

Vaja 1: Računanje z napakami

Matej Bažec

9. oktober 2015

Povzetek

Spoznali bomo osnove računanja z napakami. Obnovili bomo znanje o absolutnih in relativnih napakah, smiselnosti zapisa decimalnih mest in pravila računanja z aritmetičnimi operacijami. Znanje bomo nadgradili z računanjem napak poljubne funkcijske odvisnosti. Naučili se bomo, kaj je to sistematična in kaj statistična napaka. Slednji bomo posvetili nekaj več pozornosti in se naučili nekaj praktičnih posledic centralnega limitnega izreka.

1 Uvod

Velika večina meritev - izjema so tiste, kjer je rezultat diskretna vrednost - je obremenjenih z napakami, ki pa lahko imajo različen izvor. Lahko izvirajo iz nenatančnosti samega inštrumenta, zunanjih motenj, napake merske metode, računskih postopkov ...

1.1 Navajanje napak

Recimo, da nam nekdo pove, da je dolžina palice $100\text{cm} \pm 5\text{mm}$. To nam pove, da lahko z veliko gotovostjo trdimo, da je dolžina palice med 99,5cm in 100,5cm. Mar to pomeni, če je dejanska dolžina palice 99,4cm, da se nam ta meritev ne sklada z realnostjo? Odvisno. Če bi vam policist na radarju izmeril 51km/h, čeprav veste, da ste vozili 50km/h, bi vam bilo v bolj slabo tolažbo, da je sicer natančnost merilca 0,5km/h, ampak je vaša meritev padla ven iz območja velike gotovosti. Podobno bi se najbrž pritožili v trgovini, če bi trgovec jamčil, da je palica metrska znotraj območja 5mm. Če pa vam merska naprava z natančnostjo 5mm izmeri 100cm palico, ki je sicer 99,4cm dolga, po navadi ni s tem nič narobe.

Zakaj? Kot smo že prej povedali, je napaka meritve območje, znotraj katerega se z veliko gotovostjo nahaja pravi rezultat. Pri meritvah je lahko velika gotovost že 60%, kar pomeni dokaj veliko verjetnost, da bodo nekatere meritve izven navedenega območja. Pomembno je torej, da se zavedamo, da napaka meritve ne pomeni nujno jamstva, da je merjeni rezultat znotraj območja napake. Če bomo torej želeli jamčiti, da je dolžina palice znotraj območja 5mm, jo bomo morali izmeriti z napravo, ki bo, recimo, na 1mm natančna.

Zato je nesmiselno podajati napako več kot na velikostni red natančno in jo vedno podamo na eno decimalno mesto natančno. Napake tipa $\pm 4,62$ so nesmiselne!

Iz istega razloga nikoli ne navajamo rezultata na tista decimalna mesta, ki so manjša od napake. Zato je zapis npr. $190,563 \pm 0,3$ nesmiseln in zavajajoč. Take rezultate moramo smiselno zaokrožiti. Torej bi bil pravilen zapis $190,6 \pm 0,3$. Prav tako ne smemo po drugi strani preveč zaokroževati. Rezultat 200 bi bil pregroba zaokrožitev tega rezultata.

Nekateri študenti imajo po navadi težave pri smiselnem zaokroževanju velikih in majhnih števil. Nakažimo to na primeru človeške višine. Denimo, da smo osebo izmerili, da je 179cm visoka na 1cm natančno, kar je precej smiseln rezultat. Tako, da temu rezultatu ni kaj za dodati, niti za odvzeti. Ampak ta isti rezultat lahko po drugi strani podamo v km, m ali μm (pustimo ob strani smiselnost tega početja). Torej dobimo, da je oseba visoka 0,00179km, 1,79m ali $179000\mu\text{m}$. V vseh štirih primerih

Napaka je območje, znotraj katerega je velika verjetnost, da se vrednost nahaja. Nikakor pa ni zagotovilo, da se v njem zares nahaja.

Napako vedno podajamo na eno decimalno mesto natančno. Npr. $\pm 4, \pm 30, \pm 0,2$ itd. Rezultat smiselno zaokrožimo na velikostni red napake.

gre za en in isti rezultat. Zanima nas, kakšen lok bo ta oseba narisala, če jo zavrtimo za 60° . To izračunamo tako, da pomnožimo njeno višino s $\pi/3 = 1,0471975512$. Vstavimo v kalkulator in dobimo 187,448361664cm. Po drugi strani je to lahko 0,00187448361664km, 1,87448361664m ali 1874483,61664 μ m. Kako te rezultate smiselno zaokrožimo? Jasno je, da bo tudi ta rezultat imel natančnost 1cm. Torej je smiselni odgovor 187cm. Kako pa je to pri drugih enotah? Pri kilometrih moramo prvih pet mest za decimalno vejico pustiti na miru, pri metrih dve, pri mikronih pa lahko zanemarimo kar štiri mesta pred njo: 0,00187km, 1,87m in 1870000 μ m.

Kakšen lok bi narisal s svojo višino? Smiselno zapiši rezultat v sledečih enotah:

_____ cm
 _____ km
 _____ m
 _____ mm
 _____ μ m

1.2 Absolutna in relativna napaka

V dozdašnjih primerih smo napako navajali s konkretnim odstopanjem od verjetne vrednosti. Tako podani napaki pravimo absolutna. Alternativno lahko napako podamo z deležem odstopanja prave vrednosti. Tako podani napaki pravimo relativna. V prejšnjem primeru je bila meritev 179cm $\pm$ 1cm podana z absolutno napako (1cm). Kakšen delež predstavlja 1cm glede na celotno vrednost 179cm?

$$\frac{1\text{cm}}{179\text{cm}} = 0,0055866 \quad (1)$$

Relativna napaka je torej 0,6%. Tudi 0,5% bi bil pravilen rezultat, nikakor pa ne 0,56%. Razmisli zakaj. Celotna meritev podana z relativno napako bi potem bila 179cm(1 $\pm$ 0,6%).

Če je torej prava vrednost x , izmerjena vrednost x_0 , absolutna napaka pa δx , je relativna napaka $\delta x/x_0$. Popoln merski rezultat lahko zapišemo z absolutno napako kot

$$x = x_0 \pm \delta x \quad (2)$$

ali z relativno kot

$$x = x_0 \left(1 \pm \frac{\delta x}{x_0} \right). \quad (3)$$

Medtem ko je rezultat podan z absolutno napako vedno smiseln, pri relativni ni vedno tako. Relativna napaka ni smiselna pri količinah, ki se podajajo s skalo. Tipična primera sta podajanje smeri v stopinjah ali temperature v celzijevi skali. Nakažimo to na primeru. Recimo, da lahko določimo ladji smer plutja na 1° natančno. Če bi ladja plula v smeri 10° , bi bila relativna napaka 10%, če pa bi plula v smeri 100° , pa bi bila 1%. V še večjo zagato bi prišli, če bi želeli na podoben način dobiti relativno napako za ladjo, ki pluje proti severu. Pomisli zakaj.

Ker gre v vseh primerih za isto natančnost, je seveda podajanje relativne napake nesmiselno. Podoben razmislek velja za termometre na celzijevo skalo. Če je ta občutljiv na 1°C , bi bila relativna napaka tik nad lediščem zelo velika pri vrelišču pa komaj 1%. Mimogrede, za termometre, ki merijo absolutno temperaturo, ta razmislek ne velja.

Kljub temu to ne pomeni, da ko gre za kote in stopinje celzija, da je relativna napaka vedno nesmiselna. Če rečemo, da smo izmerili temperaturno razliko 20°C z natančnostjo 5%, je rezultat smiseln. Prav tako, če rečemo, da smo spremenili kot plutja za 5° z natančnostjo 10%. Zakaj?

Relativno napako dobimo iz absolutne tako, da jo delimo s pričakovanim rezultatom; absolutno iz relativne pa, da pomnožimo relativno napako s pričakovanim rezultatom.

1.3 Implicitno podana napaka

Kljub temu, da imamo pri podanih količinah skoraj vedno opravka z napakami, jih v fiziki in tehniki pogosto ne zapisujemo neposredno. Raje uporabljamo dogovor, da je rezultat natančen na toliko decimalnih mest, kolikor jih je podanih. Če imamo torej zapisan rezultat 21,7 je s tem mišljeno, da je rezultat natančen na 0,1. Predvsem zato je pomembno, da smiselno zaokrožujemo. Če računamo kvadratni koren od 21,7, je torej narobe, če zapišemo $\sqrt{21,7} = 4,65832587954$. Čeprav je rezultat numerično točen, se da iz zapisa razbrati, da je njegova natančnost 0,00000000001, kar pa ni res. Pravilno je torej $\sqrt{21,7} = 4,7$ in nobene decimalke več (ali manj).

Iz istega razloga ne smemo opuščati decimalnih ničel. 2, 2,0 in 2,00 so numerično iste vrednosti. Toda iz zapisa sklepamo, da so napake 1, 0,1 in 0,01 po istem vrstnem redu. Zato moramo te ničle obvezno pisati (če je seveda to utemeljeno z natančnostjo). Narobe je torej zapis $1,233 + 1,767 = 3$, ker smo pri rezultatu opustili preveč ničel. Pravilno bi bilo $1,233 + 1,767 = 3,000$.

2 Računanje z napakami

2.1 Osnovne aritmetične operacije

Recimo, da smo izmerili dve količini $x = x_0 \pm \delta x$ in $y = y_0 \pm \delta y$. Radi bi dobili oceno napake za vsoto obeh. Ker se x z veliko verjetnostjo nahaja med $x_0 - \delta x$ in $x_0 + \delta x$ in podobno tudi y , se bo vsota z veliko verjetnostjo nahajala med $(x_0 - \delta x) + (y_0 - \delta y)$ in $(x_0 + \delta x) + (y_0 + \delta y)$. Torej bo absolutna napaka $\delta x + \delta y$. Izpeljimo to še s formulo.

$$x + y = (x_0 \pm \delta x) + (y_0 \pm \delta y) = (x_0 + y_0) \pm (\delta x + \delta y) \quad (4)$$

Tukaj je na mestu še odstavek za zahtevnejše bralce. Verjetnost, da je x nad izmerjeno vrednostjo x_0 je 50% in 50% pod njo. Isto velja za y . Zdaj pa se postavi vprašanje, kako sta meritvi x in y povezani - pri meritvah rečemo korelirani, ker je to dobro definirana matematična količina, v katero pa se na tem mestu ne bomo spuščali. Če merimo dve dolžini s kovinskim metrom, ki je pred tem bil izpostavljen močni vročini, lahko upravičeno sklepamo, da smo obe dolžini izmerili prekratki (ker se je meter na vročini raztegnil). Torej sta napaki v tem primeru med seboj povezani. Če pa merimo maso dveh vijakov v škatli, nimamo nobenega razloga, da bi sklepali, da sta oba vijaka težja (ali lažja) od povprečja v škatli. Torej sta meritvi nekorelirani. Verjetnost, da sta oba težja (oziroma lažja) od povprečja je tedaj 25%. V 50% pa bo en vijak težji, drugi pa lažji. Torej se bo pri seštevanju nekoreliranih količin napaka v povprečju popravljala. Za korelirane napake bomo torej uporabljali enačbo (4), za nekorelirane pa

$$\delta_{x+y} = \sqrt{\delta x^2 + \delta y^2}. \quad (5)$$

Zadnje enačbe ne bomo izpeljevali, ker presega naše zanimanje[2]. Zaradi narave napak, ki se zaokrožujejo le na prvo decimalno mesto, je prazaprov vseeno, katero od obeh enačb uporabljamo, dokler seštevamo samo dve količini. Če pa seštevamo 10 ali več količin, pa je razlika občutna.

Pri odštevanju postopamo podobno kot pri seštevanju:

$$x - y = (x_0 \pm \delta x) - (y_0 \pm \delta y) = (x_0 - y_0) \pm (\delta x + \delta y). \quad (6)$$

Razmisli, zakaj je v zadnjem oklepaju plus namesto minusa. Torej tudi pri odštevanju seštevamo absolutne napake. Vendar zna biti odštevanje s stališča napak problematično. Zakaj? Recimo, da se odločite, da boste za naslednje praznike pekli potico. Recept zahteva 1,5kg moke. Najprej bomo stehtali skledo - npr. 107g. Nato bomo vsuli noter toliko moke, da bo tehtnica pokazala 1607g. Tehnica ima 5g natančnost,

Pri seštevanju in odštevanju se absolutne napake seštevajo

Odštevanje dveh količin, ki sta si podobni, je s stališča napak problematično.

torej boste lahko upravičeno sklepali, da je v skledi $1500\text{g} \pm 10\text{g}$ moke, pri čemer je bila natančnost tehtnice kar 3‰.

Kaj pa če se vam je pokvarila kuhinjska tehtnica? Stopili boste na tehtnico za spremljanje telesne teže in izmerili 82,4kg. Ta tehtnica je natančna na 0,2kg. Nato boste vzeli v naročje toliko moke, da bo tehtnica pokazala 83,9kg. Torej imate na sebi $1500\text{g} \pm 400\text{g}$ moke. V tem primeru predstavlja napaka že skoraj tretjino izmerjene vrednosti, tudi če je natančnost tehtnice 2‰. Vidimo, da odštevanje količin samo po sebi ni problematično, lahko pa povzroča težave, če sta si odštevani količini preblizu. Takrat bo tudi rezultat blizu 0 in relativna napaka bo poskočila.

Pri množenju bomo raje uporabili zapis z relativno napako.

Pri množenju in deljenju seštevamo relativne napake.

$$xy = x_0 \left(1 \pm \frac{\delta x}{x_0}\right) y_0 \left(1 \pm \frac{\delta y}{y_0}\right) = x_0 y_0 \left(1 \pm \frac{\delta x}{x_0} \pm \frac{\delta y}{y_0} \pm \frac{\delta x \delta y}{x_0 y_0}\right) \quad (7)$$

Zadnji člen v oklepaju lahko zanemarimo - pravzaprav ga moramo zanemariti, saj napake zaokrožujemo na eno decimalno mesto. Recimo, da sta obe relativni napaki $\delta x/x_0 = \delta y/y_0 = 10\%$. Potem je njun produkt $(\delta x/x_0)(\delta y/y_0) = 1\%$. Zadosti je torej, da upoštevamo tiste prispevke, ki so linearni v napakah. Zato je produkt dveh merjenih količin enak

$$xy = x_0 y_0 \left(1 \pm \left(\frac{\delta x}{x_0} + \frac{\delta y}{y_0}\right)\right) \quad (8)$$

Torej se pri množenju relativne napake seštevajo.

Podobno je pri deljenju. Upoštevamo linearno aproksimacijo za $1/(1+h)$:

$$\frac{1}{1+h} \approx 1 - h. \quad (9)$$

Relativna napaka je namreč vedno manjša od 1, v nasprotnem primeru lahko govorimo le o oceni in nikakor ne o meritvi. Torej lahko uporabljamo linearno aproksimacijo za relativno napako. Pravzaprav je nesmiselno uporabiti karkoli natančnejšega. Razmisli zakaj. Zato lahko tudi $1/y$ zapišemo kot

$$\frac{1}{y} = \frac{1}{y_0} \frac{1}{1 \pm \frac{\delta y}{y_0}} = \frac{1}{y_0} \left(1 \pm \frac{\delta y}{y_0}\right). \quad (10)$$

Kvocien x/y nato obravnavamo kot produkt x in $1/y$.

$$\frac{x}{y} = \frac{x_0}{y_0} \left(1 \pm \left(\frac{\delta x}{x_0} + \frac{\delta y}{y_0}\right)\right). \quad (11)$$

Vpiši rezultate računov, ki jih dobiš na <http://vaje.fpp.uni-lj.si/tm>, vaja 1, naloga 1.

1. 1.: _____
1. 2.: _____
1. 3.: _____
1. 4.: _____
1. 5.: _____

2.2 Napaka splošne funkcije

Včasih lahko nastopa pri pretvorbi iz meritve v rezultat kakšna zahtevnejša operacija. Lep primer je funkcija sinus. Recimo torej, da potrebujemo $\sin \theta$ v rezultatu, izmerili pa smo kot θ , ki je enak 30° . Če je θ natančna na 1° , na koliko je natančen izraz $\sin \theta$? Ena možnost je, da izračunamo vrednost sinusa pri 29° in pri 31° ($\sin 29^\circ = 0,4848$ in $\sin 31^\circ = 0,5150$). Obe vrednosti odstopata za približno 0,015, torej je napaka 0,02.

Obstaja tudi elegantnejši (in bolj splošen) način, kako napako izračunamo. Podobno kot pri deljenju si lahko tudi tukaj izposodimo linearno aproksimacijo. Za splošno funkcijo f velja

$$f(x_0 + \Delta x) \approx f(x_0) + \Delta x f'(x_0). \quad (12)$$

Če je x_0 točna vrednost x in Δx odstopanje naše meritve od prave vrednosti (in je torej istega velikostnega reda kot napaka meritve), potem lahko dobimo napako količine f iz velikostnega reda razlike $f(x_0 + \Delta x) - f(x_0)$. Torej je napaka f (označimo jo z ϵ_f)

$$\epsilon_f = f'(x_0)\epsilon_x, \quad (13)$$

kjer je ϵ_x napaka merjenja količine x .

Vrnimo se na naš primer. Izračunamo odvod sinusa in dobimo napako meritve $\cos(30^\circ)0,017 = 0,015 \approx 0,02$. Razmisli, zakaj smo uporabili $\epsilon_x = 0,017$ in ne $\epsilon_x = 1^\circ$. Odkod 0,017 sploh pride?

Vpiši rezultate računov, ki jih dobiš na <http://vaje.fpp.uni-lj.si/tm>, vaja 1, naloga 2.

2. 1.: _____
 2. 2.: _____
 2. 3.: _____

Napaka količine, ki je v funkcijski zvezi z merjeno količino, je enaka produktu odvoda funkcije in napake merjene količine.

3 Sistematična in statistična napaka

3.1 Izvori napak

Napake lahko imajo različne izvore, ampak v grobem jih delimo v dve skupini: sistematične in statistične. Sistematične so tiste, ki se s ponavljanjem meritve ne spreminjajo. Tisti del, ki se spreminja pa je statistična napaka. Sistematična napaka je tipično napaka merske naprave. Razlogov za to je lahko več. Do napake lahko pride pri umeritvi (kalibraciji) inštrumenta, lahko pa je sama merska metoda natančna. Do slednjega pride največkrat zaradi uporabe linearnih zvez za količine, ki niso v linearni odvisnosti. Lep primer je raztegnjen meter. Vedno bo kazal manjšo dolžino, kot je pravilna.

Po drugi strani pa lahko meritev motijo pojavi, ki so bolj kot ne naključni. Največkrat je to električni šum. Lahko pa so to glede na tip meritve mehanske vibracije, radioaktivni razpad, nelinearni pojavi pri pretakanju fluidov ... Ker se tovrstne motnje pojavljajo časovno naključno, lahko upravičeno pričakujemo, da bomo s ponavljanjem meritev izboljšali merski rezultat.

Statistično napako torej lahko dobimo tako, da meritev večkrat ponovimo, iz odstopanja pa določimo statistično napako. S sistematično napako pa ni tako enostavno. Nekega splošnega pravila, kako jo dobimo, ni. Največkrat se poslužujemo specifikacije inštrumenta, kjer je tipično navedena natančnost. Če tega nimamo, ocenimo natančnost inštrumenta tako, da opravimo meritev že znane količine, vendar to ni vedno enostavno.

Sistematične napake ne popravimo s ponavljanjem meritve, statistično lahko.

3.2 Statistična napaka

Medtem ko je sistematična napaka odvisna od same meritve in jo moramo določiti za vsako meritev posebej, veljajo za statistično napako določena računska pravila. Najprej si bomo pogledali, kako lahko zmanjšamo statistično napako meritve s ponavljanjem. Recimo, da merimo količino x . Njena prava vrednost naj bo x_0 . Odstopanje naše meritve x od prave naj bo δx in naj ima statistično napako σ (njeno točno definicijo si bomo pogledali kasneje). Torej morata biti δx in σ istega velikostnega razreda in $x = x_0 + \delta x$. Pozor: δx je odstopanje meritve od pravega rezultata, torej

je lahko pozitivno ali negativno, σ pa nam pove širino intervala v katero bo večji del teh meritev padel in je vedno pozitiven.

Opravimo sedaj n takih meritev. i -to meritev bomo označili z x_i in njeno odstopanje od x_0 z δx_i . Ker so meritve iste imajo vse isto napako σ . Privzeli bomo tudi, da so meritve med seboj nekorelirane. Zanima nas, kakšna bo napaka povprečja n takih meritev.

Označimo z $\bar{A}$ povprečje količine A . Ločiti moramo povprečje količine, ki je neodvisno od konkretne meritve, od povprečja konkretne serije n meritev. Jasno je, da je $\bar{x}_i = x_0$ in je torej $\overline{\delta x_i} = 0$. Statistično napako σ definiramo kot koren povprečja kvadratov odstopanj $\sigma^2 = \overline{\delta x_i^2}$. Povprečje povprečja bo enako x_0 :

$$\overline{\frac{1}{n} \sum_{i=1}^n x_i} = \frac{1}{n} \sum_{i=1}^n \bar{x}_i = \frac{1}{n} \sum_{i=1}^n x_0 = x_0, \quad (14)$$

kar je pričakovan rezultat.

Izračunajmo statistično napako σ_n povprečja n meritev.

$$\sigma_n^2 = \overline{\left(\frac{1}{n} \sum_{i=1}^n x_i - x_0 \right)^2} = \overline{\left(\frac{1}{n} \sum_{i=1}^n \delta x_i \right)^2} \quad (15)$$

Kvadrat vsote razvijemo po dvojni vsoti.

$$\sigma_n^2 = \overline{\left(\frac{1}{n} \sum_{i=1}^n x_i - x_0 \right)^2} = \frac{1}{n^2} \sum_{i=1}^n \overline{\delta x_i \sum_{j=1}^n \delta x_j} = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \overline{\delta x_i \delta x_j} \quad (16)$$

Ustavimo se malo pri produktu $\delta x_i \delta x_j$. Če sta meritvi nekorelirani med sabo, je $\overline{\delta x_i \delta x_j} = \overline{\delta x_i} \overline{\delta x_j}$ (razen, ko je $i = j$, ker gre takrat za eno in isto meritev). Toda povprečje odmika je enako nič. Torej bo $\overline{\delta x_i \delta x_j}$ enak 0 za vse vrednosti razen za tiste z $i = j$. Torej lahko seštevanje po j opustimo.

$$\sigma_n^2 = \frac{1}{n^2} \sum_{i=1}^n \overline{\delta x_i^2} = \frac{1}{n^2} \sum_{i=1}^n \sigma^2 = \frac{\sigma^2}{n} \quad (17)$$

Korenimo levo in desno stran, pa dobimo

$$\sigma_n = \frac{\sigma}{\sqrt{n}}. \quad (18)$$

Iz tega rezultata lepo vidimo, da statistična napaka pada s ponavljanjem meritev. Vendar je to padanje korensko. To pomeni, da bomo z malim številom meritev bistveno izboljšali napako, za večjo natančnost pa bomo morali meritev ponavljati velikokrat. Tako s štirimi meritvami napako razpolovimo, če pa jo želimo zmanjšati za 10-krat, pa moramo opraviti 100 meritev. Tudi sicer je zmanjšanje statistične napake pod velikostni red sistematične nesmiselno.

3.3 Centralni limitni izrek

Še ena lepa lastnost odlikuje statistično napako, izvira pa iz centralnega limitnega izreka, ki velja za temeljni izrek statistike. Da se namreč pokazati, da gre porazdelitev povprečja statistično nekoreliranih dogodkov (v našem primeru meritev) proti normalni (reče se ji tudi Gaussova) porazdelitvi $P(x)$, ko gre število dogodkov proti neskončnosti:

$$P(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-x_0)^2}{2\sigma^2}}, \quad (19)$$

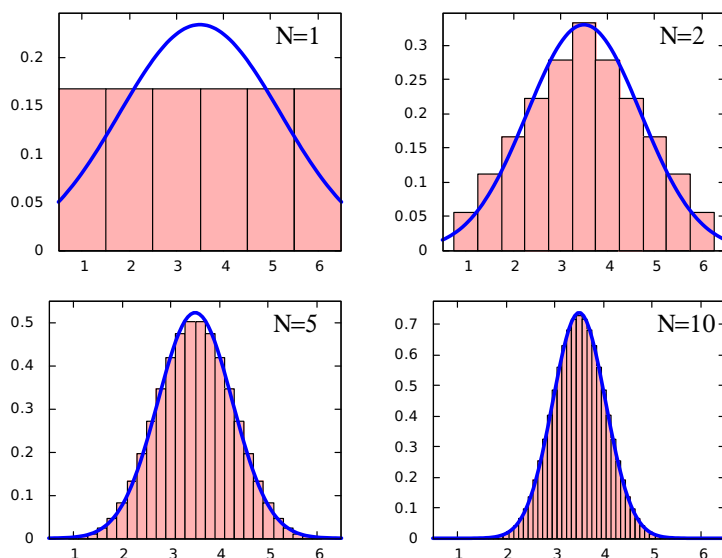
kjer je x_0 povprečje porazdelitve, σ pa standardna deviacija (v našem primeru je to statistična napaka).

Odstavek za zahtevnejše bralce: konvergenca pri centralnem limitnem izreku je mišljena kot konvergenca po distribuciji (meri) in ne kot konvergenca po točkah.

Statistična napaka je definirana kot koren povprečja kvadratov odstopanj $\sigma = \sqrt{\overline{\delta x^2}}$.

Statistična napaka pada z inverzom korena števila ponovitev meritve $\sigma_n = \frac{\sigma}{\sqrt{n}}$.

Povprečje več meritev konvergira h Gaussovi porazdelitvi.



Slika 1: Porazdelitev povprečja meta kocke pri enem, dveh, petih in dvajsetih metih. Zraven je prikazana Gaussova porazdelitev z isto srednjo vrednostjo (3,5) in standardno deviacijo (1,708, 1,208, 0,764 in 0,540).

Prvo pomeni, da konvergira povprečna porazdelitev na poljubno majhnem (ampak še vedno končnem) intervalu proti neki vrednosti, drugo pa, da konvergira vrednost porazdelitve v posamezni točki x .

Najlepše pokažemo centralni limitni izrek na primeru meta kocke. En sam met kocke je daleč od Gaussove porazdelitve. Verjetnost, da bo padla posamezna številka je enako $1/6$. Torej bo povprečje enega meta 3,5 s standardno deviacijo $\sigma = 1,708$:

$$\sigma^2 = \frac{1}{6} \sum_{i=1}^6 (i - 3,5)^2 = 2,917 \quad (20)$$

Sedaj pa nas zanima, kako je s povprečjem več metov kocke. Kocko bomo metali večkrat in izračunali povprečje več metov. Jasno je, da je v vseh primerih ne glede na število metov povprečje povprečne vrednosti enako 3,5. Standardna deviacija pa se zmanjšuje s številom meteov v skladu s formulo (18). Na sliki 1 vidimo, kako se s številom meta kock porazdelitev vse bolj približuje Gaussovi.

3.4 Pravilo dveh tretjin

Statistično napako σ ene meritve lahko ocenimo iz odstopanja rezultatov več meritev. Če imamo N meritev količine x , pri čemer je x_i i -ta meritev, lahko izračunamo povprečje teh meritev z

$$x_p = \frac{1}{N} \sum_{i=1}^N x_i. \quad (21)$$

Tako izračunano povprečje x_p seveda ne bo enako točni vrednosti x_0 , saj je tudi samo obremenjeno z napako σ_N , ki pa ne vemo, kolikšna je. Oceno za napako lahko dobimo iz

$$\sigma^2 \approx \frac{1}{N-1} \sum_{i=1}^N (x_i - x_p)^2. \quad (22)$$

Ne bomo se spuščali v razloge, zakaj je $N-1$ v imenovalcu in ne N , ima pa to nekaj za opraviiti z dejstvom, da odštevamo x_p od naših meritev in ne x_0 . Zahtevnejši bralec si lahko kaj več o tem pogleda v [1].

S takim postopkom sicer lahko določimo statistično napako, vendar je ta način zamuden. Obstaja namreč eno pravilo, ki nam delo olajša, je pa neposredna posledica centralnega limitnega izreka. Najprej izračunajmo verjetnost P , da je meritev

Statistično napako lahko dobimo tako, da poiščemo srednji interval, v katerega bo padlo $2/3$ meritev.

znotraj intervala med $x_0 - \sigma$ in $x_0 + \sigma$. To naredimo tako, da integriramo verjetnostno porazdelitev na tem intervalu. Pravo porazdelitev nadomestimo z Gaussovo in dobimo izraz

$$P = \int_{x_0 - \sigma}^{x_0 + \sigma} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-x_0)^2}{2\sigma^2}} dx. \quad (23)$$

Integral preuredimo z vpeljavo nove spremenljivke $u = (x - x_0)/\sigma$. Dobimo

$$P = \frac{1}{\sqrt{2\pi}} \int_{-1}^1 e^{-\frac{u^2}{2}} du. \quad (24)$$

Niti tega integrala se ne da izraziti z elementarnimi funkcijami (lahko ga sicer izrazimo s funkcijo erf). Vidimo pa, da izraz ni odvisen ne od x_0 ne od σ in je enak za vse Gaussove porazdelitve. Njegova vrednost je 0,682689. Ker se nam gre pri oceni napake le za njen velikostni razred, lahko to vrednost zaokrožimo na dve tretjini. To pomeni, da bo v eni seriji meritev približno 2/3 meritev znotraj intervala. Če bomo torej zavrgli 1/3 meritev, ki bodo najbolj odstopale od povprečja, bodo preostale meritve odstopale za največ σ od povprečja.

Poberi podatke tretje naloge s spletne strani. Z enim od urejevalnikov razpredelnic (MS Excel, LO Calc ...) izračunaj povprečje in oceni statistično napako po formulah (21) in (22). Slednjo izračunaj tudi s pravilom dveh tretjin. Razpredelnico natisni in jo priloži poročilu za vaje.

$x_0 =$ _____

$\sigma =$ _____ (po formuli)

$\sigma =$ _____ (pravilo dveh tretjin)

4 Slovnična opazka

Po slovenskem pravopisu (pa ne samo slovenskem) se med količino in enoto piše presledek npr. 3 m in ne 3m. To pravilo zavestno kršim, ker se z njim ne strinjam in se zavzemam za njegovo spremembo.

Literatura

- [1] I. Kuščer in A. Kodre. *Matematika v fiziki in tehniki*. DMFA, Ljubljana, 1994.
- [2] A. Likar. *Osnove fizikalnih merjenj in merilnih sistemov*. DMFA, Ljubljana, 1992.
- [3] M. Valič. *Fizikalne meritve*.