bude u desno, a n_L u lijevo. Ili, matematički rečeno: problem je kako N objekata od kojih je n_D jedne vrste, a n_L druge vrste smjestiti na N mogućih mjesta.

Razmišljamo na sljedeći način, od N objekata koje imamo prvo mjesto može zauzeti bilo koji od njih, kada za prvo mjesto odaberemo jedan objekt, znači da se na drugom mjestu može naći bilo koji od $N-1$ preostalih objekata. Prema tome, prva dva mjesta mogu biti zauzeta na $N \cdot (N-1)$ način. Sada nam je preostalo $N-2$ objekta koja možemo smjestiti na treće mjesto, dakle, prva tri mjesta mogu biti zauzeta na $N \cdot (N-1) \cdot (N-2)$ moguća načina. Analogno, posljednje mjesto može biti zauzeto na samo jedan način. Prema tome N mjesta može biti zauzeto na

$$N \cdot (N-1) \cdot (N-2) \cdot \dots \cdot 1 = N!$$

mogućih načina. Dakle,broj mogućih načina na koji N objekata možemo posložiti na N mogućih mjesta zapravo je permutacija skupa od N elemenata.

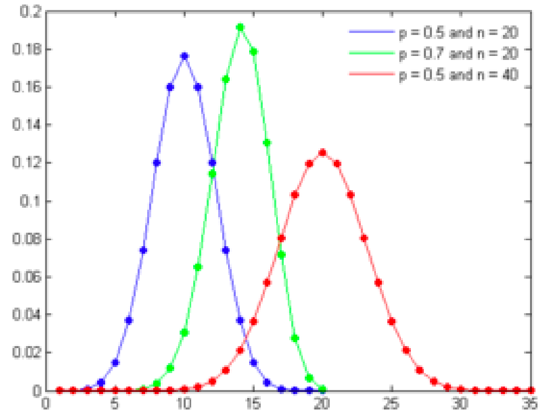
Kako u našem slučaju imamo N objekata od kojih su svi objekti n_D jednaki te svi objekti n_L jednaki, to znači da će sve permutacije istog objekta (bilo n_D ili n_L) biti jednake. Stoga je broj permutacija skupa od N elemenata od kojih su n_D jedne vrste, a n_L druge vrste umjesto $N!$ jednak

$$\frac{N!}{n_D! \cdot n_L!}$$

Upravo je to broj načina na koji čestica može napraviti N koraka, tako da je n_D koraka u desno a n_L koraka u lijevo.

Dakle, vjerojatnost $W(n_D)$ da je od ukupno N koraka n_D koraka u desno dana je umnoškom vjerojatnosti određenog niza koraka i brojem načina na koji taj niz možemo napraviti. Odnosno,

$$W(n_D) = \frac{N!}{n_D! \cdot n_L!} p^{n_D} q^{n_L} \tag{2.2.5}$$



Slika 6: Binomna raspodjela vjerojatnosti za različite vrijednosti ukupnog broja koraka N i vjerojatnosti koraka u desno p

Primjer:
Neka je ukupan broj koraka $N = 3$. Treba odrediti vjerojatnost $W(n_D)$ kada je $n_D = 3$, odnosno kada su sva tri koraka koje čestica napravi u desnu stranu.
Ako je $n_D = 3$, onda je, očigledno, $n_L = 0$, dakle, broj koraka u lijevo je nula.
Prvo odredimo broj načina na koji sva tri koraka mogu biti udesno; to izračunamo iz izraza

$$\frac{N!}{n_D! \cdot n_L!} = \frac{3!}{3! \cdot 0!} = 1$$

Dakle, postoji samo jedan način na koji sva tri koraka mogu biti u desno.
Do istog rezultat možemo doći ako razmišljamo ovako: tražimo broj načina na koji tri objekta (koraka) možemo rasporediti na tri različita mjesta. Opet, prvo mjesto može zauzeti bilo koji od tri objekta itd. Odnosno, ukupan broj načina je

$$3 \cdot 2 \cdot 1 = 3! = 6$$

S obzirom da su sva tri koraka udesno, $n_D = 3$, to znači da broj načina više neće biti 6 jer su ishodi koji se razlikuju samo po permutaciji istih objekata (koraka u desno) jednaki. Stoga je ukupan broj načina na koji sva tri koraka mogu biti udesno jednak

$$\frac{3!}{3!} = 1$$

Nadalje, kako su sva tri koraka udesno, ukupnu vjerojatnost takvog niza koraka dobit ćemo upravo množenjem vjerojatnosti svakog pojedinog koraka. Odnosno

$$p \cdot p \cdot p = p^3$$

Dakle, vjerojatnost $W(3)$, da su od ukupno N koraka, tri koraka u desno, jednaka je

$$W(3) = p^3$$

$W(n_D)$ je, kao što smo već rekli, vjerojatnost da je od ukupnog broja koraka, n_D koraka u desno. Pomak čestice od ishodišta ovisi o broju koraka koje ta čestica napravi na desnu stranu i jednak je

$$m = 2n_D - N$$

Prema tome, možemo zaključiti kako je vjerojatnost $P_N(m)$ da se čestica nađe u položaju m jednaka vjerojatnosti $W(n_D)$. Dakle,

$$P_N(m) = W(n_D) \tag{2.2.6}$$

Broj koraka u desno n_D , kao i broj koraka u lijevo n_L možemo izraziti preko pomaka čestice, tada je

$$n_D = \frac{1}{2}(N+m) \quad n_L = \frac{1}{2}(N-m) \tag{2.2.7}$$

Uvrstimo li to u izraz (2.2.6), slijedi

$$P_N(m) = \frac{N!}{[(N+m)/2]! \cdot [(N-m)/2]!}\, p^{(N+m)/2}\,(1-p)^{(N-m)/2} \tag{2.2.8}$$

Dobili smo vjerojatnost $P_N(m)$ da je pomak čestice koja se nasumično giba u jednoj dimenizji jednak m. Ta je vjerojatnost dana binomnom raspodjelom jer relacija (2.2.5) predstavlja karakterističan izraz koji pronalazimo u razvoju $(p+q)^N$ prema binomnom teoremu. U posebnom slučaju, kada su vjerojatnosti pravljenja koraka u desno i u lijevo jednake, dakle, $p = q = \frac{1}{2}$ vjerojatnost $P_N(m)$ jednaka je

$$P_N(m) = \frac{N!}{[(N+m)/2]! \cdot [(N-m)/2]!} \left(\frac{1}{2}\right)^N \tag{2.2.9}$$

Primjer:
Neka je ponovno broj koraka koje čestica napravi $N = 3$, ali neka su sada zadane i vjerojatnosti p i q. Štoviše, vrijedi da je $p = q = \frac{1}{2}$, odnosno, vjerojatnosti koraka u lijevu i desnu stranu su jednake.
Treba odrediti vjerojatnost $W(n_D)$, da je n_D koraka u desno, odnosno $P_N(m)$, vjerojatnost

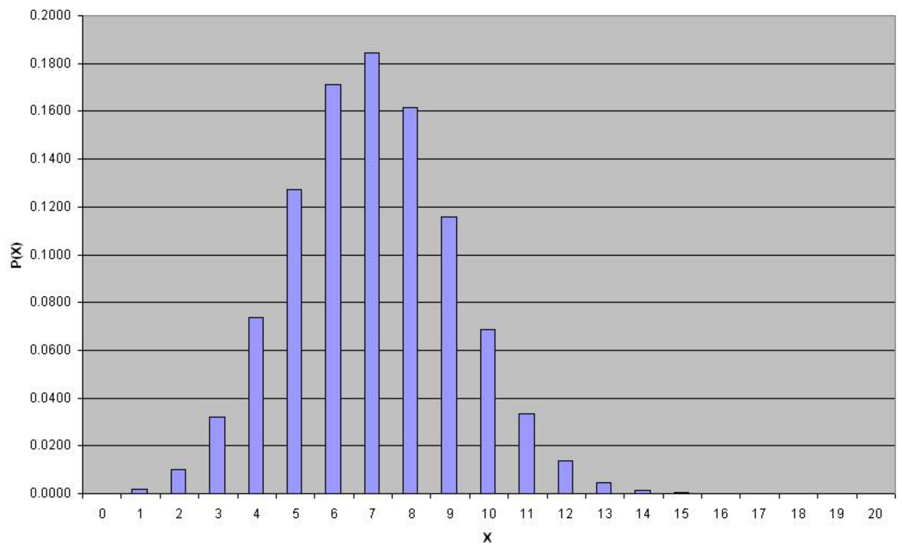
da je pomak čestice od ishodišta jednak m.
Pomak, kao što znamo, ovisi o ukupnom broju koraka N i o broju koraka koje čestica napravi na desnu stranu, n_D. Od ukupno $N = 3$ koraka, čestica u desnu stranu može napraviti niti jedan korak, zatim jedan, dva ili tri koraka. Odnosno

$$n_D = 0,\ 1,\ 2,\ 3$$

Ovisno o tome koliko koraka je napravljeno u desno, pomak čestice može biti

$$m = -3,\ -1,\ 1,\ 3$$

Vjerojatnost $W(n_D)$, odnosno $P_N(m)$, računamo iz izraza (2.2.9) jer je to specijalan slučaj izraza (2.2.8) za vjerojatnost $P_N(m)$, kada su vjerojatnosti p i q jednake. Prema tome slijedi: Vjerojatnost da od ukupnog broja koraka niti jedan korak nije udesno jednaka je vjerojatnosti da su sva tri napravljena koraka u desno, tj. $W(0) = W(3) = \frac{1}{8}$.
Isto tako, vjerojatnost da čestica napravi jedan korak u desno jednaka je vjerojatnosti pravljenja dva koraka, odnosno $W(1) = W(2) = \frac{3}{8}$.



Slika 7: Binomna raspodjela vjerojatnosti kada je $p = q = \frac{1}{2}$ i $N = 20$. Graf pokazuje vjerojatnost $W(n_D)$ da je n_D koraka u desno, odnosno da je $P_N(m)$ da je m pomak od ishodišta

2.3 Općenito o srednjim vrijednostima

Neka je u diskretna slučajna varijabla koja poprima bilo koju od M diskretnih vrijednosti

$$u_1,\ u_2,\ \dots\ ,\ u_M$$

Te neka su definirane pripadne vjerojatnosti

$$P(u_1),\ P(u_2),\ \dots\ ,\ P(u_M)$$

Želimo odrediti karakteristike raspodjele varijable u, što je moguće ako znamo raspodjelu vjerojatnosti $P(u)$. Ipak, tu je raspodjelu vrlo često teško ili gotovo nemoguće izračunati. Zato ćemo umjesto računanja vjerojatnosti $P(u)$ računati momente varijable u jer nam poznavanje n-tog momenta varijable omogućuje potpuno poznavanje karakteristika raspodjele varijable. Ovdje ćemo definirati neke od tih momenata.
Dakle, promatramo diskretnu varijablu u, $P(u_i)$ je vjerojatnost da varijabla u poprima vrijednost u_i. Prema tome, vjerojatnost da u može poprimiti bilo koju vrijednost od u_i, kada je $i \in [1,\ M]$ jednaka je

$$P(u_1) + P(u_2) + \cdots + P(u_M) \equiv \sum_{i=1}^{M} P(u_i)$$

Da bismo $P(u_i)$ mogli definirati kao vjerojatnost dane varijable, ona mora ispunjavati uvjet normiranja, odnosno mora vrijediti

$$\sum_{i=1}^{M} P(u_i) = 1 \tag{2.3.1}$$

Taj uvjet mora ispunjavati svaka funkcija vjerojatnosti; odnosno, mora vrijediti da je zbroj vjerojatnosti sigurnog događaja jednak je jedan.
Definirajmo sada srednju vrijednost, tj. prvi moment raspodjele varijable u.
Pojednostavljeno, srednju vrijednost varijable možemo objasniti kao npr. srednju ocjenu učenika na nekom ispitu. Ako sa u označimo ocjenu, a sa $P(u)$ broj učenika koji su tu ocjenu dobili, tada srednju ocjenu na ispitu dobijemo tako da svaku ocjenu pomnožimo brojem učenika koji imaju tu ocjenu, zatim sve zbrojimo i na kraju podijelimo brojem svih učenika. Dakako, to je samo primjer kojim opisujemo pojam srednje vrijednosti.
U teoriji vjerojatnosti taj se pojam definira pomoću raspodjele vjerojatnosti $P(u)$ dane varijable, i to kao umnožak varijable i raspodjele te varijable. Dakle, srednja vrijednost $\overline{u}$ varijable u jednaka je

$$\overline{u} = P(u_1)\,u_1 + P(u_2)\,u_2 + \cdots + P(u_M)\,u_M = \sum_{i=1}^{M} P(u_i)\,u_i \tag{2.3.2}$$

Raspodjela vjerojatnosti $P(u_i)$ opisuje kako su raspoređene vrijednosti varijable u, kojih vrijednosti ima, kojih nema, kojih ima više itd.
Na sličan način možemo definirati i srednju vrijednost bilo koje funkcije diskretne varijable. Neka je f funkcija diskretne varijable u. Tada je srednja vrijednost funkcije $f(u)$

$$\overline{f(u)} \equiv \sum_{i=1}^{M} P(u_i)\, f(u_i) \tag{2.3.3}$$

Srednja vrijednost ima sljedeća svojstva:

(i) neka je c bilo koja konstanta i $f(u)$ funkcija diskretne varijable u, tada vrijedi

$$\overline{c \cdot f(u)} = c \cdot \overline{f(u)} \tag{2.3.4}$$

(ii) neka su $f(u)$ i $g(u)$ bilo koje dvije funkcije od u, tada za srednju vrijednost njihovog zbroja vrijedi

$$\overline{f(u) + g(u)} = \sum_{i=1}^{M} P(u_i)\, [f(u_i)+g(u_i)] = \sum_{i=1}^{M} P(u_i)\, f(u_i)+\sum_{i=1}^{M} P(u_i)\, g(u_i) = \overline{f(u)}+\overline{g(u)} \tag{2.3.5}$$

Dakle, $\overline{u}$ je srednja vrijednost dane varijable, sve ostale vrijednosti varijable odstupaju od $\overline{u}$. To odstupanje označavamo sa Δu i pišemo

$$\Delta u \equiv u - \overline{u} \tag{2.3.6}$$

Srednje odstupanje varijable u od njezine srednje vrijednosti $\overline{u}$, iščezava

$$\overline{\Delta u} = \overline{u - \overline{u}} = \sum_{i=1}^{M} (u_i - \overline{u_i})\, P(u_i) = \sum_{i=1}^{M} P(u_i)\, u_i + \overline{u} \sum_{i=1}^{M} P(u_i) = \overline{u} - \overline{u} = 0$$

Zbog toga se $\overline{u}$ i zove srednja vrijednost, jer predstavlja sredinu skupa $\{u\}$ oko koje su raspoređene sve ostale vrijednosti od u.
Zanima nas kako su te vrijednosti raspoređene oko srednje vrijednosti. Dakle, želimo odrediti neko odstupanje pojedine varijable od njezine srednje vrijednosti.
Kako je iznad pokazano, srednje odstupanje $\overline{\Delta u}$ jednako je nuli, dakle to nam ne pomaže kada želimo odrediti odstupanje varijable u od $\overline{u}$. Ono što nam hoće koristiti, dobijemo ako kvadriramo Δu , te tražimo srednju vrijednost te veličine. Dakle,

$$(\Delta u)^2 = (u - \overline{u})^2$$

Odnosno, srednja vrijednost tog izraza jednaka je

$$\overline{(\Delta u)^2} = \overline{(u-\overline{u})^2} = \overline{u^2 - 2u\overline{u} + \overline{u}^2}$$

Što primjenom izraza (2.3.5) za srednju vrijednost zbroja, odnosno, razlike, postaje

$$\overline{(\Delta u)^2} = \overline{u^2} - 2\overline{u}^2 + \overline{u}^2 = \overline{u^2} - \overline{u}^2 \tag{2.3.7}$$

Prema definiciji za srednju vrijednost, gornji izraz jednak je

$$\overline{(\Delta u)^2} = \overline{(u-\overline{u})^2} = \sum_{i=1}^{M} (u_i - \overline{u})^2 \, P(u_i) \tag{2.3.8}$$

Veličina $\overline{(\Delta u)^2}$ naziva se srednje kvadratno odstupanje varijable u. Zovemo ju još disperzija ili varijanca. Varijanca je drugi moment od u oko njegove srednje vrijednosti i mjera je raspršenosti varijable oko srednje vrijednosti $\overline{u}$, odnosno pokazuje kolika je srednja vrijednost razlike svake pojedine varijable od njezine srednje vrijednosti. Varijanca je uvijek nenegativan broj

$$\overline{(\Delta u)^2} \geq 0$$

Drugi korijen iz varijance zovemo standardna devijacija i označavamo σ. Standardna devijacija jednaka je

$$\sigma = \left[\overline{(\Delta u)^2}\right]^{\frac{1}{2}} = \left[\overline{u^2} - \overline{u}^2\right]^{\frac{1}{2}} \tag{2.3.9}$$

i ona je linearna mjera širine raspona dane varijable.
Često se kaže kako vrijednosti dane varijable fluktuiraju oko njezine srednje vrijednosti, a standardna je devijacija upravo mjera tih fluktuacija.
Mogli bismo sada dalje definirati i treći, četvrti i ostale momente dane varijable. Da bi mogli potpuno odrediti obilježja raspodjele varijable u, trebali bismo poznavati n-ti moment od u. Ali ti se momenti rjeđe koriste, a i za problem nasumičnog hoda dovoljno je poznavanje prvog i drugog momenta raspodjele.

2.4 Računanje srednjih vrijednosti za problem nasumičnog hoda

U prethodnom poglavlju dali smo općenitu definiciju srednje vrijednosti i disperzije dane varijable, odnosno funkcije. Ovdje ćemo te izraze primijeniti na problem nasumičnog hoda. Tako ćemo definirati srednji broj koraka koje čestica napravi, zatim srednji pomak čestice od ishodišta i druge veličine.
Promatramo primjer nasumičnog hoda čestice koja učini ukupno N koraka, od kojih je n_D

koraka u desno, a n_L koraka u lijevo. Vjerojatnost $W(n_D)$ da je n_D koraka u desno, definirana je u poglavlju (2.2) kao

$$W(n_D) = \frac{N!}{n_D!\,(N-n_D)!}\,p^{n_D}\,q^{N-n_D} \tag{2.4.1}$$

Kako bi $W(n_D)$ uopće bila vjerojatnosti varijable n_D, ona mora zadovoljavati uvjet normiranja. Dakle, mora vrijediti

$$\sum_{n_D=0}^{N} W(n_D) = 1 \tag{2.4.2}$$

Uvrstimo u uvjet normiranja izraz (2.4.1) te slijedi

$$\sum_{n_D=0}^{N} W(n_D) = \sum_{n_D=0}^{N} \frac{N!}{n_D!\,(N-n_D)!}\,p^{n_D}\,q^{N-n_D}$$

Gornja je suma prema binomnom teoremu jednaka

$$(p+q)^N$$

S obzirom da su p i q vjerojatnosti, njihov je zbroj jednak 1. Dakle, ispunjen je uvjet normiranja, odnosno $W(n_D)$ je raspodjela vjerojatnosti varijable n_D.
Promatramo česticu koja u svom nasumičnom gibanju učini n_D koraka na desnu stranu, te nas zanima kolika je srednja vrijednost od n_D.
Prema definiciji srednja je vrijednost broja koraka $\overline{n_D}$ jednaka sumi umnoška varijable n_D i njezine funkcije raspodjele $W(n_D)$, ili simbolima

$$\overline{n_D} = \sum_{n_D=0}^{N} W(n_D)\,n_D$$

Što uvrštavanjem izraza za $W(n_D)$ postaje

$$\overline{n_D} = \sum_{n_D=0}^{N} \frac{N!}{n_D!\,(N-n_D)!}\,p^{n_D}\,q^{N-n_D}\,n_D \tag{2.4.3}$$

Kada ne bi bilo dodatnog člana n_D, rješenje sume bilo bi trivijalno; prema binomnom teoremu suma bi bila jednaka jedan. Dakle, problem je kako riješiti taj dodatni član a da suma bude što jednostavnija. To ćemo učiniti uvođenjem sljedeće zamjene

$$n_D\,p^{n_D} = p\,\frac{\partial}{\partial p}\,(p^{n_D})$$

koja mora vrijediti za bilo koje vrijednosti p i q.
Uvrstimo tu zamjenu u izraz (2.4.3), pa slijedi

$$\overline{n_D} = \sum_{n_D=0}^{N} \frac{N!}{n_D!\,(N-n_D)!}\left[p\,\frac{\partial}{\partial p}\,(p^{\,n_D})\right] q^{\,N-n_D}$$

Zamjenom redoslijeda zbrajanja i deriviranja gornji izraz postaje

$$= p\,\frac{\partial}{\partial p}\left[\sum_{n_D=0}^{N} \frac{N!}{n_D!\,(N-n_D)!}\,p^{\,n_D}\,q^{\,N-n_D}\right]$$

Iz čega parcijalnim deriviranjem po p dobivamo izraz za srednji broj koraka u desno n_D

$$\overline{n_D} = p\,N\,(p+q)^{N-1}$$

Kako taj izraz vrijedi za bilo koje vrijednosti p i q, vrijedi dakle i kada je p točno određena konstanta, a $q = 1 - p$.
Prema tome, uvrstimo li sada u prethodni izraz da je $p + q = 1$, dobit ćemo konačno da je srednji broj koraka u desno nasumičnog hoda čestice

$$\overline{n_D} = Np \tag{2.4.4}$$

Dakle, srednji broj koraka u desno ovisi samo o ukupnom broju koraka koji čestica u svom gibanju napravi i vjerojatnosti p da je korak čestice na desnu stranu.
Analogno, možemo zaključiti da je srednji broj koraka u lijevo $\overline{n_L}$ određen ukupnim brojem koraka i vjerojatnosti q da je korak čestice u lijevo. Odnosno,

$$\overline{n_L} = Nq \tag{2.4.5}$$

Očigledno, zbroj srednjih koraka u lijevo i u desno jednak je ukupnom broju koraka koje čestica napravi

$$N = \overline{n_D} + \overline{n_L}$$

Sada kada smo odredili srednju vrijednost broja koraka u desno želimo naći disperziju koraka, odnosno raspršenost vrijednosti oko srednje vrijednosti $\overline{n_D}$.
Disperzija broja koraka u desno jednaka je

$$\overline{(\Delta n_D)^2} = \overline{n_D^{\,2}} - \overline{n_D}^{\,2} \tag{2.4.6}$$

Kako smo srednju vrijednost broja koraka n_D upravo odredili, sada je još potrebno naći koliko je $\overline{n_D^{\,2}}$.

Prema definiciji srednje vrijednosti, srednja vrijednost kvadrata broja koraka $\overline{n_D{}^2}$ jednaka je

$$\overline{n_D{}^2} = \sum_{n_D=0}^{N} W(n_D)\, n_D{}^2$$

Što uvrštavanjem izraza (2.4.1) postaje

$$\overline{n_D{}^2} = \sum_{n_D=0}^{N} \frac{N!}{n_D!\,(N-n_D)!}\, p^{\,n_D}\, q^{\,N-n_D}\, n_D{}^2 \tag{2.4.7}$$

Ponovno, imamo sličnu situaciju kao i kod računanja srednje vrijednosti $\overline{n_D}$, smeta nam član $n_D{}^2$; jer da toga člana nema rješenje sume bilo bi trivijalno. Zato ponovno uvodimo zamjenu

$$p^{\,n_D}\, n_D{}^2 = \left(p\,\frac{\partial}{\partial p}\right)^2 (p^{\,n_D})$$

koja vrijedi za sve vrijednosti p i q.
Uvrštavanjem u (2.4.7), slijedi

$$\overline{n_D{}^2} = \sum_{n_D=0}^{N} \frac{N!}{n_D!\,(N-n_D)!} \left(p\,\frac{\partial}{\partial p}\right)^2 p^{\,n_D}\, q^{\,N-n_D}$$

A to je, nakon zamjene redoslijeda zbrajanja i deriviranja, jednako

$$= \left(p\,\frac{\partial}{\partial p}\right)^2 \sum_{n_D=0}^{N} \frac{N!}{n_D!\,(N-n_D)!}\, p^{\,n_D}\, q^{\,N-n_D}$$

Kako je dana suma prema binomnom teoremu jednaka $(p+q)^N$, gornji izraz postaje

$$= \left(p\,\frac{\partial}{\partial p}\right)^2 (p+q)^N$$

Što parcijalno deriviramo po p te dobijemo

$$= \left(p\,\frac{\partial}{\partial p}\right) \left[p\,N\,(p+q)^{\,N-1}\right] = p\left[N\,(p+q)^{N-1} + p\,N\,(N-1)\,(p+q)^{N-2}\right]$$

Kada uvrstimo da je $p+q=1$, dobit ćemo da je srednja vrijednost kvadrata broja koraka u desno jednaka

$$\overline{n_D{}^2} = p\,[N + p\,N\,(N-1)] = p\,N\,[1 + p\,N - p]$$

Što će uvrštavanjem za $q = 1-p$ te jednostavnim matematičkim manipulacijama postati

$$\overline{n_D{}^2} = (Np)^2 + Npq$$

Dakle, disperzija broja koraka u desno jednaka je

$$\overline{\Delta n_D{}^2} = (Np)^2 + Npq - (Np)^2 = Npq \tag{2.4.8}$$

Prema tome, raspršenost broja koraka oko njihove srednje vrijednosti $\overline{n_D}$ ovisit će samo o ukupnom broju koraka koje čestica u svome gibanju učini i o vjerojatnostima p i q.
Standardna devijacija σ broja koraka u desno jednaka je drugom korijenu disperzije te iznosi

$$\sigma = \sqrt{Npq}$$

Ona opisuje širinu raspona u kojemu su vrijednosti varijable raspršene oko njezine srednje vrijednosti.
Ovdje nas neće toliko zanimati raspršenost varijable oko srednje vrijednosti već relativna širina raspodjele te varijable.
Relativna širina raspodjele zapravo je relativna veličina rasipanja (tj. standardne devijacije) u odnosu na srednju vrijednost varijable. Prema tome, relativna širina raspodjele srednjeg broja koraka u desno jednaka je omjeru standardne devijacije i srednjeg broja koraka u desno

$$\frac{\sigma}{\overline{n_D}} = \frac{\sqrt{Npq}}{Np} = \sqrt{\frac{q}{p}}\frac{1}{\sqrt{N}}$$

Ako je vjerojatnost da je korak koji čestica napravi u desnu stranu jednaka vjerojatnosti da je korak u lijevu stranu, odnosno ako je $p = q = \frac{1}{2}$, tada je relativna širina raspodjele jednaka

$$\frac{\sigma}{\overline{n_D}} = \frac{1}{\sqrt{N}}$$

Dakle, relativna širina raspodjele ovisi samo o ukupnom broju koraka N; kada N raste relativna širina $\sigma/\overline{n_D}$ se smanjuje kao $N^{-\frac{1}{2}}$.

Srednja vrijednost i disperzija mogu se odrediti i za pomak čestice od ishodišta, m.
Pomak čestice od ishodišta, ako ga mjerimo u desno, dan je izrazom $m = n_D - n_L$. A srednja je vrijednost tog pomaka jednaka

$$\overline{m} = \overline{n_D - n_L} = \overline{n_D} - \overline{n_L}$$

Uvrstimo li sada ranije dobivene izraze za srednji broj koraka u desno $\overline{n_D}$, odnosno u lijevo $\overline{n_L}$, srednji je pomak $\overline{m}$ jednak

$$\overline{m} = Np - Nq = N\,(p - q) \tag{2.4.9}$$

Kada je vjerojatnost koraka u desnu i u lijevu stranu jednaka, odnosno $p = q = \frac{1}{2}$, srednji je pomak jednak nuli. Što je zapravo i očekivano jer tada postoji simetrija između lijevog i

desnog smjera, odnosno oba su smjera jednako vjerojatna.
Pomak se može izraziti i samo u ovisnosti o broju koraka u desno, tada je prema (2.2.3)

$$m = 2n_D - N \tag{2.4.10}$$

Odstupanje od srednje vrijednosti pomaka dano je razlikom neke vrijednosti pomaka i njegove srednje vrijednosti, tj.

$$\Delta m = m - \overline{m} \tag{2.4.11}$$

Odnosno,

$$\Delta m = (2n_D - N) - \overline{(2n_D - N)} = 2(n_D - \overline{n_D}) = 2\Delta n_D \tag{2.4.12}$$

Dakle, odstupanje od srednje vrijednosti pomaka ovisi o odstupanju od srednje vrijednosti koraka u desno. Kvadriranjem te uzimanjem srednje vrijednosti gornjeg izraza dobivamo disperziju pomaka od ishodišta, odnosno

$$\overline{(\Delta m)^2} = 4\overline{(\Delta n_D)^2} \tag{2.4.13}$$

Kada u gornju relaciju uvrstimo disperziju broja koraka u desno $\overline{(\Delta n_D)^2}$ danu izrazom (2.4.8) za disperziju pomaka dobivamo

$$\overline{(\Delta m)^2} = 4Npq \tag{2.4.14}$$

Standardna devijacija σ za pomak čestice od ishodišta jednaka je drugom korijenu iz varijance (disperzije)

$$\sigma = \left[\overline{(\Delta m)^2}\right]^{\frac{1}{2}} = 2\sqrt{Npq} \tag{2.4.15}$$

U slučaju kada su vjerojatnosti p i q jednake, tj. $p = q = \frac{1}{2}$, disperzija pomaka jednaka je

$$\overline{(\Delta m)^2} = N$$

Dok je standardna devijacija u tom posebnom slučaju jednaka

$$\sigma = \sqrt{N}$$

Dakle, kada su oba smjera čestice jednako vjerojatna, disperzija pomaka jednaka je ukupnom broju koraka N.

Primjer:

Zadatak je naći srednju vrijednost i disperziju broja koraka i pomaka čestice od ishodišta ako je $N = 100$, a vjerojatnosti p i q su jednake, tj. $p = q = \frac{1}{2}$.
Tada je srednji broj koraka u desno $\overline{n_D}$ jednak

$$\overline{n_D} = Np = 50.$$

Dakle, i srednji broj koraka u lijevo jednak je $\overline{n_L} = 50$.
S obzirom da su oba smjera gibanja čestice jednako vjerojatna, srednji je pomak jednak nula

$$\overline{m} = 0$$

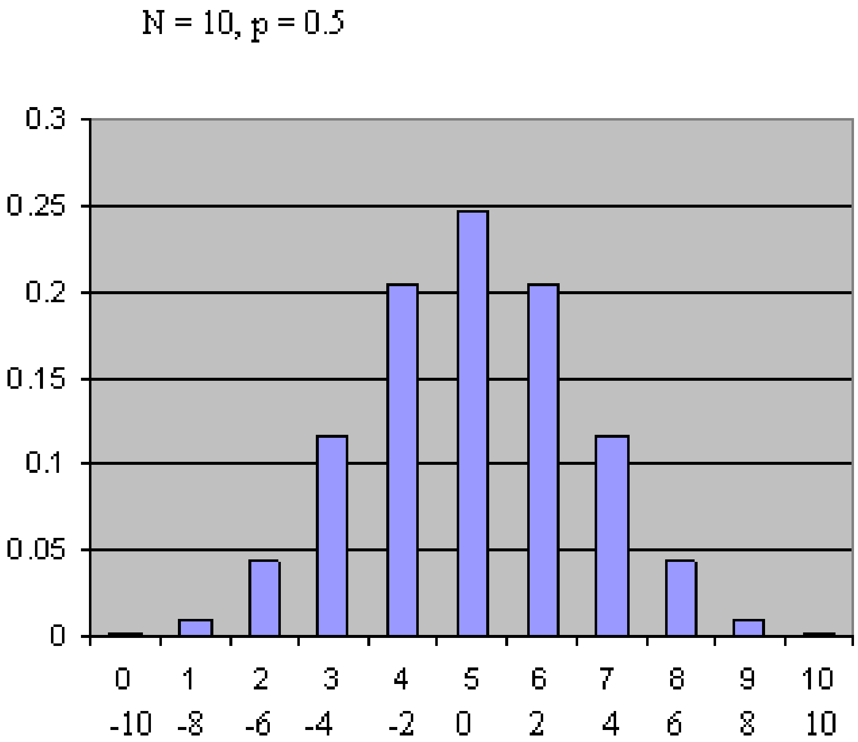
Disperziju broja koraka u desno računamo prema (2.4.8) te je ona

$$\overline{(\Delta n_D)^2} = Npq = 25$$

A disperzija pomaka će prema (2.4.14) biti

$$\overline{(\Delta m)^2} = 4Npq = 100$$

Ako bismo računali standardnu devijaciju, odnosno raspršenost varijabli oko srednje vrijednosti za broj koraka u desno ona bi bila $\sigma_{n_D} = 5$, dok je za pomak $\sigma_m = 100$.



Slika 8: Binomna raspodjela vjerojatnosti za $N = 10$ koraka i $p = 0.5$. U gornjem retku na x-osi prikazan je broj koraka u desno n_D, a u donjem pomak od ishodišta m. Na y-osi prikazana je vjerojatnost $W(n_D)$, odnosno $P_N(m)$

2.5 Raspodjela vjerojatnosti za veliki N

U poglavlju (2.2) definirali smo vjerojatnost $W(n_D)$ da je od ukupno N koraka koje čestica učini, n_D koraka u desno. Ta je vjerojatnost jednaka

$$W(n_D) = \frac{N!}{n_D!\,(N-n_D)!}\,p^{\,n_D}\,q^{\,N-n_D}$$

Dakle, varijabla n_D ima binomnu raspodjelu. Dani izraz možemo primijeniti za računanje raspodjele vjerojatnosti kada N nije velik.
Kod većih vrijednosti broja N pojavljuje se problem računanja faktorijela tako velikih brojeva.
Stoga ćemo ovdje tražiti aproksimaciju za $W(n_D)$, koja će vrijediti i za veliki N.
Promotrimo prvo kako funkcija $W(n_D)$ ovisi o varijabli n_D. Odnosno, zanima nas kada će $W(n_D)$ imati najveću vrijednost. $W(n_D)$ će biti najveća za onu vrijednost varijable najbližu srednjoj vrijednosti; dakle, upravo je srednja vrijednost varijable njezina najvjerojatnija vrijednost.
Zato će dana funkcija imati maksimalnu vrijednost kada je varijabla $n_D = \overline{n_D}$.
Za svaku drugu vrijednost varijable n_D vjerojatnost $W(n_D)$ će se smanjivati.
Promatramo slučaj kada je N dovoljno velik što znači i da je n_D velik. Kada bi se vrijednost od n_D promijenila za neki mali iznos, promjena $W(n_D)$ bila bi razmjerno mala, što se može zapisati kao

$$|W(n_D+1) - W(n_D)| \ll W(n_D) \tag{2.5.1}$$

Dakle, uz dobru aproksimaciju možemo reći da je $W(n_D)$ neprekidna funkcija neprekidne varijable n_D.
Funkcija $W(n_D)$ ima maksimum u točki $n_D = \overline{n_D}$. Zanima nas kako se ta funkcija ponaša u blizini svog maksimuma.
Neka je

$$\eta = n_D - \overline{n_D} \tag{2.5.2}$$

odstupanje varijable n_D od njezine najvjerojatnije vrijednosti $\overline{n_D}$.
Kako bismo istražili ponašanje funkcije $W(n_D)$ u blizini njenog maksimuma, potrebno je tu funkciju razviti u Taylorov red. Zapravo, u Taylorov red razvijamo funkciju $\ln W(n_D)$. Razloga za to ima više: jedan od razloga je to što je $\ln W(n_D)$ mnogo manje promjenjiva funkcija od n_D nego što je to $W(n_D)$; a onda i zbog toga što članovi u razvoju funkcije $\ln W(n_D)$ konvergiraju puno brže od onih u razvoju funkcije $W(n_D)$.

Napomena:
Pokažimo gornje razmatranje na primjeru funkcije

$$f \equiv (1+y)^{-N}$$

koja vrijedi za svaki $y \ll 1$ te za veliki N.
Želimo naći približnu aproksimaciju funkcije f, te da bi to učinili funkciju razvijamo u Taylorov red. Ako bismo u Taylorov red razvili samo funkciju f dobili bi slijedeće

$$f = 1 - Ny + \frac{1}{2}N(N+1)y^2 + \cdots$$

Kada bi sada odredili vrijednost od f za veliki N i $y \ll 1$, vidjeli bi da je čak i za male vrijednosti od y, $Ny \geq 1$.
Dakle, članovi toga razvoja ne konvergiraju.
Zato ćemo prvo definirati prirodni logaritam dane funkcije, odnosno $\ln f$ je jednako

$$\ln f = \ln(1+y)^{-N} = -N\ln(1+y)$$

Razvojem funkcije $\ln f$ u Taylorov red imamo

$$\ln f = -N\left(y - \frac{1}{2}y^2 + \frac{1}{3}y^3 - \cdots\right)$$

Dakle, sama funkcija f jednaka je

$$f = e^{-N(y - \frac{1}{2}y^2 + \cdots)}$$

što vrijedi za svaki $y \leq 1$. Dakle, članovi ovakvoga razvoja konvergiraju.

Slično razmišljanje možemo primijeniti i na danu nam funkciju $W(n_D)$. S obzirom da je ta funkcija neprekidna, kada ju razvijemo u Taylorov red članovi toga reda moraju biti konvergentni. Dakle, to je razlog što ćemo tražiti razvoj funkcije $\ln W(n_D)$ umjesto razvoja funkcije $W(n_D)$.
Prema tome, funkciju $\ln W(n_D)$ napisat ćemo kao sumu reda potencija, odnosno razvit ćemo ju u Taylorov red u okolini točke $\overline{n_D}$. Dakle,

$$\begin{aligned}\ln W(n_D) &= \ln W(\overline{n_D}) + \left.\frac{d\ln W(n_D)}{dn_D}\right|_{n_D=\overline{n_D}}\eta + \frac{1}{2!}\left.\frac{d^2\ln W(n_D)}{dn_D^2}\right|_{n_D=\overline{n_D}}\eta^2 \\ &+ \frac{1}{3!}\left.\frac{d^3\ln W(n_D)}{dn_D^3}\right|_{n_D=\overline{n_D}}\eta^3 + \cdots\end{aligned}$$

Ako sa

$$B_k \equiv \left.\frac{d^k\ln W(n_D)}{dn_D^k}\right|_{n_D=\overline{n_D}}$$

označimo k-tu derivaciju funkcije definirane u točki $n_D = \overline{n_D}$, gornji izraz postaje

$$\ln W(n_D) = \ln W(\overline{n_D}) + B_1\eta + \frac{1}{2!}B_2\eta^2 + \frac{1}{3!}B_3\eta^3 + \cdots \tag{2.5.3}$$

S obzirom da je u točki $n_D = \overline{n_D}$ definiran maksimum funkcije $W(n_D)$, to znači da prva derivacija funkcije u toj točki mora biti jednaka nula i da druga derivacija mora biti manja od nule, odnosno vrijedi

$$\left.\frac{d\ln W(n_D)}{dn_D}\right|_{n_D=\overline{n_D}} = 0$$
$$\left.\frac{d^2\ln W(n_D)}{dn_D^2}\right|_{n_D=\overline{n_D}} < 0$$

Stoga proizlazi, koeficijent B_1 u gornjem izrazu (2.5.3) jednak je nula, dok je B_2 negativan. Zato B_2 zapisujemo pomoću apsolutne vrijednosti kao $B_2 = -|B_2|$.
Kako bi pojednostavili izraz (2.5.3) $W(\overline{n_D})$ zamjenjujemo sa $\widetilde{W}$ pa kada sve uvrstimo dani razvoj od $\ln W(n_D)$ slijedi

$$\ln W(n_D) = \ln \widetilde{W} - \frac{1}{2}|B_2|\,\eta^2 + \frac{1}{6}B_3\eta^3 + \cdots \tag{2.5.4}$$

Iz čega sređivanjem konačno dobijemo funkciju $W(n_D)$

$$W(n_D) = \widetilde{W}\, e^{-\frac{1}{2}|B_2|\eta^2 + \frac{1}{6}B_3\eta^3 + \cdots} \tag{2.5.5}$$

Kada je η dovoljno malen, članovi trećeg reda i viši mogu se zanemariti. Stoga bi aproksimacija funkcije $W(n_D)$ glasila

$$W(n_D) = \widetilde{W}\, e^{-\frac{1}{2}|B_2|\eta^2} \tag{2.5.6}$$

Dakle, dobili smo aproksimaciju raspodjele vjerojatnosti $W(n_D)$ koja je primjenjiva i kada je vrijednost od N velika.

Zanima nas sada za koje se vrijednosti veličine η članovi višeg reda mogu zanemariti. Promotrimo funkciju raspodjele $W(n_D)$ danu relacijom (2.4.1). Logaritmiranje toga izraza daje

$$\ln W(n_D) = \ln \frac{N!}{n_D!\,(N-n_D)!}\, p^{\,n_D}\, q^{\,N-n_D}$$

Što primjenom svojstava logaritamske funkcije postaje

$$\ln W(n_D) = \ln N! - \ln n_D! - \ln(N-n_D)! + n_D\,\ln p + (N-n_D)\,\ln q \tag{2.5.7}$$

Napomena:
Neka je n bilo koji veliki cijeli broj takav da je $n \gg 1$. Ako se vrijednost od n promijeni za

neki mali iznos, tada će i promjena vrijednosti funkcije $\ln n!$ biti mala. Odnosno, možemo reći kako je $\ln n!$ gotovo neprekidna funkcija od n. Stoga vrijedi

$$\frac{d\ln n!}{dn} \approx \frac{\ln(n+1)! - \ln n!}{1} = \ln(n+1)$$

Kako je $n \gg 1$, vrijedi

$$\frac{d\ln n!}{dn} \approx \ln n \tag{2.5.8}$$

Deriviranjem izraza (2.5.7) po n_D dobivamo sljedeće

$$\left.\frac{d\ln W(n_D)}{dn_D}\right|_{n_D=\overline{n_D}} = \frac{d\ln N!}{dn_D} - \frac{d\ln n_D!}{dn_D} - \frac{d\ln(N-n_D)!}{dn_D} + \frac{d(n_D\ln p)}{dn_D} + \frac{d[(N-n_D)\ln q]}{dn_D}$$

S obzirom da su u našem slučaju vrijednosti N i n_D velike, možemo primijeniti rezultat dobiven u prethodnoj napomeni. Dakle,

$$\left.\frac{d\ln W(n_D)}{dn_D}\right|_{n_D=\overline{n_D}} = -\ln n_D + \ln(N-n_D) + \ln p - \ln q = \ln\frac{(N-n_D)\,p}{n_D\,q} \tag{2.5.9}$$

Tražimo maksimum funkcije $W(n_D)$, stoga prva derivacija funkcije u točki maksimuma mora biti jednaka nuli

$$\left.\frac{dW(n_D)}{dn_D}\right|_{n_D=\overline{n_D}} = 0$$

Ili, ekvivalentno, prva derivacija $\ln W(n_D)$ mora biti jednaka nuli

$$\left.\frac{d\ln W(n_D)}{dn_D}\right|_{n_D=\overline{n_D}} = 0 \tag{2.5.10}$$

a derivacije su definirane u točki $n_D = \overline{n_D}$.
Uvrštavanjem izraza (2.5.9) u uvjet (2.5.10) slijedi

$$\ln\frac{(N-\overline{n_D})\;p}{\overline{n_D}\,q} = 0$$

Iz čega proizlazi

$$\frac{(N-\overline{n_D})\,p}{\overline{n_D}\,q} = 1$$

Odnosno,

$$(N-\overline{n_D})\;p = \overline{n_D}\,q$$

Sređivanjem gornjeg izraza imamo

$$\overline{n_D}(p+q) = N\,p$$

Kao što znamo $p+q=1$, te je naposljetku $\overline{n_D}$ jednako

$$\overline{n_D} = N\,p \tag{2.5.11}$$

Isti smo rezultat dobili i u poglavlju (2.4) pri računanju srednjeg broja koraka koje čestica u svom gibanju učini na desnu stranu. Dakle, vidimo da isti izraz vrijedi i kada je broj koraka N koje čestica napravi velik.
Kao što je već navedeno, prva derivacija funkcije $\ln W(n_D)$ u točki $n_D = \overline{n_D}$ jednaka je nuli, stoga je i koeficijent $B_1 = 0$. Želimo sada naći i koeficijente B_2, B_3, B_4 i tako dalje, kako bi utvrdili mogu li se viši članovi Taylorova reda zanemariti.
Tražimo drugu derivaciju funkcije $\ln W(n_D)$ kako bismo dobili koeficijent B_2.
Prema relaciji (2.5.9) druga je derivacija jednaka

$$\left.\frac{d^2 \ln W(n_D)}{dn_D^2}\right|_{n_D=\overline{n_D}} = -\frac{1}{n_D} - \frac{1}{N-n_D} \tag{2.5.12}$$

Ponovno, derivacija je definirana u točki $n_D = \overline{n_D}$, a $\overline{n_D}$ je dan izrazom (2.5.11) pa kada to uvrstimo u gornju relaciju, za koeficijent B_2 imamo

$$B_2 = -\frac{1}{Np} - \frac{1}{N-Np} = -\frac{1}{N}\left(\frac{1}{p} - \frac{1}{1-p}\right)$$

Kako je $q = 1-p$, gornji izraz postaje

$$B_2 = -\frac{1}{N}\left(\frac{1}{p} + \frac{1}{q}\right) = -\frac{q+p}{Npq}$$

Dakle, primjenom već poznate relacije $p+q=1$, konačno je koeficijent B_2 jednak

$$B_2 = -\frac{1}{Npq}$$

S obzirom da N, p i q mogu imati samo pozitivne vrijednosti, očito je da je koeficijent B_2 negativan. Što upravo i mora vrijediti da bi funkcija $W(n_D)$ imala maksimum u točki $\overline{n_D}$. Koeficijent B_2 pišemo pod znakom apsolutne vrijednosti, dakle

$$|B_2| = \frac{1}{Npq} \tag{2.5.13}$$

Nađimo sada koeficijent B_3. Dakle, tražimo treću derivaciju funkcije $\ln W(n_D)$.
Polazeći od izraza (2.5.12) ponovnim deriviranjem po n_D dobije se

$$\left.\frac{d^3 \ln W(n_D)}{dn_D^3}\right|_{n_D=\overline{n_D}} = \frac{1}{n_D^2} - \frac{1}{(N-n_D)^2}$$

I opet je derivacije definirana u točki $n_D = \overline{n_D}$, gdje je $\overline{n_D} = Np$.
Stoga je koeficijent B_3 jednak

$$B_3 = \frac{1}{(Np)^2} - \frac{1}{(N - Np)^2}$$

Iz danog izraza sređivanjem i ponovno iz činjenice da je zbroj vjerojatnosti jednak jedan, odnosno $p + q = 1$, za koeficijent B_3 slijedi

$$B_3 = \frac{q^2 - p^2}{N^2 p^2 q^2}$$

Kako brojnik danog razlomka može biti negativan, potrebno je B_3 pisati sa apsolutnom vrijednosti

$$|B_3| = \frac{|q^2 - p^2|}{N^2 p^2 q^2}$$

Iz činjenice da je $p + q = 1$ proizlazi kako je $|q^2 - p^2|$ uvijek manje od jedan. Stoga možemo zaključiti kako je

$$|B_3| < \frac{1}{N^2 p^2 q^2} \tag{2.5.14}$$

Koeficijent je B_4 jednak četvrtoj derivaciji $\ln W(n_D)$. Dakle,

$$B_4 = \left. \frac{d^4 \ln W(n_D)}{dn_D^4} \right|_{n_D = \overline{n_D}}$$

Dakako, derivaciju ponovno tražimo u točki $n_D = \overline{n_D}$. Koeficijent B_4 ćemo, dakle, dobiti deriviranjem po n_D izraza za B_3. Prema tome, slijedi

$$B_4 = -\frac{2}{n_D^3} - \frac{2}{(N - n_D)^3}$$

Što će, kada se uvrsti da je $n_D = \overline{n_D}$, odnosno, $\overline{n_D} = Np$ postati

$$B_4 = -2\left(\frac{1}{N^3 p^3} + \frac{1}{N^3 q^3}\right) = -2\frac{(q^3 + p^3)}{N^3 p^3 q^3}$$

Ponovno, dani izraz zbog negativnog predznaka pišemo pod apsolutnom vrijednosti

$$|B_4| = \frac{2(p^3 + q^3)}{N^3 p^3 q^3}$$

I ponovno, zbog činjenice da je $p + q = 1$, brojnik gornjeg razlomka mora biti manji od 4. Dakle, vrijedi

$$|B_4| < \frac{4}{N^3 p^3 q^3} \tag{2.5.15}$$

Stoga, na temelju dobivenih rezutata (2.5.14) i (2.5.15) možemo zaključiti kako za k-ti član izraza (2.5.4) vrijedi

$$\frac{1}{k!} B_k \cdot \eta^k \sim \frac{\eta^k}{(Npq)^{k-1}} < 1$$

Dakle, zanemarivanje članova višeg reda u Taylorovom razvoju opravdano je kada je η dovoljno malen tako da vrijedi

$$\eta \ll Npq \tag{2.5.16}$$

Jer je tada vrijednost članova višeg reda (višeg od drugog reda) zanemarivo malena u odnosu na prethodne članove. Obzirom da smo zanemarili članove reda većeg od dva, funkcija $W(n_D)$ imat će oblik

$$W(n_D) = \widetilde{W}\, e^{-\frac{1}{2}|B_2|\eta^2} \tag{2.5.17}$$

Očigledno, vrijednost izraza (2.5.17) ovisi o eksponencijalnom članu $e^{-\frac{1}{2}|B_2|\eta^2}$.
Kada je B_2 velik i kada $|\eta|$ raste tako da vrijedi

$$|B_2|\, \eta^2 = \frac{\eta^2}{Npq} \gg 1$$

odnosno, kada je vrijednost eksponenta puno veća od 1, vrijednost će funkcije padati vrlo brzo. Stoga smo dobili još jedan uvjet za vrijednost veličine η

$$\eta^2 \gg Npq \qquad \Rightarrow \qquad \eta \gg \sqrt{Npq} \tag{2.5.18}$$

Prema tome, zaključujemo, na čitavom području gdje W ima značajne vrijednosti (dakle, gdje vrijednost W nije zanemarivo mala), te ako η zadovoljava uvjete (2.5.16) i (2.5.18), odnosno ako je

$$\sqrt{Npq} \ll \eta \ll Npq \tag{2.5.19}$$

možemo reći da je izraz (2.5.17) dobra aproksimacija izraza (2.4.1) za veliki N.
Dana aproksimacija vrijedi kada je $Npq \gg 1$, odnosno, kada je vrijednost od N velika i kada p i q nisu premali.
U slučaju kada je $p \ll 1$ i $q \ll 1$, funkciju W aproksimiramo Poissonovom raspodjelom.

Nakon što smo pronašli vrijednost η za koju možemo zanemariti članove trećeg i viših redova u aproksimaciji funkcije W, da bismo tu aproksimaciju potpuno odredili potrebno je još pronaći vrijednost veličine $\widetilde{W}$.
Vrijednost te veličine odrediti odredit ćemo iz uvjeta normiranja funkcije $W(n_D)$.

Kako bi dana funkcija predstavljala vjerojatnost varijable n_D, ona mora ispunjavati uvjet normiranja. Mora vrijediti

$$\sum_{n_D=0}^{N} W(n_D) = 1$$

Općenito, kada govorimo o neprekidnoj varijabli, u uvjetu normiranja (a tako i u svim ostalim izračunima), zbroj po svim vrijednostima varijable možemo zamijeniti integralom.
S obzirom da smo na početku ovoga poglavlja, već pretpostavili da je $W(n_D)$ neprekidna funkcija neprekidne varijable n_D, uvjet normiranja dane funkcije bit će

$$\int W(n_D) dn_D = 1$$

Kako je $n_D = \overline{n_D} + \eta$, gornji je integral približno jednak

$$\int W(n_D) dn_D \approx \int W(\overline{n_D} + \eta) d\eta = 1. \tag{2.5.20}$$

Ako je odstupanje $|\eta| = |n_D - \overline{n_D}|$ od srednje vrijednosti varijable dovoljno veliko, tako da je vrijednost W daleko od njezinog maksimuma, tada integrand ne doprinosi značajno vrijednosti integrala. Prema tome, područje integracije po η može se proširiti od $-\infty$ do ∞.
Uvrstimo sada u uvjet normiranja izraz (2.5.17), odnosno, funkciju $W(n_D)$

$$\int_{-\infty}^{\infty} W(n_D) dn_D = \int_{-\infty}^{\infty} \widetilde{W} \, e^{-\frac{1}{2}|B_2|\eta^2} d\eta = 1$$

Što kada izračunamo dani integral postaje

$$\widetilde{W} \sqrt{\frac{2\pi}{|B_2|}} = 1$$

Iz čega je $\widetilde{W}$ jednako

$$\widetilde{W} = \sqrt{\frac{|B_2|}{2\pi}}$$

Prema tome, aproksimacija funkcije $W(n_D)$ za veliki N jednaka je

$$W(n_D) = \sqrt{\frac{|B_2|}{2\pi}} \, e^{-\frac{1}{2}|B_2|\eta^2} \tag{2.5.21}$$

Izraz (2.5.21) naziva se Gaussova raspodjela i zbog svoje se općenitosti vrlo često primjenjuje u statistici. O Gaussovoj ćemo raspodjeli reći nešto više u idućem poglavlju.
Dobivenu aproksimaciju možemo zapisati i u drugom obliku tako što primijenimo neke od

izraza koje smo ranije dobili, kao npr. za srednju vrijednost, disperziju i slično. Naime, ako umjesto $\overline{n_D}$ i $|B_2|$ uvrstimo redom relacije (2.5.11) i (2.5.13), raspodjela će imati oblik

$$W(n_D) = (2\pi Npq)^{-\frac{1}{2}}\ e^{-\left[\frac{(n_D - Np)^2}{2Npq}\right]}$$

Aproksimaciju možemo zapisati i u terminima srednje vrijednosti $\overline{n_D}$ i disperzije $\overline{(\Delta n_D)^2}$, dakle

$$W(n_D) = \left(2\pi\overline{(\Delta n_D)^2}\right)^{-\frac{1}{2}}\ e^{-\left[\frac{(n_D - \overline{n_D})^2}{2\overline{(\Delta n_D)^2}}\right]}$$

U kojem god obliku zapisali Gaussovu raspodjelu, bitno je reći kako je to vrlo općenita i najčešće primjenjiva raspodjela u teoriji vjerojatnosti i statistici, ali također i u prirodi jer su mnoge stvari i pojave u prirodi podložne upravo Gaussovoj raspodjeli.

2.6 Gaussova raspodjela vjerojatnosti

Dosada smo o problemu nasumičnog hoda govorili u smislu diskretne varijable. Dakle, dani broj koraka ili pomak čestice bili su diskretne varijable. Stoga je i raspodjela takvih varijabli diskretna; štoviše, raspodjela je danih varijabli, poznato je, binomna raspodjela.
Ali što kada varijabla koju promatramo i za koju tražimo vjerojatnost, nije diskretna nego kontinuirana? Tada će i raspodjela te varijable morati biti kontinuirana.
Kao što smo u prethodnom poglavlju utvrdili, vjerojatnost kontinuirane varijable dana je Gaussovom raspodjelom.
Vjerojatnost kontinuirane varijable nije uvijek dana Gaussovom raspodjelom, mogu to biti i neke druge raspodjele, Poissonova ili neka druga. Ali u našem slučaju, kada je broj koraka N čestice velik, te kada vjerojatnosti p i q nisu male, koristit će nam upravo Gaussova raspodjela.
Dakle, ponovno se razmatra problem nasumičnog hoda u jednoj dimenzij: neka čestica koja se giba napravi ukupno N koraka od koji je n_D koraka u desnu stranu, a n_L koraka u lijevu stranu. I opet je pomak čestice, uz pretpostavku da svi koraci jednake duljine l i pomak mjerimo u desno jednak

$$m = 2n_D - N$$

gdje je $-N \leq m \leq N$, takav da se njegove susjedne vrijednosti razlikuju za $\Delta m = 2$.
Prema Gaussovoj raspodjeli vjerojatnost $W(n_D)$ broja koraka u desno jednaka je

$$W(n_D) = (2\pi Npq)^{-\frac{1}{2}}\ e^{-\left[\frac{(n_D - Np)^2}{2Npq}\right]}$$

Vjerojatnost broja koraka u desno $W(n_D)$ jednaka je vjerojatnosti pomaka čestice $P(m)$. Prema tome, ako broj koraka u desno n_D izrazimo preko pomaka m, dakle $n_D = \frac{1}{2}(N+m)$, vjerojatnost pomaka čestice $P(m)$ bit će

$$P(m) = W\left(\frac{N+m}{2}\right) = [2\pi Npq]^{-\frac{1}{2}}\ e^{-\left\{\frac{[m-N(p-q)]^2}{8Npq}\right]} \tag{2.6.1}$$

Položaj, x, čestice ovisi o pomaku m, odnosno vrijedi da je $x = ml$, gdje je l duljina koraka koji čestica napravi, koja je u ovom slučaju jednaka za svaki korak.
Zanima nas koji će biti položaj čestice u odnosu na njezino ishodište, odnosno kolika je vjerojatnost da se čestica nakon N koraka nađe u intervalu $(x, x+dx)$.
Karakteristična duljina koraka l ovisit će o sustavu koji promatramo; tako će u nekom mikroskopskom sustavu karakteristična duljina od l biti reda veličine 10^{-8} cm, dok će u makroskopskim sustavima biti reda veličine 10^{-4} cm. Ako promatramo makroskopski sustav tada je duljina koraka l zanemarivo mala u odnosu na najmanje područje koje razmatramo, te stoga diskretne vrijednosti pomaka za $2l$ možemo zanemariti. Odnosno, možemo pretpostaviti kako je promatrano područje kontinuirano.
Također, za velike vrijednosti broja N, promjena pomaka za $\Delta m = 2$ neće značajno utjecati na vjerojatnost $P(m)$ te možemo reći da vrijedi

$$|P(m+2) - P(m)| \ll P(m)$$

Odnosno, $P(m)$ se može smatrati glatkom funkcijom od x. Prema tome, iz prethodnog razmatranja proizlazi kako, u makroskopskom smislu x možemo smatrati kontinuiranom varijablom te tražiti vjerojatnost da je položaj čestice nakon N koraka između x i $x+dx$.
Diferencijal dx područja koje promatramo mora biti manji od karakteristične vrijednosti duljine koraka L u makroskopskom sustavu, ali veći od karakteristične vrijednosti duljine koraka l u mikroskopskom sustavu. Dakle

$$l \ll dx \ll L$$

U intervalu $(x, x+dx)$ nalazi se $\frac{dx}{2l}$ mogućih vrijednosti pomaka m. Razlog tomu je što se svaka vrijednost pomaka od susjedne razlikuje upravo za 2. Vjerojatnost $P(m)$ je približno jednaka za svaku vrijednost iz područja dx te se stoga vjerojatnost da je položaj čestice u intervalu $(x, x+dx)$ dobije množenjem vjerojatnosti $P(m)$ ukupnim brojem vrijednosti pomaka koje interval sadrži

$$\rho(x)dx = P(m)\,\frac{dx}{2l} \tag{2.6.2}$$

Ovdje se uvodi nova veličina $\rho(x)$, zovemo ju gustoća vjerojatnosti, i definiramo kada je promatrana varijabla kontinuirana. Kao što je vidljivo iz gornjeg izraza gustoća vjerojatnosti

ne ovisi o diferencijalu područja dx.
Općenito, vjerojatnost diskretne varijable označavamo sa $P(x)$ a uvjet normiranja i srednje vrijednosti računamo zbrajanjem po svim vrijednostima varijable, dok se kod kontinuirane varijable vjerojatnost izražava pomoću gustoće vjerojatnosti $\rho(x)$ a zbrojevi po vrijednostima varijable zamjenjuju integralima.
Uvrstimo sada u gornji relaciju izraz (2.6.1) pa za vjerojatnost slijedi

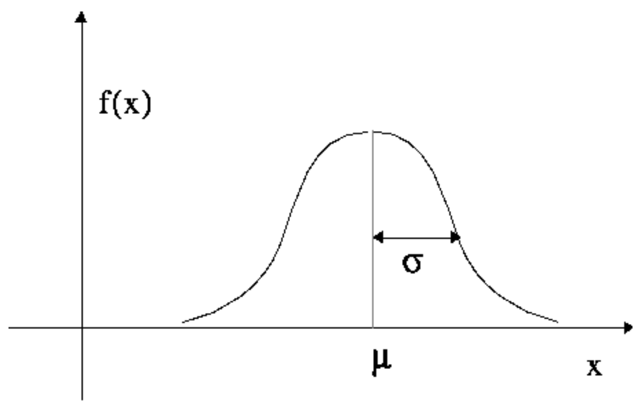
$$\rho(x)dx = [2\pi Np\,q]^{-\frac{1}{2}}\; e^{-\left\{\frac{[m-N(p-q)]^2}{8Npq}\right\}}\,\frac{dx}{2l}$$

Što će uvođenjem zamjena

$$\begin{aligned}\mu &\equiv (p-q)\,Nl\\ \sigma &\equiv 2\sqrt{Npq}\,l\end{aligned}$$

postati

$$\rho(x)dx = \frac{1}{\sqrt{2\pi}\,\sigma}\; e^{-(x-\mu)^2/2\sigma^2}\,dx \tag{2.6.3}$$



Slika 9: Gaussova (normalna) raspodjel vjerojatnosti. Na slici je sa μ označena srednja vrijednost varijable x, a sa σ standardna devijacija

Gornji se izraz naziva Gaussova raspodjela vjerojatnosti i jedna je od najvažnijih raspodjela u teoriji vjerojatnosti i statistici.
Primjenjujući Gaussovu raspodjelu možemo izračunati srednju vrijednost $\overline{x}$ i disperziju $\overline{(\Delta x)^2}$ varijable.
Kako je varijabla x kontinuirana, pri računanju ćemo zbroj zamijeniti integralom. Granice u kojima integriramo su od $-\infty$ do ∞ zato jer gustoća vjerojatnosti $\rho(x)$ postaje zanemarivo

mala kada je $|x|$ toliko velik da je pomak nedostižan u N koraka.
Dakle, uvjet normiranja kada je dana varijabla kontinuirana glasi

$$\int_{-\infty}^{\infty} \rho(x)dx = 1$$

Što, kada za $\rho(x)$ uvrstimo relaciju (2.6.3), postaje

$$\int_{-\infty}^{\infty} \rho(x)dx = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{\infty} e^{-(x-\mu)^2/2\sigma^2}\, dx$$

Da bi izračunali integral uvodimo zamjenu $y \equiv x - \mu$. Dakle,

$$\int_{-\infty}^{\infty} \rho(x)dx = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{\infty} e^{-y^2/2\sigma^2} dy = \frac{1}{\sqrt{2\pi}\,\sigma}\ \sqrt{2\pi}\,\sigma = 1$$

Prema tome, zaključujemo da je $\rho(x)dx$ pravilno normirana te možemo računati srednju vrijednost i disperziju. Općenito, srednja je vrijednost diskretne varijable dana izrazom

$$\overline{u} = \sum_{i=1}^{M} P(u_i)u_i$$

Kao što znamo, kada je varijabla kontinuirana sa zbroja prelazimo na integral, a raspodjelu vjerojatnosti zamjenjujemo gustoćom vjerojatnosti. Dakle, srednja vrijednost varijable x jednaka je

$$\overline{x} \equiv \int_{-\infty}^{\infty} x\rho(x)dx \tag{2.6.4}$$

Uvrštavanjem izraza (2.6.3) to postaje

$$\overline{x} = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{\infty} x\, e^{-(x-\mu)^2/2\sigma^2} dx$$

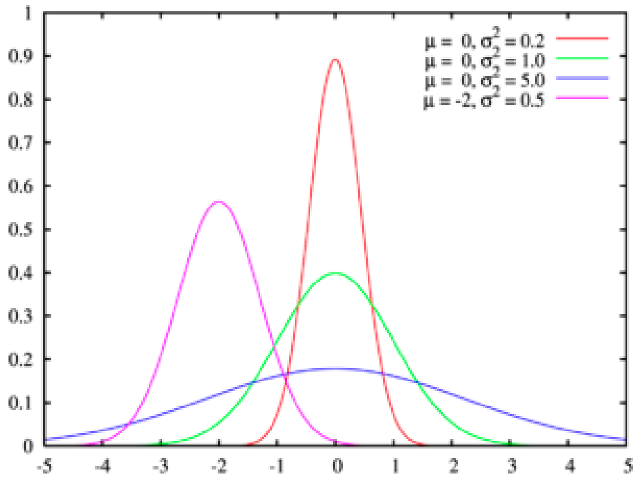
Ponovno, uvodimo zamjenu $y \equiv x - \mu$ te integral postaje

$$\overline{x} = \frac{1}{\sqrt{2\pi}\,\sigma} \left[\int_{-\infty}^{\infty} y\, e^{-y^2/2\sigma^2} dy + \mu \int_{-\infty}^{\infty} e^{-y^2/2\sigma^2} dy\right]$$

Integrand prvog integrala neparna je funkcija, pa je taj integral jednak nuli. Drugi je integral jednak jedan, kao što smo ranije izračunali uvjetom normiranja. Dakle, srednja vrijednost varijable x jednaka je

$$\overline{x} = \mu \tag{2.6.5}$$

gdje je $\mu \equiv (p - q)\, Nl$.



Slika 10: Gaussova raspodjela vjerojatnosti za različite srednje vrijednosti μ i disperzije σ^2

Ovaj smo rezultat mogli i pretpostaviti jer je gustoća vjerojatnosti $\rho(x)$ samo funkcija od $|x-\mu|$ te je stoga simetrična s obzirom na svoj maksimum, koji je upravo u točki $x = \mu$.Dakle, μ je srednja vrijednost od x.

Slično razmatranje možemo dati i za disperziju varijable, $\overline{(\Delta x)^2}$. Kao i kod računanja srednje vrijednosti varijable, i u izračunu disperzije kontinuirane varijable, sa zbroja prelazimo na integral. Prema tome, disperzija od x jednaka je

$$\overline{(\Delta x)^2} = \overline{(x-\overline{x})^2} = \int_{-\infty}^{\infty} (x-\overline{x})^2 \rho(x) dx \tag{2.6.6}$$

Kako je $\overline{x} = \mu$, slijedi

$$\overline{(x-\mu)^2} = \int_{-\infty}^{\infty} (x-\mu)^2 \rho(x) dx \tag{2.6.7}$$

Što, kada se uvrsti ranije dobivena relacija za $\rho(x)dx$, postaje

$$\overline{(x-\mu)^2} = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{\infty} y^2\, e^{-y^2/2\sigma^2} dy$$

Da bismo disperziju mogli izračunati, treba riješiti integral $\int_{-\infty}^{\infty} y^2\, e^{-y^2/2\sigma^2} dy$.
Integral oblika $\int_0^{\infty} x^2\, e^{-a^2x^2}$ za $a > 0$, nam je poznat i njegovo je rješenje prema Bronštejn-Semendjajev jednako $\frac{\sqrt{\pi}}{4a^3}$.
U našem je slučaju konstanata $a = \frac{1}{\sqrt{2}\,\sigma}$, također, integrand je parna funkcija te vrijedi

$$\int_{-\infty}^{\infty} y^2\, e^{-y^2/2\sigma^2} dy = 2 \int_{0}^{\infty} y^2\, e^{-y^2/2\sigma^2} dy$$

Prema tome, za rješenje danog integrala dobivamo

$$\int_{-\infty}^{\infty} y^2\, e^{-y^2/2\sigma^2} dy = 2\, \frac{\sqrt{\pi}}{4}\, \left(\sqrt{\pi}\,\sigma\right)^3$$

Dakle, nakon uvršavanja u izraz (2.6.7) disperzija varijable x jednaka je

$$\overline{(x-\mu)^2} = \sigma^2 \tag{2.6.8}$$

Dobiveni rezultat je očito valjan jer je standardna devijacija σ jednaka drugom korijenu iz disperzije. Kako je standardna devijacija jednaka

$$\sigma \equiv 2\sqrt{N p\, q} l$$

proizlazi da je disperzija varijable x jednaka

$$\overline{(\Delta x)^2} = 4\, N\, p\, q\, l^2$$

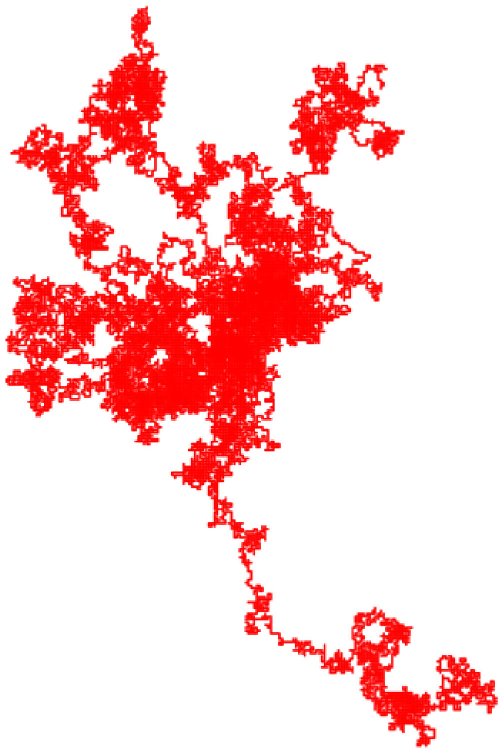
Prema tome je srednja vrijednost i disperzija položaja čestice od ishodišta

$$\begin{aligned} \overline{x} &= N\,(p-q)\,l \\ \overline{(\Delta x)^2} &= 4\, N\, p\, q\, l^2 \end{aligned}$$

Stoga zaključujemo da se dobiveni izrazi za srednju vrijednost i disperziju varijable x, izvedeni za veliki N, slažu s onima dobivenim u poglavlju (2.4).

3 POOPĆENJE PROBLEMA NASUMIČNOG HODA

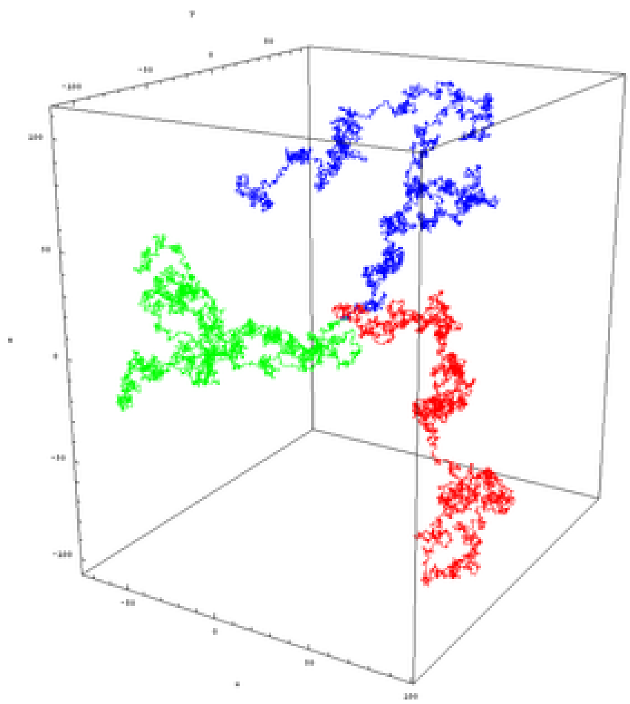
U prethodnim smo se poglavljima bavili problemom nasumičnog hoda u jednoj dimenziji te smo za taj slučaj odredili neke osnovne pojmove i karakteristike; tako smo, na primjer, računali vjerojatnost pronalaženja čestice nakon određenog broja koraka koje ta čestica učini, zatim smo određivali njezin pomak od ishodišta i slično.
Pri tome smo, radi jednostavnosti i lakšeg računanja, pretpostavili kako se dana čestica giba u jednoj dimenziji i kako je svaki njezin korak jednake duljine. Te pretpostavke ograničile su primjenu dobivenih rezultata tako da se oni ne mogu koristiti u slučaju kada, na primjer, koraci čestice nisu iste duljine ili kada se čestica giba u više dimenzija. Upravo ćemo o tome govoriti u narednim poglavljima. Dakle, problem nasumičnog hoda poopćit ćemo tako da dobiveni izrazi imaju širu primjenu.



Slika 11: Nasumični hod u dvije dimenzije s većim brojem manjih koraka. Za vrlo malu duljinu koraka dobije se Brownovo gibanje

3.1 Raspodjela vjerojatnosti više varijabli

U dosadašnjem smo se razmatranju bavili raspodjelom vjerojatnosti samo jedne varijable, bilo diskretne ili kontinuirane, i svojstvima te raspodjele. U ovom ćemo poglavlju definirati raspodjelu vjerojatnosti više varijabli. Iako ćemo ovdje razmatrati slučaj dvije varijable, dobiveni se rezultati izravno mogu poopćiti na tri i više varijabli.



Slika 12: Slika prikazuje tri nasumična hoda u tri dimenzije

Dakle, promatramo diskretne slučajne varijable u i v koje mogu poprimiti vrijednosti

$$\begin{aligned} & u_i \quad za \;\; i = 1, 2, \ldots, M \\ & v_j \quad za \;\; j = 1, 2, \ldots, N \end{aligned}$$

Tada je $P(u_i, v_j)$ vjerojatnost da u poprima vrijednost u_i i da v poprima vrijednost v_j. Prema tome, ona mora ispunjavati uvjet normiranja. Odnosno, mora biti

$$\sum_{i=1}^{M} \sum_{j=1}^{N} P(u_i, v_j) = 1 \tag{3.1.1}$$

Ako su varijable u i v statistički neovisne, odnosno, vjerojatnost da jedna varijabla poprima neku vrijednost ne ovisi o vjerojatnosti druge varijable, onda vjerojatnost $P(u_i, v_j)$ možemo zapisati kao umnožak vjerojatnosti pojedine varijable. Dakle, vrijedi

$$P(u_i, v_j) = P_u(u_i)\, P_v(v_j) \tag{3.1.2}$$

Gdje $P_u(u_i)$ predstavlja vjerojatnost da varijabla u poprima vrijednosti od u_i, i neovisna je o vjerojatnosti da v poprima vrijednosti od v_j. Što zapisujemo kao

$$P_u(u_i) = \sum_{j=1}^{N} P(u_i, v_j) \tag{3.1.3}$$

Isto tako, $P_v(v_j)$ je vjerojatnost da varijabla v poprima vrijednosti od v_j, neovisno o vjerojatnosti varijable u. Ta je vjerojatnost jednaka

$$P_v(v_j) = \sum_{i=1}^{M} P(u_i, v_j) \tag{3.1.4}$$

Provjerimo vrijedi li za vjerojatnosti $P_u(u_i)$ i $P_v(v_j)$ uvjet normiranja. Dakle, za vjerojatnost $P_u(u_i)$ imamo sljedeće

$$\sum_{i=1}^{M} P_u(u_i) = \sum_{i=1}^{M} \left[\sum_{j=1}^{N} P(u_i, v_j) \right] = \sum_{i=1}^{M} \sum_{j=1}^{N} P(u_i, v_j) = 1$$

A za $P_v(v_j)$

$$\sum_{j=1}^{N} P_v(v_j) = \sum_{j=1}^{N} \left[\sum_{i=1}^{M} P(u_i, v_j) \right] = \sum_{j=1}^{N} \sum_{i=1}^{M} P(u_i, v_j) = 1$$

Dakle, zbog uvjeta normiranja (3.1.1) slijedi da su obje gornje sume jednake jedan, odnosno vjerojatnosti $P_u(u_i)$ i $P_v(v_j)$ pravilno su normirane.
U poglavlju (2.3) govorili smo o srednjoj vrijednosti funkcije jedne varijable i njenim osnovnim svojstvima. Uvjerimo se da ta svojstva vrijede i kada je dana funkcija više varijabli. Dakle, neka je $F(u, v)$ bilo koja funkcija varijabli u i v. Tada je srednja vrijednost $\overline{F(u, v)}$ dane funkcije jednaka

$$\overline{F(u, v)} \equiv \sum_{i=1}^{M} \sum_{j=1}^{N} P(u_i, v_j)\, F(u_i, v_j) \tag{3.1.5}$$

Nadalje, neka su $F(u, v)$ i $G(u, v)$ funkcije dvije varijable u i v, može se pokazati da za srednju vrijednost njihova zbroja vrijedi

$$\begin{aligned} \overline{F + G} &\equiv \sum_{i} \sum_{j} P(u_i, v_j)\, [F(u_i, v_j) + G(u_i, v_j)] \\ &= \sum_{i} \sum_{j} P(u_i, v_j)\, F(u_i, v_j) + \sum_{i} \sum_{j} P(u_i, v_j) G(u_i, v_j) \end{aligned}$$

što primjenom izraza (3.1.5) za srednju vrijednost funkcije dvije varijable postaje

$$\overline{F + G} = \overline{F} + \overline{G} \tag{3.1.6}$$

Dakle, srednja vrijednost zbroja funkcija dvije varijable jednaka je zbroju srednjih vrijednosti tih funkcija. Isti smo rezultat dobili i u poglavlju (2.3) kada su dane funkcije bile funkcije jedne varijable.
Na sličan način može se definirati i srednja vrijednost statistički neovisnih varijabli. Ako je $f(u)$ samo funkcija varijable u, njena je srednja vrijednost jednaka

$$\overline{f(u)} \equiv \sum_i \sum_j P(u_i, v_j)\, f(u_i)$$

Što, kada primjenimo izraz (3.1.3) za vjerojatnost statistički neovisnih varijabli, postane

$$\overline{f(u)} \equiv \sum_i P_u(u_i)\, f(u_i) \tag{3.1.7}$$

Ako je $f(u)$ funkcija varijable u, a $g(v)$ funkcija varijable v te ako su varijable statistički neovisne, možemo definirati srednju vrijednost umnoška tih varijabli. Dakle,

$$\overline{f(u) \cdot g(v)} \equiv \sum_i \sum_j P(u_i, v_j)\, f(u_i)\, g(v_j)$$

Vjerojatnost $P(u_i, v_j)$ je prema izrazu (3.1.2) jednaka je umnošku vjerojatnosti svake pojedine varijable, kada su one statistički neovisne. Pa gornji izraz postaje

$$\overline{f(u) \cdot g(v)} = \sum_i \sum_j P_u(u_i)\, P_v(v_j)\, f(u_i)\, g(v_j) = \left[\sum_i P_u(u_i)\, f(u_i)\right] \cdot \left[\sum_j P_v(v_j)\, g(v_j)\right]$$

Dakle, srednja vrijednost umnoška funkcija dvije varijable jednaka je umnošku srednjih vrijednosti tih funkcija, ili simbolima

$$\overline{f(u) \cdot g(v)} = \overline{f(u)} \cdot \overline{g(v)} \tag{3.1.8}$$

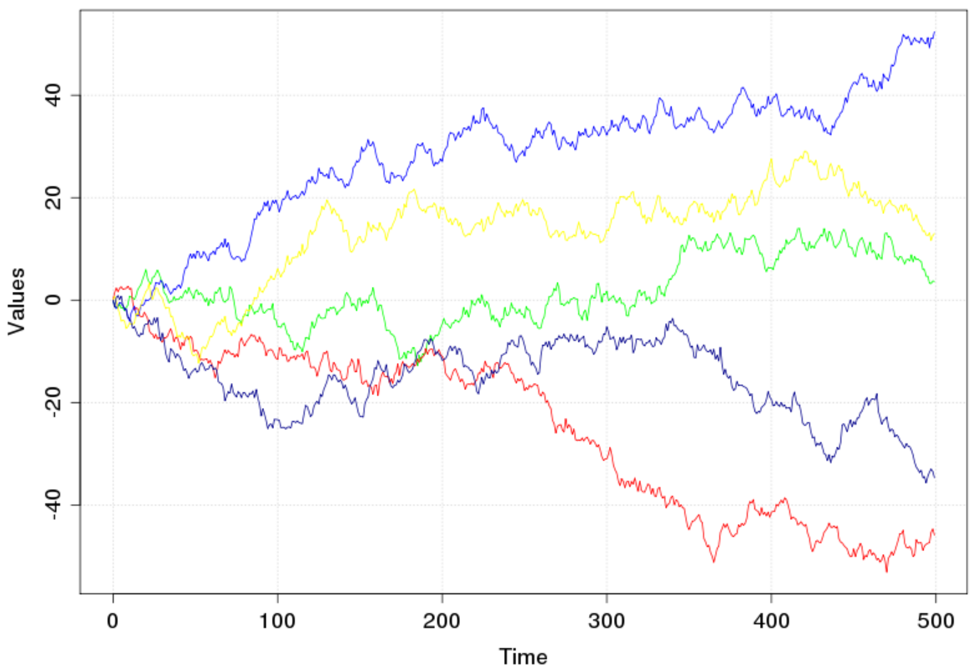
Ako varijable u i v nisu statistički neovisne, gornji izraz neće vrijediti.
U ovom smo poglavlju definirali vjerojatnost i srednju vrijednost funkcije dvije varijable, ali svi dobiveni izrazi mogu se izravno poopćiti i na više varijabli.

3.2 Kontinuirana raspodjela vjerojatnosti

Promatramo neprekidnu varijablu u koja poprima bilo koju vrijednost iz neprekidnog područja

$$a_1 < u < a_2$$

Da bismo za ovu situaciju definirali vjerojatnost, promatramo bilo koje infinitezimalno područje varijable između u i $u + du$ te tražimo vjerojatnost da varijabla poprima vrijednosti iz toga područja.



Slika 13: Nasumični hod u jednoj dimenziji kada je pomak kontinuirana varijabla

Ako je dani diferencijal du dovoljno malen, tada je vjerojatnost proporcionalna veličini du te ju možemo zapisati kao $\rho(u)du$, gdje je $\rho(u)$ gustoća vjerojatnosti koja ne ovisi o veličini diferencijala du.
Zapravo, vjerojatnost je dana razvojem u Taylorov red po potencijama od du kada taj $du \to 0$. Kako je vrijednost diferencijala du vrlo mala, to se viši članovi razvoja mogu zanemariti te nam ostane samo član $\rho(u)du$.
Kako smo se do sada uglavnom bavili diskretnim varijablama i njihovom vjerojatnosti, postavlja se pitanje kako ćemo ovdje naći vjerojatnost varijable u kada je ona kontinuirana.
Problem ćemo riješiti na sljedeći način: promatrano neprekidno područje $a_1 < u < a_2$ na kojemu je definirana varijabla u podijelit ćemo na proizvoljno malene intervale jednake veličine δu.
Vrijednost varijable u u pojedinom intervalu označit ćemo sa u_i, a vjerojatnost pronalaženja varijable u tom intervalu sa $P(u_i)$. Na taj smo način dobili prebrojiv skup vrijednosti varijable u, a svaka od tih vrijednosti odgovara jednom od fiksnih infinitezimalnih intervala.
Dakle, problem pronalaženja vjerojatnosti kontinuirane varijable sveli smo na pronalaženje vjerojatnosti diskretne varijable.
Tada je vjerojatnost $\rho(u)du$, da u poprima bilo koju vrijednost iz područja $(u, u + du)$, s

obzirom da smo to područje podjelili na intervale δu, jednaka

$$\rho(u)\delta u$$

Interval $(u, u+du)$ upravo zbog te podjele sadži

$$\frac{du}{\delta u}$$

mogućih diskretnih vrijednosti varijable u; pri čemu je vjerojatnost $P(u_i)$ svake te vrijednosti jednaka te ju možemo označiti samo sa $P(u)$.
Prema tome, vjerojatnost $\rho(u)du$ da varijabla poprima vrijednosti iz intervala $(u, u+du)$ jednaka je umnošku diskretnih vrijednosti varijable koje taj interval sadrži i vjerojatnosti $P(u)$ svake te vrijednosti, odnosno

$$\rho(u)du = P(u)\,\frac{du}{\delta u}$$

I ovdje mora vrijediti uvjet normiranja, odnosno zbroj vjerojatnosti po svim mogućim vrijednostima varijable mora biti jednak jedan

$$\sum_i P(u_i) = 1$$

Kako je u našem slučaju varijabla u kontinuirana, zbroj možemo zamijeniti integralom.
Stoga ćemo prvo provesti zbrajanje po svim mogućim vrijednostima varijable u promatranom intervalu $(u, u+du)$ te tako dobivenu vrijednost $\rho(u)du$ integrirati po čitavom neprekidnom području (a_1, a_2), odnosno integrirati preko svih takvih intervala du.
Prema tome, uvjet normiranja glasi

$$\int_{a_1}^{a_2} \rho(u)du = 1 \tag{3.2.1}$$

Također, možemo definirati i srednju vrijednost funkcije kontinuirane varijable.
Podsjetimo se, u poglavlju (2.3) definirali smo srednju vrijednost funkcije $f(u)$ kada je u diskretna varijabla kao

$$\overline{f(u)} = \sum_i P(u_i)\,f(u_i)$$

Kada je u kontinuirana varijabla, prvo ćemo zbrojiti vrijednosti varijable u području između u i $u+du$, a tada integrirati preko svih takvih područja.
Dakle, neka je $f(u)$ funkcija kontinuirane varijable u, srednja je vrijednost te funkcije jednaka

$$\overline{f(u)} = \int_{a_1}^{a_2} f(u)\rho(u)du \tag{3.2.2}$$

Prema tome, i srednja vrijednost funkcije izražena je preko gustoće vjerojatnosti.
Svojstva srednje vrijednosti funkcije diskretne varijable koja smo naveli u poglavlju (2.3) vrijedit će i kada je dana varijabla kontinuirana.
Dakle, imamo dvije funkcije $f(u)$ i $g(u)$ kontinuriane varijable u tako da je $u \in (a_1, a_2)$. Za srednju vrijednost njihovog zbroja vrijedi

$$\overline{f(u)+g(u)} = \int_{a_1}^{a_2} (f(u)+g(u))\rho(u)du = \int_{a_1}^{a_2} f(u)\rho(u)du + \int_{a_1}^{a_2} g(u)\rho(u)du = \overline{f(u)} + \overline{g(u)} \tag{3.2.3}$$

Također, ako je c bilo koja konstanta, a $f(u)$ funkcija kontinuirane varijable u, vrijedi

$$\overline{c \cdot f(u)} = c \cdot \overline{f(u)}$$

Napomena:
Ponekad se može dogoditi da je gustoća vjerojatnosti $\rho(u)$ za određene vrijednosti varijable beskonačna. Ta pojava ne predstavlja problem dok je god integral

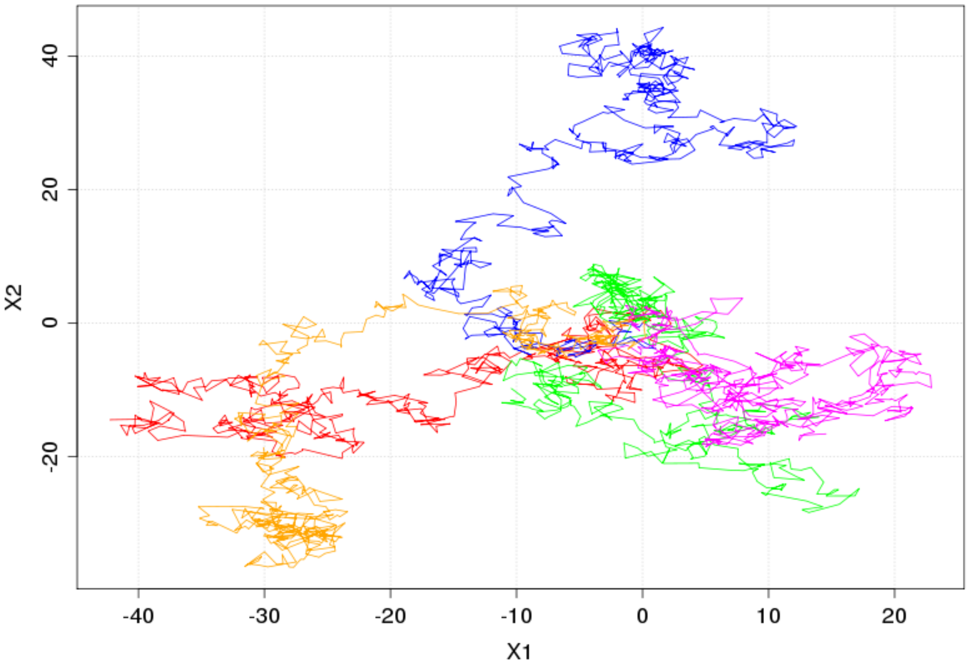
$$\int_{c_1}^{c_2} \rho(u)du$$

koji predstavlja vjerojatnost da u poprima sve vrijednosti u proizvoljnom području (c_1, c_2) konačan.

Prethodni problem traženja vjerojatnosti jedne kontinuirane varijable, može se jednostavno proširiti i na problem više varijabli.
Radi jednostavnosti promatrat ćemo slučaj dvije kontinuirane varijable, ali rezultati se lako mogu poopćiti i na više varijabli.
Dakle, promatramo dvije kontinuirane varijable u i v takve da je $a_1 < u < a_2$ i $b_1 < v < b_2$.
Pri tome je $\rho(u, v)dudv$ vjerojatnost da varijabla u poprima vrijednosti iz područja $(u, u+du)$ i varijabla v vrijednosti iz područja $(v, v + dv)$.
Problem kontinuiranih varijabli ponovno svodimo na problem prebrojivih diskretnih varijabli. Prema tome, kao i u slučaju jedne varijable, i ovdje ćemo neprekidno područje podijeliti na jednake infinitezimalne intervale kako bi dobili prebrojivi skup vrijednosti.
Kako sada imamo dvije kontinuirane varijable, potrebno je prvo područje (a_1, a_2) podijeliti na jednake infinitezimalne intervale δu, gdje je u_i vrijednost varijable u te drugo područje (b_1, b_2) podijeliti na jednake infinitezimalne intervale δv u kojima je v_j vrijednost varijable v.
Tada je $P(u_i, v_j)$ vjerojatnost da je $u = u_i$ te istodobno $v = v_j$. Kako su vjerojatnosti

$P(u_i, v_j)$ jednake u svakom infinitezimalnom području, možemo ih (kao i kod jedne varijable), označiti samo sa $P(u, v)$.
Također, diferencijal du sadrži $\frac{du}{\delta u}$ mogućih vrijednosti od u, dok diferencijal dv sadrži $\frac{dv}{\delta v}$ mogućih vrijednosti od v. Uzimajući ovo sve u obzir, proizlazi da je vjerojatnost $\rho(u, v)dudv$ jednaka

$$\rho(u, v)dudv = P(u, v)\frac{du}{\delta u}\frac{dv}{\delta v} \tag{3.2.4}$$



Slika 14: Nasumični hod u dvije dimenzije kada je pomak kontinuirana varijabla

I ovdje je uvjet normiranja vjerojatnosti dan preko gustoće vjerojatnosti, dakle mora vrijediti

$$\int_{a_1}^{a_2}\int_{b_1}^{b_2}\rho(u, v)dudv = 1 \tag{3.2.5}$$

Izrazom (3.1.5) dana je srednja vrijednost funkcije dvije diskretne varijable. Kako su ovdje dane varijable kontinuirane, sa zbroja ćemo prijeći na integral, a raspodjelu vjerojatnosti zamijenit će gustoća vjerojatnosti.
Dakle, za funkciju $F(u, v)$ kontinuiranih varijabli u i v definiramo njezinu srednju vrijednost $\overline{F(u, v)}$

$$\overline{F(u, v)} = \int_{a_1}^{a_2}\int_{b_1}^{b_2} F(u, v)\rho(u, v)dudv \tag{3.2.6}$$

Može se pokazati da za funkcije dvije kontinuirane varijable vrijede izrazi (3.1.6) i (3.1.8) za srednju vrijednost zbroja, odnosno umnoška funkcija.

3.2.1 Funkcije slučajnih varijabli

U daljnjem tekstu pozabavit ćemo se sljedećim problemom: neka je u kontinuirana slučajna varijabla te neka je $\rho(u)du$ vjerojatnost da se dana varijabla nalazi u intervalu $(u, u+du)$. Definiramo jednoznačnu realnu funkciju $\varphi(u)$ te želimo naći vjerojatnost $W(\varphi)d\varphi$ da se varijabla φ nalazi između $(\varphi, \varphi + d\varphi)$.
Tražena se vjerojatnost dobije zbrajanjem svih onih vjerojatnosti od u takvi da se φ nalazi u intervalu $(\varphi, \varphi + d\varphi)$.
Dakle, za $W(\varphi)d\varphi$ imamo

$$W(\varphi)d\varphi = \int_{d\varphi} \rho(u)du$$

Ako sada u promatramo kao funkciju od φ, tada područje integracije možemo proširiti preko svih vrijednosti od u koje leže između $u(\varphi)$ i $u(\varphi + d\varphi)$.
Tada vrijedi

$$W(\varphi)d\varphi = \int_{\varphi}^{\varphi+d\varphi} \rho(u) \left| \frac{du}{d\varphi} \right| d\varphi$$

Ovo možemo učiniti samo ako je $u(\varphi)$ jednoznačna funkcija od φ jer se tada cijeli prethodni integral može izraziti preko φ.

3.3 Općeniti izračun srednjih vrijednosti za problem nasumičnog hoda

U poglavlju (2.4) bavili smo se računanjem srednjih vrijednosti za problem nasumičnog hoda u jednoj dimenziji. U tom specijalnom slučaju pretpostavili smo kako je duljina svakog koraka hoda bila jednaka. Tako dobiveni rezultati nisu dovoljno općeniti i primjenjivi u svim situacijama. Stoga ćemo se u ovom poglavlju baviti upravo nešto općenitijom situacijom problema nasumičnog hoda. Odnosno, promatrat ćemo takav slučaj kada duljina koraka čestice koja se giba nasumično, nije nužno stalna.
Dakle, promatramo nasumično gibanje čestice u jednoj dimenziji. Neka je pri tom pomak čestice u i-tom koraku, pozitvan ili negativan, označen sa s_i.
Vjerojatnost da se pomak nalazi u intervalu $(s_i, s_i + ds_i)$ jednaka je $\omega(s_i)ds_i$ i neovisna o pomacima koji se događaju u drugim koracima.
Radi jednostavnosti pretpostavljamo kako je raspodjela vjerojatnosti ω jednaka za svaki korak i.
Dakle,u ovakvoj sitauciji više neće biti bitna duljina koraka čestice već će nas zanimati

raspodjela mogućih duljina koraka s relativnom vjerojatnosti ω.
Želimo naći ukupan pomak x nakon N koraka, zatim srednju vrijednost toga pomaka $\overline{x}$ te disperziju pomaka $\overline{(\Delta x)^2}$.
Ukupan pomak x nakon N koraka jednak je

$$x = s_1 + s_2 + \cdots + s_N$$

Odnosno,

$$x = \sum_{i=1}^{N} s_i \tag{3.3.1}$$

gdje je s_i pomak u i-tom koraku. Onda je srednja vrijednost $\overline{x}$ ukupnog pomaka jednaka

$$\overline{x} = \overline{\sum_{i=1}^{N} s_i}$$

Što, kada primjenimo svojstvo (3.2.3) za srednju vrijednost zbroja, postaje

$$\overline{x} = \sum_{i=1}^{N} \overline{s_i} \tag{3.3.2}$$

gdje veličina $\overline{s_i}$ označava srednji pomak po koraku.
Dakle, srednja vrijednost ukupnog pomaka jednaka je zbroju srednjih vrijednosti pomaka u svakom koraku.
U prethodnoj relaciji primjenili smo svojstvo za srednju vrijednost zbroja funkcija kontinuirane varijable. To smo učinli jer je pomak s_i kontinuirana varijabla, stoga ćemo ju i u daljnim izračunima smatrati takvom. To znači da ćemo pri računanju srednje vrijednosti i disperzije zbroj zamijeniti integralom.
Kao što smo u uvodu spomenuli, raspodjela vjerojatnosti pomaka $\omega(s_i)$ ista je za svaki korak pa stoga zaključujemo kako i srednja vrijednost pomaka $\overline{s_i}$ mora biti za svaki korak jednaka.
Čime onda srednja vrijednost ukupnog pomaka, $\overline{x}$, postaje zapravo zbroj N jednakih članova.
Odnosno, pišemo

$$\overline{x} = N\,\overline{s} \tag{3.3.3}$$

Gdje smo sa $\overline{s}$ označili srednji pomak po koraku koji je za svaki korak isti, dakle

$$\overline{s} \equiv \overline{s_i} = \int \omega(s)\, s\, ds$$

I ovdje smo zbog kontinuirane varijable s_i sa zbroja prešli na integral.
Prema tome, srednja je vrijednost ukupnog pomaka x za N koraka jednaka umnošku broja koraka N i srednjeg pomaka pojedinog koraka.

Odstupanje od srednje vrijednosti ukupnog pomaka Δx jednako je

$$\Delta x = x - \overline{x} \tag{3.3.4}$$

Odnosno, primjenom izraza (3.3.1) i (3.3.2) Δx postaje

$$x - \overline{x} = \sum_i s_i - \sum_i \overline{s_i} = \sum_i (s_i - \overline{s_i})$$

Kako je srednji pomak u svakom koraku isti, odnosno $\overline{s_i} \equiv \overline{s}$ gornji će izraz postati

$$x - \overline{x} = \sum_i (s_i - \overline{s})$$

Odstupanje pomaka s_i od srednje vrijednosti pomaka pojedinog koraka $\overline{s}$ jednako je

$$\Delta s_i = s_i - \overline{s}$$

Pa je, konačno, odstupanje ukupnog pomaka, Δx jednako

$$\Delta x = \sum_{i=1}^{N} \Delta s_i$$

Kada gornju relaciju kvadriramo, slijedi

$$(\Delta x)^2 = \left(\sum_i \Delta s_i\right)\left(\sum_j \Delta s_j\right) = \sum_i (\Delta s_i)^2 + \sum_i \sum_j (\Delta s_i)(\Delta s_j)$$

pri čemu vrijedi $i \neq j$.
A srednja je vrijednost toga izraza jednaka

$$\overline{(\Delta x)^2} = \overline{\sum_i (\Delta s_i)^2} + \overline{\sum_i \sum_j (\Delta s_i)\,(\Delta s_j)}$$

Odnosno, zbog svojstva (3.2.3) da je srednja vrijednost zbroja jednaka zbroju srednjih vrijednosti, slijedi

$$\overline{(\Delta x)^2} = \sum_i \overline{(\Delta s_i)^2} + \sum_i \sum_j \overline{(\Delta s_i)\,(\Delta s_j)}$$

Primjenimo li sada na mješoviti član prethodnog izraza pravilo (3.1.8), dobit ćemo sljedeće

$$\overline{(\Delta s_i) \cdot (\Delta s_j)} = \overline{(\Delta s_i)} \cdot \overline{(\Delta s_j)}$$

Svojstvo (3.1.8) definirano za diskretne varijable, vrijedi, kao što smo to u poglavlju (3.2) zaključili i za kontinuirane varijable. A ovdje ga možemo primjeniti jer vjerojatnost $\omega(s_i)ds_i$

u i-tom koraku ne ovisi o vjerojatnosti u ostalim koracima. Dakle, koraci su statistički neovisni. S obzirom da je

$$\overline{(\Delta s_i)} = \overline{s_i} - \overline{s} = \overline{s} - \overline{s} = 0$$

Proizlazi, srednja je vrijednost odstupanja od srednje vrijednosti jednaka nula. Prema tome, taj mješoviti član jednak je nula te nam za disperziju ukupnog pomaka $\overline{(\Delta x)^2}$ ostaje

$$\overline{(\Delta x)^2} = \sum_{i=1}^{N} \overline{(\Delta s_i)^2}$$

Raspodjela vjerojatnosti $\omega(s_i)$ ne ovisi o koraku i, i za svaki je korak jednaka pa zbog toga mora i disperzija pomaka svakog koraka biti jednaka. Što zapisujemo kao

$$\overline{(\Delta s)^2} \equiv \overline{(\Delta s_i)^2} = \int ds\, \omega(s)(\Delta s)^2$$

Znači, disperzija $\overline{(\Delta x)^2}$ ukupnog pomaka zapravo je jednaka zbroju N jednakih članova

$$\overline{(\Delta x)^2} = N\, \overline{(\Delta s)^2} \tag{3.3.5}$$

Disperzija pomaka $\overline{(\Delta x)^2}$ je srednje kvadratno odstupanje od srednje vrijednosti pomaka $\overline{x}$. Standardna devijacija pomaka σ omogućuje izravno mjerenje širine raspodjele pomaka oko njegove srednje vrijednosti. Standardna je devijacija jednaka

$$\sigma = \left[\overline{(\Delta x)^2}\right]^{\frac{1}{2}}$$

Odnosno

$$\sigma = \left[N\, \overline{(\Delta s)^2}\right]^{\frac{1}{2}}$$

Iz prethodnog je izraza jasno kako standardna devijacija σ raste sa vrijednosti N i to kao $N^{\frac{1}{2}}$.

Iz izraza (3.3.3) slijedi da srednja vrijednost pomaka $\overline{x}$ raste srazmjerno sa brojem varijabli (brojem koraka nasumičnog hoda čestice) N.

Relativna širina raspodjele izražava postotak odstupanja raspodjele vrijednosti x od njegove srednje vrijednosti, a jednaka omjeru standradne devijacije i srednje vrijednosti varijable

$$\frac{\sigma}{\overline{x}} = \frac{\left[N\, \overline{(\Delta s)^2}\right]^{\frac{1}{2}}}{N\, \overline{s}} = \frac{\left[\overline{(\Delta s)^2}\right]^{\frac{1}{2}}}{\sqrt{N}\, \overline{s}} = \frac{1}{\sqrt{N}} \frac{\sigma(s)}{\overline{s}}$$

Prema tome, relativna širina raspodjele $\frac{\sigma}{\overline{x}}$, kada N raste, pada kao $N^{-\frac{1}{2}}$. Odnosno, kako N raste, relativna širina postaje zanemarivo mala.

Na sljedećem ćemo primjeru vidjeti primjenu rezultata dobivenih u ovom poglavlju. Dobivene relacije primjenit ćemo na poseban slučaj nasumičnog hoda, kada je duljina svakog koraka ista.

Primjer:
Promatramo nasumičan hod u jednoj dimenziji, gdje je duljina svakog koraka čestice ista i jednaka l. Tada je vjerojatnost koraka u desno jednaka p, a vjerojatnost koraka u lijevo jednaka $q = 1 - p$.
Srednji pomak po koraku $\overline{s}$ jednak je

$$\overline{s} = p\,l + q\,(-l)$$

Kada iskoristimo činjenicu da je $q = 1 - p$, srednji pomak postaje

$$\overline{s} = (2p - 1)l$$

Srednja vrijednost kvadrata pomaka jednaka je

$$\overline{s^2} = p\,l^2 + q\,(-l)^2 = (p+q)l^2 = l^2$$

Općenito, disperzija pomaka pojedinog koraka jednaka je

$$\overline{(\Delta s)^2} = \overline{s^2} - \overline{s}^{\,2}$$

Iz čega uvrštavanjem prethodnih izraza slijedi

$$\overline{(\Delta s)^2} = l^2 - [(2p-1)l]^2 = l^2[1 - (2p-1)^2] = 4l^2p(1-p) = 4pql^2$$

Prema tome, srednja vrijednost ukupnog pomaka, $\overline{x}$ jednaka je

$$\overline{x} = N\,\overline{s} = N\,(p-q)\,l$$

Dok je disperzija ukupnog pomaka $\overline{(\Delta x)^2}$ jednaka

$$\overline{(\Delta x)^2} = N\,\overline{(\Delta s)^2} = 4Np\,ql^2$$

Kako je $x = ml$, iz gornjih bi izraza slijedilo da je

$$\overline{m} = N(p-q)$$
$$\overline{(\Delta m)^2} = 4Np\,q$$

A to su upravo izrazi za srednju vrijednost i disperziju pomaka m koje smo dobili u poglavlju (2.4) za poseban slučaj nasumičnog hoda kada je duljina koraka jednaka. Dakle, vidimo da se općeniti izrazi dobiveni u ovom poglavlju mogu primjeniti i na taj poseban slučaj te su rezultati jednaki.

3.4 Računanje raspodjele vjerojatnosti

U prošlom smo poglavlju definirali problem nasumičnog hoda na nešto općenitiji način nego ranije. Dakle, pretpostavili smo kako duljina koraka prilikom nasumičnog hoda česice ne mora uvijek biti jednaka. Ovdje ćemo za tako definiranu situaciju tražiti vjerojatnost ukupnog pomaka x.
Dakle, ponovno promatramo nasumično gibanje čestice u jednoj dimenziji. Želimo naći vjerojatnost $\rho(x)dx$ da se čestica nakon ukupno N koraka nađe u intervalu $(x, x + dx)$.
Ponovno je ukupan pomak čestice dan izrazom

$$x = \sum_{i=1}^{N} s_i$$

gdje je s_i pomak u i-tom koraku, a $\omega(s_i)ds_i$ vjerojatnost da i-ti pomak leži između $(s_i, s_i + ds_i)$.
I ponovno su koraci koje čestica pravi statistički neovisni te je vjerojatnost $\omega(s_i)ds_i$ neovisna o pomacima koji se događaju u drugim koracima.
Upravo zbog činjenice da su koraci neovisni, možemo vjerojatnost određenog niza koraka zapisati kao umnožak vjerojatnosti pojedinih koraka, odnosno

$$\omega(s_1)ds_1 \cdot \omega(s_2)ds_2 \cdot \ \cdots \ \cdot \omega(s_N)ds_N \tag{3.4.1}$$

Tada se vjerojatnost da se x nalazi između x i $x + dx$ dobije zbrajanjem vjerojatnosti svih mogućih pomaka (koraka) i ne ovisi o nizu koraka koji proizvedu ukupan pomak.
Kako je varijabla s_i kontinuirana, ovdje ćemo zbroj zamijeniti integralom, pri čemu je područje integracije od $-\infty$ do ∞. Dakle, vjerojatnost $\rho(x)dx$ jednaka je

$$\rho(x)dx = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \omega(s_1)\omega(s_2) \cdots \omega(s_N)\, ds_1 ds_2 \cdots ds_N \tag{3.4.2}$$

Integrirati možemo od $-\infty$ do ∞, ali uz uvjet da se pomak x nalazi u intervalu $(x, x + dx)$.
Odnosno, mora vrijediti

$$x < \sum_{i=1}^{N} s_i < x + dx \tag{3.4.3}$$

Traženu vjerojatnost lako bi mogli odrediti iz izraza (3.4.2), ali zbog uvjeta (3.4.3) pojavljuje se kompliciran geometrijski problem preko kojega područja integrirati tako ga granice zadovoljavaju uvjet (3.4.3).
Problem ćemo riješiti tako da bez ograničenja integriramo po svim varijablama s_i, a problem zadovoljavanja uvjeta prebacimo u integrand.
I to tako da integrand u izrazu (3.4.2) pomnožimo faktorom koji zadovoljava sljedeće: taj je

faktor jednak jedan kada je ispunjen uvjet (3.4.3), a nula kada nije ispunjen.
Dakle, dani ćemo integral pomnožiti Diracovom δ funkcijom $\delta(x - x_0)$ jer upravo ona ima to svojstvo promjenjivosti: kada je

$$|x - x_0| > \frac{1}{2}|dx|$$

onda Diracova δ funkcija $\delta(x - x_0) \to 0$.
A kada je

$$|x - x_0| < \frac{1}{2}|dx|$$

tada Diracova funkcija $\delta(x - x_0) \to \infty$, odnosno integral Diracove funkcije jednak je 1

$$\int_{-\infty}^{\infty} \delta(x - x_0) = 1$$

Prema tome, integrand izraza (3.4.2) množimo faktorom

$$\delta\left(x - \sum_{i=1}^{N} s_i\right) dx$$

Time vjerojatnost $\rho(x)dx$ postaje

$$\rho(x)dx = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \omega(s_1) \cdots \omega(s_N) \left[\delta\left(x - \sum_{i=1}^{N} s_i\right) dx\right] ds_1 \cdots ds_N$$

Odnosno, sama je gustoća vjerojatnosti $\rho(x)$ jednaka

$$\rho(x) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \omega(s_1) \cdots \omega(s_N) \left[\delta\left(x - \sum_{i=1}^{N} s_i\right)\right] ds_1 \cdots ds_N$$

Integralna reprezentacija Diracove δ funkcije je

$$\delta\left(x - \sum_{i=1}^{N} s_i\right) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ik[\sum s_i - x]}$$

Uvrstivši to u izraz za gustoću vjerojatnosti $\rho(x)$, ta gustoća postaje

$$\rho(x) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \omega(s_1) \cdots \omega(s_N) \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ik(s_1 + \cdots + s_N - x)} ds_1 \cdots ds_N$$

Iskoristimo li ovdje svojstvo množenja eksponencijalne funkcije, gornji izraz možemo napisati

$$\rho(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \omega(s_1) \cdots \omega(s_N) \int_{-\infty}^{\infty} \left(e^{iks_1} \cdot \ \cdots \ \cdot e^{iks_N} \cdot e^{-ikx}\right) dk$$

Dobiveni izraz možemo još srediti zamijenom redoslijeda integriranja te konačno za gustoću $\rho(x)$ dobijemo

$$\rho(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dk e^{-ikx} \int_{-\infty}^{\infty} ds_1 \omega(s_1) e^{iks_1} \cdots \int_{-\infty}^{\infty} ds_N \omega(s_N) e^{iks_N}$$

Kako je vjerojatnost $\omega(s_i)ds_i$ jednaka za svaki i, tada mora i svaki integral oblika

$$\int_{-\infty}^{\infty} ds_i \omega(s_i) e^{iks_i}$$

biti jednak. Odnosno,

$$\int_{-\infty}^{\infty} ds_1 \omega(s_1) e^{iks_1} = \cdots = \int_{-\infty}^{\infty} ds_N \omega(s_N) e^{iks_N}$$

Dakle, možemo svaki takav integral označiti sa $Q(k)$

$$Q(k) \equiv \int_{-\infty}^{\infty} ds \omega(s) e^{iks}$$

Prema tome, množenjem N takvih integrala dobije se

$$Q^N(k) \equiv \int_{-\infty}^{\infty} ds_1 \omega(s_1) e^{iks_1} \cdot \cdots \cdot \int_{-\infty}^{\infty} ds_N \omega(s_N) e^{iks_N}$$

Konačno, gustoća je vjerojatnosti $\rho(x)$ jednaka

$$\rho(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dk e^{-ikx} Q^N(k) \tag{3.4.4}$$

Time smo odredili izraz za gustoću vjerojatnosti $\rho(x)$ u ovom općenitom slučaju, kada duljina koraka nasumičnog hoda nije konstantna. Dakle, vjerojatnost $\rho(x)dx$ da se ukupni pomak x nalazi u intervalu $(x, x+dx)$ dobije se računanjem dvaju integrala u izrazu (3.4.4).

Literatura

[1] F. Rief, *Fundamentals of statistical and thermal physics*, McGraw-Hill Book Company, New York,1965.

[2] V. Vranić, *Vjerojatnost i statistika*, Tehnička knjiga, Zagreb, 1971.

[3] http://www.math.dartmouth.edu/ prob/prob/prob.pdf

[4] http://coco.ccu.uniovi.es/malva/sketchbook/lssketchbook/examples/fractal/ fractal.htm

[5] http://www.websters-online-dictionary.org/definitions/Random

[6] http://doc.openturns.org/openturns-latest/html/UseCasesGuide/cid5.xhtml

[7] http://pages.stern.nyu.edu

[8] http://faculty.ccc.edu/colleges/hwashington/math/links.html

[9] http://jwilson.coe.uga.edu/EMAT6680Fa06/Crumley/Normal/Normal1.html

[10] http://www.phy.ornl.gov/csep/mc/node19.html

[11] http://www.nationmaster.com/encyclopedia/random-walk

[12] http://doc.openturns.org/openturns-latest/html/UseCasesGuide/cid5.xhtml

[13] http://simple.wikipedia.org/wiki/Normaldistribution

4 SAŽETAK

U diplomskom radu opisan je problem nasumičnog hoda u jednoj dimenziji. Izvedeni su izrazi za vjerojatnost pomaka čestice od ishodišta, zatim za vjerojatnost broja koraka u desno, odnosno u lijevo te za disperziju i standardnu devijaciju tih vrijednsoti kada je pomak čestice diskretna varijabla.
Također, dan je izraz za računanje vjerojatnosti nekog položaja čestice kada je njezin pomak kontinuirana varijabla.
U radu se, također, govori i o poopćenju nasumičnog hoda u dvije i više dimenzija, te su i za takav slučaj dani izrazi za računanje vjerojatnosti.

5 SUMMARY

This paper describes the problem of random walk in one dimension. There have been given the expressions for the probability of net displacement, the probability of right or left steps and the dispersion and the standard deviation of these values when the displacement of particles is discrete variable. There has also been given the expression to calculate the probability of any position of the particle when the displacement is a continuous variable. The paper also describes the generalization of random walk in two or more dimensions and it gives the equations for calculating the probability of those cases.

6 ŽIVOTOPIS

Petra Pandža, rođena 11. listopada 1985. u Đakovu. Od 1992. g. pohađa Osnovnu školu "I. G. Kovačić" u Đakovu. Školske godine 2000./ 2001. upisuje Gimnaziju A. G. Matoša u Đakovu koju 2004. g. završava. Iste godine upisuje na Sveučilištu J. J. Strossmayera u Osijeku, Odjel za matematiku, smjer matematika - fizika.