



ÖREBRO
UNIVERSITET

Institutionen för naturvetenskap och teknik

Konsten att bestämma π genom slumpen

En jämförelse mellan olika stokastiska processer som genererar π

Olivia Bergman

Konsten att bestämma π genom slumpen

En jämförelse mellan olika stokastiska processer som genererar π

Olivia Bergman

Maj 2018

Handledare: Niklas Eriksen

Examinator: Mårten Gulliksson

Sammanfattning

Syftet med detta arbete var att jämföra olika stokastiska processer som genererar π , de olika modellerna som studerades var modeller baserade på Buffons-, Buffon-Laplaces- och kasta pil-problem och förutom att lösa respektive problem så undersöktes bland annat om det gick att skatta π med motsvarande problem i högre dimensioner.

Innan problemen undersöktes fick läsaren en bakgrund med nödvändiga definitioner och satser inom sannolikheteori, där bland annat Kolmogorovs axiomsystem, fördelningars täthetsfunktion, väntevärde och varians presenterades, men också olika definitioner av konvergens som i sin tur ledde till en sats: Deltametoden, en generalisering av Centrala gränsvärdessatsen. Allt detta tillsammans gjorde det möjligt att lösa problemen och sedan kunna jämföra dem med hjälp av respektive skattnings varians.

Innehåll

1 Inledning	5
1.1 Konstanten πs historia	5
2 Bakgrund	8
2.1 Sannolikhetssteori	8
2.1.1 Grundläggande definitioner i sannolikhetsläran	9
2.1.2 Stokastiska variabler och dess egenskaper	10
2.1.3 Centrala gränsvärdessatsen och en generalisering	14
2.2 Statistisk inferens	19
2.2.1 Estimatorer	19
2.3 Buffons problem	20
2.4 Buffon-Laplaces problem	20
2.5 Kasta pil-problemet	21
2.6 Problemen i högre dimensioner	21
2.6.1 Buffons problem i 3 dimensioner	21
2.6.2 Kasta pil-problemet i 3 dimensioner	21
2.6.3 Kasta pil-problemet i k dimensioner	22
3 Resultat	23
3.1 Buffons problem	23
3.2 Buffon-Laplaces problem	25
3.3 Buffon-Laplace-kast som Buffon-kast	29
3.4 Kasta pil-problemet	32
3.5 Buffons problem i 3D	34
3.6 Kasta pil-problemet i 3 dimensioner	35
3.7 Kasta pil-problemet i k dimensioner	37
3.8 Jämförelse mellan modellerna	39
Referenslista	41

1 Inledning

Vad har slumpen gemensamt med konstanten π ? Ett irrationellt tal som har många miljoner kända decimaler och kan tyckas vara väldigt långt ifrån någon slump.

I slutet på 1700-talet kom en fransk naturforskare vid namn Buffon på att *om en nål kastas upprepade gånger på ett golv med parallella plankor där plankornas bredd är större än nålen och antalet gånger nålen korsar en linje noteras, så kan π uppskattas*. Problemet är känt som Buffons nålproblem och en skattning av π fås genom att besvara frågan: *Vad är sannolikheten att nålen korsar en linje?* [1].

Buffon vidareutvecklade sina idéer med parallella linjer och satte upp liknande problem med andra geometriska rutnätsmönster [2], så som exempelvis rektangulära-, hexagonala- och triangulära rutnätsmönster. Buffon publicerade sina lösningar på respektive problem i sin bok *Naturhistoria: allmänt och i detalj*, men Laplace upptäckte ett fel i beräkningarna för det rektangulära rutnätsmönstret och han rättade till det i början av 1800-talet [3]. Problemet går idag under namnet Buffon-Laplaces problem. Dessa problem klassas som så kallade *Monte Carlo-metoder*. Monte Carlo-metoder är metoder som löser matematiska problem genom experiment med slumpantal [4]. Hur ser lösningarna på Buffons och Buffon-Laplaces problem ut? Ett annat likvärdigt problem är en kvadratisk piltavla. Tavlan har en inskriven cirkel med radie r och om pilen träffar innanför cirkeln så räknas det som en träff. Även det här en Monte Carlo-metod som kan användas för att skatta π . Går det att generalisera problemen till högre dimensioner?

En fråga som faller sig naturligt när någonting skattas är hur stor spridningen för skattningen är, ju mindre spridning en skattning väntas ha desto bättre precision har skattningen. Hur ter sig variansen i problemen, är något av problemen att föredra med hänsyn till variansen? Vad händer med variansen om ett kast i Buffon-Laplace problem istället ses som två kast i Buffons problem?

Frågorna som ställts här i inledningen ska försöka besvaras i detta arbete, men innan dess behöver läsaren rustas med kunskap om grunderna i sannolikhetsteorin och om estimatorer (avsnitt 2) för att sedan kunna läsa resultatet med beräkningar (avsnitt 3.1–3.7) med behållning. Arbetet avslutas med en jämförelse mellan metoderna, hur de förhåller sig till varandra (Avsnitt 3.8).

1.1 Konstanten π s historia

Varför är just π intressant? Definitionen av π är förhållandet mellan en cirkels omkrets och dess diameter. Talet är irrationellt utan någon regelbundenhet i decimalutvecklingen. Det speciella med π är att talet dyker upp på ställen som inte alltid uppenbart har kopplingar till talets geometri, exempelvis i normalfördelningens täthetsfunktion och i flertalet oändliga summor och produkter. Fascinationen av π har funnits länge, redan i Bibeln finns det referenser till att π uppskattas till tre. Därefter försökte Babylonierna och

Egypterna på varsina håll finna skattningar på π , $3 + \frac{1}{8} = 3.125$ respektive $\frac{2^8}{3^4} = 3.1605$ på vissa håll och annars återfinns approximationen $\frac{22}{7} \approx 3,1429$ från Egypterna [5]. Genom århundraden har det upptäckts mängder med formler som kan approximera π , ett exempel är formeln som Newton fann på 1600-talet [6]

$$\pi = \frac{3}{4}\sqrt{3} + 24 \int_0^{\frac{1}{4}} \sqrt{x - x^2} dx.$$

Att försöka hitta approximationer på π har många matematiker försökt göra under flera hundratals år. Arkimedes försökte genom att stänga in en cirkel mellan två polygoner för att på så sätt få en undre- och övre gräns på π . Genom att använda två 96-goner (en regelbunden månghörning med 96 hörn), fick Arkimedes fram att π ligger i mellan $3 + \frac{10}{71} < \pi < 3 + \frac{1}{7}$. En annan metod är just Buffons problem, och liknande problem [5].

Efter Arkimedes, Egypterna och alla andra som sökt svaret på vad π är, har matematikerna fortsatt jaga decimaler på π genom att hitta nya summor och formler. Från 1700-talet och framåt har formler med arctangens använts, 1706 var det en matematiker vid namn Machin som med följande formler lyckades bestämma de 100 första decimalerna på π [7],

$$\frac{\pi}{4} = 4 \arctan \frac{1}{5} - \arctan \frac{1}{239}.$$

I början på 1900-talet fann den Indiska matematikern Ramanujan serien nedan. Serien konvergerar snabbt till π , där varje term ger hela åtta decimaler [7].

$$\frac{1}{\pi} = \frac{\sqrt{8}}{9801} \sum_{n=0}^{\infty} \frac{(4n)! (1103 + 26390n)}{(n!)^4 396^{4n}}.$$

Under 1900-talet har antal kända decimaler ökat explosionsartat. I november 2016 slogs det tidigare världsrekordet på antal decimaler på π från 2013. Det var π -fantasten Peter Trueb som byggt en dator som klarade av att bestämma drygt 22,4 biljoner decimaler och det tog drygt 100 dagar för datorn att göra beräkningarna. Datorn krävde enormt mycket minne och till sin hjälp användes ett program vid namn γ -cruncher framtaget av Alexander Yee [8]. Programmet bygger på Chudnovskys algoritmen som utvecklades under 1980-talet och algoritmen bygger på tre delar [9], Heegnernummer $d = -163$, j -funktionen $j\left(\frac{1+\sqrt{-163}}{2}\right) = -640320^3$ samt den generaliserade hypergeometriska serien

$$\frac{1}{\pi} = 12 \sum_{n=0}^{\infty} \frac{(-1)^n (6n)! (545140134n + 13591409)}{(3n)! (n!)^3 (650320)^{3n+3/2}}.$$

Med dessa tre delar kan ett närmevärde på π beräknas.

Att fortsätta söka efter decimaler på π har inte med att matematiker vill få noggrannare beräkningar utan det handlar om att klara av att bygga ännu starkare datorer med hög datakraft. Exempelvis NASA använder omkring 15 decimaler vid sina beräkningar och om jordens omkrets ska beräknas med en felmarginal på mindre än en atomkärnas diameter behövs bara cirka 40 decimaler på π [8].

2 Bakgrund

2.1 Sannolikhetsteori

Sannolikhetsteorin har en lång historia som sträcker sig tillbaka ända till 1500-talets Italien där matematikern Cardano skrev på en skrift om *hasardspel*, spel som beror av slumpen. Teorin om dessa spel bygger på sannolikhetsproblem och har efterhand byggts upp till en hel teori som inte bara innefattar hasardspel. I Frankrike på 1600-talet var det också just spel som fick matematikerna Pascal och Fermat att brevväxla för att försöka lösa problem som berör sannolikhetsteorin. På 1700- och 1800-talet bidrog matematiker som Bernoulli, de Moivre och Laplace till att sannolikhetsteorin utvecklades ännu mer och bidrog till att utvidga tillämpningar av området till bland annat astronomi och mekanik.

Som synes är sannolikhetsteorin gammal, men har matematiska svagheter. Begreppen *händelser* och *sannolikhet* definierades aldrig matematiskt vilket ledde till att ett gäng matematiker omformulerade den *klassiska sannolikhetsteorin* till den *moderna* i början av 1900-talet. Kolmogorov upptäckte att med hjälp av teori från Cantors mängdteori och Lebesgues måtteori kunde begreppet sannolikhetsrum definieras och beskriva den klassiska sannolikhetsteorin. Värt att nämna är att andra områden inom matematiken utvecklades under slutet av 1800-talet så som måtteorin och mängdläran [10], så att tidigare än så var det svårt att utveckla sannolikhetsteorin. Kolmogorov formulerade tre axiom som satte fart på utvecklingen inom området och med bidrag av många matematiker runt om i världen har en rad många nya tillämpningsområden tillkommit, bland annat inom bioinformatik, bildbehandling och finansiell matematik och utvecklingen av hela teorin pågår än idag [11].

Följande kapitel kommer att behandla *grunderna i sannolikhetsteorin* med nödvändiga definitioner och satser för att kunna gå vidare till arbetets syfte – Buffons problem med flera. Efter att grunderna har behandlas tas en av sannolikhetsteorins viktigaste sats upp – *Centrala gränsvärdessatsen* (CGS). I samma avsnitt presenteras olika *konvergensbegrepp* för att slutligen kunna bevisa *Deltametoden*, en generalisering av CGS som kommer behövas för att beräkna variansen till skattningarna för de olika problemen som tas upp i arbetet. Efter sannolikhetsteorin kommer ett kortare avsnitt om *statistisk inferens* där *estimatorer* diskuteras, innan kapitlet avslutas med att de olika problemen som arbetet syftar att lösa presenteras.

2.1.1 Grundläggande definitioner i sannolikhetsläran

Definition 2.1.1.1. Resultatet av ett slumpmässigt försök kallas ett *utfall*.

Mängden av möjliga utfall kallas *utfallsrummet*. En *händelse* är en samling av utfall.

Ett centralt begrepp i en modell för ett slumpmässigt försök är begreppet *sannolikhet* och hur begreppet definieras rent matematiskt. I början av 1900-talet formulerade en rysk matematiker Kolmogorov följande axiom, som är grunden för den moderna sannolikhetsteorin [4].

Kolmogorovs tre axiom

1. (Icke-negativitet). För en godtycklig händelse $A \subseteq \Omega$, där Ω är utfallsrummet, gäller att $P(A) \in \mathbb{R} \wedge P(A) \geq 0$.
2. (Normalisering). För hela utfallsrummet Ω gäller att $P(\Omega) = 1$.
3. (Additionsformeln). Om A_1, A_2, \dots är en ändlig eller uppräknelig oändlig följd av händelser och $A_i \cap A_j = \emptyset \forall i \neq j$, gäller att

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

Anmärkning: I axiom 1 räcker det med kravet att sannolikheter är icke-negativa, för att få att sannolikheter dessutom begränsas till att vara mindre eller lika med 1 följer av axiomen, se följsatserna nedan.

Från Kolmogorovs axiom fås viktiga följsatser.

Lemma 2.1.1.1. Om $A \subseteq B$ gäller att

$$P(A) \leq P(B).$$

Bevis. Mängden A är en delmängd till B , så mängden B kan skrivas som $B = A \cup (B \setminus A)$. Kolmogorovs axiom 3 ger då följande,

$$P(A \cup (B \setminus A)) = P(A) + P(B \setminus A)$$

$$P(B) = P(A) + P(B \setminus A)$$

$$P(A) = P(B) - P(B \setminus A)$$

och enligt Kolmogorovs axiom 1 fås att samtliga sannolikheter är positiva, och då följer det att $P(A) \leq P(B)$.

□

Sats 2.1.1.1. För en händelse $A \subseteq \Omega$ gäller att

$$0 \leq P(A) \leq 1.$$

Bevis. Enligt Kolmogorovs axiom 2 fås att $P(\Omega) = 1$ och enligt lemma 2.1.1.1. fås att $P(A) \leq P(\Omega)$. Tillsammans med axiom 1, $P(A) \geq 0$, fås att $0 \leq P(A) \leq P(\Omega) = 1$ vilket bevisar satsen.

□

Sats 2.1.1.2. (Komplementsatsen). För komplementet A^* till A gäller att

$$P(A^*) = 1 - P(A).$$

Bevis. Komplementet A^* är de händelserna som inte finns i A , $A^* \cap A = \emptyset$ och $A^* \cup A = \Omega$. Enligt Kolmogorovs axiom 2 och 3 fås att $P(A^*) + P(A) = P(A^* \cup A) = P(\Omega) = 1$. Då följer att $P(A^*) = 1 - P(A)$ och satsen är bevisad.

□

Anmärkning: Det finns flera olika notationer på komplementet, andra vanligt förekommande notationer är A^C och \bar{A} .

2.1.2 Stokastiska variabler och dess egenskaper

Definition 2.1.2.1. En funktion $P(\cdot)$ kallas för ett *sannolikhetsmått* om Kolmogorovs tre axiom är uppfyllda.

Utfallsrummet Ω , händelserna A_1, A_2, \dots och sannolikhetsmättet $P(\cdot)$ säges utgöra ett *sannolikhetsrum*.

Definition 2.1.2.2. En slumpvariabel X är *diskret* om den endast kan anta ett ändligt eller uppräknelig oändligt antal värden x_1, x_2, \dots

För att beskriva hur sannolika de olika utfallen för en diskret slumpvariabel är, används dess så kallade *sannolikhetsfunktion*, alltså sannolikheten för att respektive utfall ska inträffa.

Definition 2.1.2.3. Sannolikhetsfunktionen, p_X , för en diskret slumpvariabel X definieras av

$$p_X(x) := P(X = x), \quad \text{där } x = x_1, x_2, \dots$$

Definition 2.1.2.4. Om den stokastiska variabeln X har sannolikhetsfunktion

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \dots, n,$$

och där n är ett positivt heltal och att $0 < p < 1$ sägs X vara *binomialfördelad*.

Om en stokastisk variabel X är binomialfördelad med parametrar n och p , används notationen $X \sim \text{Bin}(n, p)$.

Stokastiska variabler har vissa egenskaper kopplat till deras sannolikhetsfördelning, dels väntevärde som kan tolkas som ett försöks

medelvärde om försöket upprepas ett oändligt antal och dels variansen som beskriver hur mycket de observerade värdena avviker från medelvärdet.

Definition 2.1.2.5. Väntevärdet μ och variansen σ^2 för en diskret stokastisk variabel X definieras som

$$\mu = E(X) = \sum_{k=1}^n k \cdot p_X(k),$$

$$\sigma^2 = V(X) = E((X - E(X))^2)$$

Sats 2.1.2.1. En stokastisk variabel X som är binomialfördelad, $X \sim \text{Bin}(n, p)$, har väntevärde $E(X) = np$ och varians $V(X) = np(1 - p)$.

Bevis. Kom ihåg att sannolikhetsfunktionen för en binomialfördelad variabel X är $p_X(k) = \binom{n}{k} p^k (1 - p)^{n-k}$, $k = 0, 1, 2, \dots, n$, så att väntevärdet för X fås som

$$E(X) = \sum_{k=1}^n k \cdot \binom{n}{k} p^k (1 - p)^{n-k}.$$

Termen $k = 0$ kommer inte ge något bidrag till summan, så den kan bortses ifrån.

$$k \binom{n}{k} = k \frac{n!}{k! (n-k)!} = k \frac{n!}{k \cdot (k-1)! (n-k)!} = \frac{n \cdot (n-1)!}{(k-1)! (n-k)!}$$

$$= n \binom{n-1}{k-1}$$

Notera att för sannolikhetsfunktionen summerar till ett, ty

$$\sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} = (p + (1-p))^n = 1^n = 1$$

Genom att bryta ut ett p ur summan och skifta index ett steg nedåt fås väntevärdet för en $\text{Bin}(n-1, p)$ -variabel, så summan nedan summerar till ett.

$$E(X) = np \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{n-1-k} = np$$

Variansen kan beräknas enligt $V(X) = E(X^2) - (E(X))^2$, ty

$$V(X) = E((X - E(X))^2)$$

$$= E(X^2 - 2XE(X) + (E(X))^2)$$

$$= E(X^2) - 2E(X)E(X) + E((E(X))^2)$$

$$= E(X^2) - (E(X))^2.$$

För att härleda variansen behövs alltså $E(X^2)$.

$$\begin{aligned} E(X^2) &= \sum_{k=1}^n k^2 \cdot \binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{k=1}^n (k^2 - k + k) \cdot \binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{k=1}^n k(k-1) \cdot \binom{n}{k} p^k (1-p)^{n-k} + \sum_{k=1}^n k \cdot \binom{n}{k} p^k (1-p)^{n-k} \end{aligned}$$

Den senare summan blir enligt ovan np och på liknande sätt kan det visas att den första summan blir $n(n-1)p^2$ men det lämnas som övning till läsaren, så att variansen för en binomialfördelad variabel blir $V(X) = n(n-1)p^2 + np - (np)^2 = n((n-1)p^2 + p - np^2) = np(1-p)$.

□

Ibland kan stokastiska variabler, så kallade *kontinuerliga* stokastiska variabler, anta värden på ett intervall. Variabelns olika utfall ligger då oändligt tätt och sannolikheten att ett givet utfall ska inträffa kommer vara noll. Alltså kommer det inte existera någon sannolikhetsfunktion. Teori och praktik går inte alltid hand i hand, för i praktiken kommer det inte existera några riktiga kontinuerliga stokastiska variabler, eftersom att oavsett hur noggranna mätningar som görs kommer de möjliga utfallen vara uppräknligt många. Trots det är det väldigt användbart att se att utfallen utgör ett helt kontinuum av värden, och anledningen till det är att det oftast är lättare att integrera än att summera [11].

Definition 2.1.2.6. En stokastisk variabel X sägs vara kontinuerlig om det finns en funktion $f_X(x)$ så att det för alla mängder A gäller att

$$P(X \in A) = \int_A f_X(t) dt.$$

Funktionen $f_X(\cdot)$ kallas för den stokastiska variabelns *täthetsfunktion*.

Kontinuerliga stokastiska variabler följer olika *sannolikhetsfördelningar*, ibland bara fördelning. En sannolikhetsfördelning är som en slags beskrivning, ofta en funktion, av hur sannolika de olika utfallen i ett givet utfallsrum är.

Om alla utfall i ett utfallsrum är lika sannolika, sägs den stokastiska variabeln följa en *likformig fördelning*.

Definition 2.1.2.7. Om den stokastiska variabeln X har täthetsfunktion

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{om } a < x < b \\ 0 & \text{annars} \end{cases}$$

sägs X vara *likformigt fördelad*.

I slumpmässiga försök där fler än en storhet är av intresse kallas variablerna *flerdimensionella stokastiska variabler*.

Definition 2.1.2.8. En n -dimensionell stokastisk variabel är en funktion (X_1, X_2, \dots, X_n) definierad på ett utfallsrum Ω och som antar värden i \mathbb{R}^n .

Precis som i det endimensionella fallet så finns det täthetsfunktioner för flerdimensionella stokastiska variabler.

Definition 2.1.2.9. Om det finns en funktion $f_{X_1, \dots, X_n}(x_1, \dots, x_n)$ sådan att

$$P((X_1, \dots, X_n) \in A) = \int \dots \int_A f_{X_1, \dots, X_n}(x_1, \dots, x_n) dx_1 \dots dx_n$$

för alla A sägs den n -dimensionella stokastiska variabeln (X_1, X_2, \dots, X_n) vara kontinuerlig och $f_{X_1, \dots, X_n}(x_1, \dots, x_n)$ är den simultana täthetsfunktionen för (X_1, X_2, \dots, X_n) .

Om variablerna är oberoende fås den simultana täthetsfunktionen som produkten av respektive variablers täthetsfunktion $f_{X_1, \dots, X_n}(x_1, \dots, x_n) = f_{X_1}(x_1) \cdot \dots \cdot f_{X_n}(x_n)$.

I många situationer finns det ett intresse att veta hur en följd, en *stokastisk process*, av stokastiska variabler beter sig vid upprepade slumpförsök, och genom att modellera en viss händelse som en stokastisk variabel med lämplig fördelning kan modellen användas för att analysera händelsen beroende på vad som är av intresse. Om ett slumpförsök upprepas tillräckligt många gånger för en stokastisk variabel, kommer det aritmetiska medelvärdet av dessa observationer med stor sannolikhet ligga nära stokastiska variabelns väntevärde, enligt lagen om *de stora talens lag*. Uttrycket ”det jämnar ut sig i det långa loppet” sägs under vissa omständigheter motsvaras av lagen. Men innan behövs ett resultat, *Chebyshevs olikhet*, för att kunna bevisa satsen [4].

Lemma 2.1.2.1. (Chebyshevs olikhet). Låt X vara en stokastisk variabel med ändligt väntevärde μ och standardavvikelse σ . Då gäller, för varje $a > 0$, att

$$P(|X - \mu| \geq a) \leq \frac{\sigma^2}{a^2}.$$

Bevis. Anta att X är kontinuerlig. Från definitionen av varians fås att

$$\begin{aligned}\sigma^2 &= \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx \\ &= \int_{|x-\mu|<a} (x - \mu)^2 f(x) dx + \int_{|x-\mu|\geq a} (x - \mu)^2 f(x) dx \\ &\geq 0 + a^2 \int_{|x-\mu|\geq a} f(x) dx \\ &= a^2 P(|X - \mu| \geq a).\end{aligned}$$

vilket bevisar olikheten. □

Sats 2.1.2.2. (De stora talens lag). Låt X_1, X_2, \dots vara en följd av inbördes oberoende stokastiska variabler. Om väntevärdet är $\mu = E(X_i)$ och standardavvikelsen $\sigma = \sqrt{V(X)}$ är ändligt. Låt $\bar{X}_n = \sum_{i=1}^n X_i/n$ är medelvärdet av de n första variablerna. Då gäller för godtyckligt $\varepsilon > 0$ att

$$\lim_{n \rightarrow \infty} P(|\bar{X}_n - \mu| > \varepsilon) \rightarrow 0, \quad \text{då } n \rightarrow \infty.$$

Bevis. Betrakta den stokastiska variabeln \bar{X}_n . Dess medelvärde respektive standardavvikelse är $E(\bar{X}_n) = \mu$ och $\sqrt{V(\bar{X}_n)} = \frac{\sigma}{\sqrt{n}}$. Låt $\varepsilon > 0$ vara fixt. Med Chebyshevs olikhet (lemma 2.1.2.1) fås att

$$P(|\bar{X}_n - \mu| > \varepsilon) \leq \frac{V(\bar{X}_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}$$

och då $n \rightarrow \infty$, kommer $P(|\bar{X}_n - \mu| > \varepsilon)$ gå mot noll, vilket bevisar satsen. □

Korollarium 2.1.2.1. (Bernoullis sats). Betrakta en följd oberoende upprepningar av ett försök där händelsen A inträffar med sannolikhet $P(A)$ i ett enskilt försök. Då konvergerar den relativa frekvensen för A mot $P(A)$ då antalet försök går mot oändligheten.

Bevis. Låt $X_k := \chi_k(A)$, alltså en indikator för händelsen A i försök k . Då är $\mu = E(X_k) = E(\chi_k(A)) = 1 \cdot P(A) + 0 \cdot P(A^*) = P(A)$ och \bar{X}_n är den relativa frekvensen. Stora talens lag (sats 2.1.2.2) bevisar därför satsen. □

2.1.3 Centrala gränsvärdessatsen och en generalisering

Ett av sannolikhetssteorins viktigaste resultat, är *Centrala gränsvärdessatsen* [4]. I De stora talens lag (sats 2.1.2.2) visades att medelvärdet för en följd av stokastiska variabler konvergerar i sannolikhet mot väntevärdet då $n \rightarrow \infty$, om variablerna har samma medelvärde och standardavvikelse. Nu läggs fokus på hur

fördelningen för $\bar{X}_n - \mu$ blir när n växer, under antagandet att alla variablerna följer samma fördelning.

Sats 2.1.3.1. (Centrala gränsvärdessatsen). Låt X_1, X_2, \dots vara oberoende och likafördelade stokastiska variabler med $E(X_i) = \mu$ och att standardavvikelsen $\sqrt{V(X)} = \sigma$ är ändligt. Låt sedan $\bar{X}_n := \sum_{i=1}^n X_i/n$. För godtyckliga $a < b$ gäller då att

$$P\left(a < \frac{\sqrt{n}}{\sigma} (\bar{X}_n - \mu) < b\right) \rightarrow \Phi(b) - \Phi(a), \text{ då } n \rightarrow \infty$$

där Φ är fördelningsfunktionen för $N(0,1)$.

Bevis. För ett bevis av centrala gränsvärdessatsen hänvisas läsaren till någon bok i sannolikhets teori, förslagsvis [4].

Centrala gränsvärdessatsen är mycket användbar, då det följer att om n är ”tillräckligt stor” kommer medelvärdet \bar{X}_n approximativt vara normalfördelat som $\bar{X}_n \stackrel{\text{appr}}{\sim} N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$ [4]. Ett annat sätt att uttrycka det på är att säga att $\sqrt{n}(\bar{X}_n - \mu)$ konvergerar i fördelning till $N(0, \sigma)$ [12].

Definition 2.1.3.1. En sekvens av stokastiska variabler X_1, X_2, X_3, \dots sägs konvergera i fördelning till en stokastisk variabel X om

$$\lim_{n \rightarrow \infty} F_{X_n}(x) = F_X(x),$$

i alla punkter x där $F_X(x)$ är kontinuerlig.

En starkare form av konvergens är konvergens i sannolikhet.

Definition 2.1.3.2. En sekvens av stokastiska variabler X_1, X_2, X_3, \dots sägs konvergera i sannolikhet till en stokastisk variabel X om, för varje $\varepsilon > 0$, gäller att

$$\lim_{n \rightarrow \infty} P(|X_n - X| < \varepsilon) = 1.$$

Konvergens i sannolikhet implicerar alltid konvergens i fördelning, men omvändningen gäller bara i specialfall [12].

Sats 2.1.3.2. Sekvensen av stokastiska variabler X_1, X_2, X_3, \dots , konvergerar i sannolikhet till en konstant μ om och endast om sekvensen också konvergerar i fördelning till μ .

Sats 2.1.3.3. Antag att X_1, X_2, X_3, \dots konvergerar i sannolikhet till en stokastisk variabel X och att g är en kontinuerlig funktion. Då kommer $g(X_1), g(X_2), g(X_3), \dots$ konvergera i sannolikhet till $g(X)$.

Bevis. För ett givet $\varepsilon > 0$ finns det ett $\delta > 0$ så att

$$|X_n - X| < \delta \implies |g(X_n) - g(X)| < \varepsilon.$$

Enligt antagande fås att $P\left(\lim_{n \rightarrow \infty} |X_n - X| < \delta\right) = 1$, vilket ger att $P\left(\lim_{n \rightarrow \infty} |g(X_n) - g(X)| < \varepsilon\right) = 1$ och därmed är satsen bevisad. □

Definitionerna av konvergens tillsammans med Slutskys sats, som presenteras på nästa sida, leder till en användbar generalisering av centrala gränsvärdessatsen.

Sats 2.1.3.4. (Slutskys sats). *Om $X_n \rightarrow X$ i fördelning och $Y_n \rightarrow c$ i sannolikhet, där c är en konstant, då gäller*

- (I) $Y_n X_n \rightarrow cX$ i fördelning
- (II) $X_n + Y_n \rightarrow X + c$ i fördelning.

Bevis. För fullständigt bevis vänligen se exempelvis [13]. Beviset bygger på att visa att (X_n, Y_n) konvergerar i fördelning till (X, c) för sedan använda en sats (som går under namnet *continuous mapping theorem*) som säger att om X_n konvergerar i fördelning till X kommer $g(X_n)$ konvergera till $g(X)$ i fördelning om g är en kontinuerlig funktion.

Den senare satsen kan då användas genom att sätta funktionen till $g(x, y) = x + y$ och $g(x, y) = xy$, som då kommer att bevisa Slutskys sats. □

Centrala gränsvärdessatsen är inte alltid möjlig att använda eftersom den inte säger någonting om hur sekvensen av stokastiska variabler beter sig om sekvensen består av en funktion $g(X_n)$ och inte bara av X_n . Något mer behövs:

Deltametoden. Deltametoden bygger på att funktionen g Taylorutvecklas kring μ till första ordningen [12].

Sats 2.1.3.5. (Deltametoden). *Låt X_n vara en sekvens av stokastiska variabler som uppfyller att $\sqrt{n}(X_n - \mu) \rightarrow N(0, \sigma)$ i fördelning. För en given funktion g och för ett specifikt värde på μ , anta att $g'(\mu)$ existerar och är nollskild. Då gäller att*

$$\sqrt{n}[g(X_n) - g(\mu)] \rightarrow N(0, g'(\mu)\sigma)$$

konvergerar i fördelning.

Bevis. Genom att använda *Medelvårdessatsen* (se [14]) fås att (d.v.s. en Taylorutveckling av $g(X_n)$ kring μ) ger att

$$g(X_n) = g(\mu) + g'(\tilde{\mu})(X_n - \mu)$$

för något $\tilde{\mu}$ som ligger mellan $X_n < \tilde{\mu} < \mu$.

Enligt antagandet konvergerar $X_n \rightarrow \mu$ i fördelning, och sats 2.1.3.2. ger då att $X_n \rightarrow \mu$ även konvergerar i sannolikhet. Eftersom $\tilde{\mu}$ stängs in mellan X_n och μ kommer även $\tilde{\mu} \rightarrow \mu$ i sannolikhet. Enligt sats 2.1.3.3. fås att $g(\tilde{\mu}) \rightarrow g(\mu)$ i sannolikhet.

Taylorutvecklingen kring μ kan skrivas om och multipliceras med \sqrt{n} till

$$\sqrt{n}[g(X_n) - g(\mu)] = g'(\tilde{\mu}) \cdot \sqrt{n}(X_n - \mu)$$

Eftersom $\sqrt{n}(X_n - \mu) \rightarrow \sqrt{n}(X - \mu)$ i fördelning enligt antagande och $g(\tilde{\mu}) \rightarrow g(\mu)$ i sannolikhet kan del (I) i Slutskys sats (sats 2.1.3.4) användas, så att

$$g'(\tilde{\mu}) \cdot \sqrt{n}(X_n - \mu) \rightarrow g'(\mu) \cdot \sqrt{n}(X - \mu)$$

i fördelning, det vill säga

$$\sqrt{n}[g(X_n) - g(\mu)] \rightarrow g'(\mu) \cdot \sqrt{n}(X - \mu)$$

konvergerar i fördelning, vilket bevisar satsen. □

Som redan nämnt bygger Deltametoden på Taylorutveckling, och variansen går att härleda fram. Nedan visas härledningen för fallet med flera variabler.

Anta att $\mathbf{X} = X_1, \dots, X_k$ är stokastiska variabler med medelvärde $\boldsymbol{\mu} = \mu_1, \dots, \mu_k$. Anta även att det finns en deriverbar funktion $g(\boldsymbol{\mu})$, alltså en estimator för någon parameter och syftet är att bestämma en approximation av estimatorns varians.

Låt

$$g'_i(\boldsymbol{\mu}) = \frac{\partial}{\partial x_i} g(\mathbf{x})|_{x_1=\mu_1, \dots, x_k=\mu_k}$$

Första ordningens Taylorutveckling av g kring $\boldsymbol{\mu}$ ger

$$g(\mathbf{x}) = g(\boldsymbol{\mu}) + \sum_{i=1}^k g'_i(\boldsymbol{\mu})(x_i - \mu_i) + \text{rest}$$

I och med den statistiska användningen av polynomet går det att tänka att resten är försumbar [12]. Eftersom syftet är att finna variansen för funktionen, är det lämpligt att börja med att ta fram funktionens väntevärde.

$$E(g(\mathbf{X})) \approx g(\boldsymbol{\mu}) + \sum_{i=1}^k g'_i(\boldsymbol{\mu}) E((X_i - \mu_i)) = g(\boldsymbol{\mu}),$$

eftersom X_i har medelvärde μ_i . Enligt definition 2.1.2.5. fås variansen som

$$\begin{aligned} V(g(\mathbf{X})) &\approx E\left((g(\mathbf{X}) - g(\boldsymbol{\mu}))^2\right) \\ &\approx E\left(\left(\sum_{i=1}^k g'_i(\boldsymbol{\mu}) E((X_i - \mu_i))\right)^2\right) \\ &= \sum_{i=1}^k (g'_i(\boldsymbol{\mu}))^2 V(X_i) + 2 \sum_{i>j} g'_i(\boldsymbol{\mu}) g'_j(\boldsymbol{\mu}) \text{Cov}(X_i, X_j) \end{aligned}$$

där $\text{Cov}(\cdot)$ är kovariansen, och sista steget fås av variansoperatorns räkneregler för linjärbidrag, se [4].

Resultatet stämmer med Deltametoden, för om $k = 1$, så att det bara är en variabel, kommer andra summan ge bidrag noll. Med ovanstående härledning fås att väntevärde och varians för X fås som $E(g(X)) \approx g(\mu)$ och $V(g(X)) \approx (g'(\mu))^2 V(X)$.

För en kvot mellan två stokastiska variabler X_1 och X_2 fås med ovanstående härledning av variansen som

$$E\left(\frac{X_1}{X_2}\right) \approx \frac{\mu_{X_1}}{\mu_{X_2}}$$

Där μ_{X_1} och μ_{X_2} är väntevärdet för respektive variabel och kvoten skattas med funktionen $g(\mu_{X_1}, \mu_{X_2}) = \frac{\mu_{X_1}}{\mu_{X_2}}$.

De partiella derivatorna för funktionen blir följande,

$$\begin{aligned} \frac{\partial}{\partial X_1} g(\mu_{X_1}, \mu_{X_2}) &= \frac{1}{\mu_{X_2}} \\ \frac{\partial}{\partial X_2} g(\mu_{X_1}, \mu_{X_2}) &= -\frac{\mu_{X_1}}{\mu_{X_2}^2}. \end{aligned}$$

Från detta fås, med härledningen ovan, att

$$\begin{aligned} V\left(\frac{X_1}{X_2}\right) &\approx \frac{1}{\mu_{X_2}^2} V(X_1) + \frac{\mu_{X_1}^2}{\mu_{X_2}^4} V(X_2) + 2 \frac{1}{\mu_{X_2}} \left(-\frac{\mu_{X_1}}{\mu_{X_2}^2}\right) \text{Cov}(X_1, X_2) \\ &= \left(\frac{\mu_{X_1}}{\mu_{X_2}}\right)^2 \left(\frac{V(X_1)}{\mu_{X_1}^2} + \frac{V(X_2)}{\mu_{X_2}^2} + 2 \frac{\text{Cov}(X_1, X_2)}{\mu_{X_1} \mu_{X_2}}\right). \end{aligned}$$

Notera att om X_1 är konstant, fås att $V(X_1) = 0$ samt att $\text{Cov}(X_1, X_2) = 0$.

Det kommer visa sig senare i arbetet att variansen för funktionen $f(\mu) = \frac{1}{\mu}$ behövs, låt funktionen estimeras med $f(X) = \frac{1}{X}$, $f'(X) = -\frac{1}{X^2}$. Då fås enligt ovan att $E\left(\frac{1}{X}\right) \approx \frac{1}{\mu}$ och $V\left(\frac{1}{X}\right) \approx \left(\frac{1}{\mu}\right)^4 V(X)$.

2.2 Statistisk inferens

Sannolikhets teorin syftar till att utifrån en given sannolikhetsmodell försöka förutse utfallen i olika slumpförsök eller bestämma med vilken sannolikhet olika händelser kommer att inträffa. För statistisk inferens gäller det omvända, utifrån ett eller flera slumpförsök önskas information om den underliggande sannolikhetsmodellen.

2.2.1 Estimatorer

Att skatta, eller estimeras, olika parametrar är ett sätt att få information om parametrarna. Att skatta π handlar om att försöka få en uppfattning om konstantens storlek, alltså att bestämma så många korrekta decimaler som möjligt. Skillnaden mot att skatta en (okänd) parameter och π är att i det senare fallet finns det ett känt svar, det är lätt att kontrollera om skattningarna är bra vilket det inte alltid är vid annan form av inferens. Så syftet med att hitta stokastiska processer som kan ge skattningar på π är inte att försöka slå rekord i antal kända decimaler, det finns betydligt mer effektiva metoder för det, utan det handlar om matematikerns vetgirighet och att det går att skatta π med hjälp av slumpen [4].

Definition 2.2.1.1. En skattning, $\hat{\theta} = \hat{\theta}(X)$, av θ är en funktion av stickprovet X . Skattningen, eller *estimatet*, är en observation av *estimatorn* $\hat{\theta}(X)$.

Vid en jämförelse mellan olika skattningar är det två delar som avgör vilken skattning som är "bäst". En bra skattning skattar korrekt värde på parametern och skattningens slumpmässiga variation är liten.

Definition 2.2.1.2. Om $\hat{\theta}_1$ och $\hat{\theta}_2$ är *väntevärdesriktiga* skattningar, det vill säga $E(\hat{\theta}) = \theta$ för alla möjliga parameterskattningar, och $V(\hat{\theta}_1) \leq V(\hat{\theta}_2)$ för alla möjliga skattningar och att det för något θ är strikt olikhet, sägs $\hat{\theta}_1$ vara *effektivare* än $\hat{\theta}_2$.

Alltså, ju mindre variansen är, desto effektivare (bättre) kommer skattningen vara.

Som tidigare nämnt klassas problem som Buffons problem som Monte Carlo-metoder. Dessa metoder har som fördel att de är lätta att implementera med simuleringar i datorer, men de *konvergerar långsamt*. I exempelvis Buffons problem ger en miljon upprepningar bara omkring 4 korrekta decimaler [7]. Så alltså, det krävs ett stort antal försök för att få en någorlunda bra skattning på π med dessa metoder.

2.3 Buffons problem

Som nämnts i inledningen formulerade Buffon ett problem vars lösning kan användas för att skatta π . Problemet kan formuleras som [15]:

”Anta att vi har ett mönster av ekvidistanta parallella linjer (se figur 1). Låt b vara bredden mellan två parallella linjer och $\ell \leq b$ är nålens längd. Nålens position kan bestämmas genom avståndet y från nålens mittpunkt till den närmsta linjen samt den spetsiga vinkeln θ som skapats mellan nålen och en linje genom nålens mittpunkt som är parallell med de övriga linjerna. Vad är sannolikheten att nålen korsar en av linjerna?”



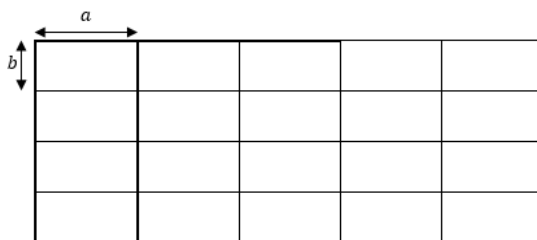
Figur 1 Mönstret med ekvidistanta parallella linjer som avses i Buffons problem, genom att beräkna sannolikheten att en nål med längd $\ell \leq b$ landar på mönstret så att nålen korsar en linje kan π skattas genom upprepade försök.

Genom att upprepa försöket kommer relativa frekvensen av antal träffar att, i enlighet med Bernoullis sats (korollarium 2.1.2.1), gå mot den sanna sannolikheten att nålen träffar linjen. Den teoretiska sannolikheten för att nålen ska träffa en linje beräknas i avsnitt 3.1.

2.4 Buffon-Laplaces problem

Buffon fortsatte att formulera problem som kan användas till skattningar av π . Ett annat problem är följande:

”Anta att vi har ett mönster av kongruenta rektanglar (se figur 2) med sidorna a och b och anta vidare att $\ell \leq \min(a, b)$ är nålens längd. Vad är sannolikheten att nålen korsar minst en av linjerna?”



Figur 2 Mönstret med kongruenta rektanglar som avses vid Buffon-Laplace problem, genom att beräkna sannolikheten att en nål med längd $\ell \leq \min(a, b)$ så att den korsar minst en linje kan π skattas genom upprepade försök.

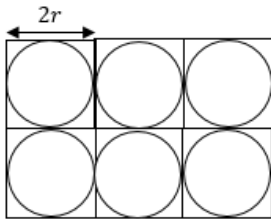
Det sägs att Buffon publicerade en lösning på problemet i sin bok i slutet på 1700-talet, men lösningen innehöll några fel. I början på 1800-talet löste Laplace problemet korrekt [3].

Hur ser kopplingen ut mellan Buffons problem och Buffon-Laplaces problem, går det att se ett kast i det senare problemet som två kast i det första problemet?

2.5 Kasta pil-problemet

Ett liknande problem som problemen Buffon formulerade är en darttavla med en cirkel med radien r är inskriven i en kvadrat med sidan $2r$. ”Vad är sannolikheten att en pil som kastas mot tavlan träffar innanför cirkeln?”

Genom att sätta ihop många tavlor blidas ett mönster (se figur 3) likt mönstren i de två tidigare problemen med inskrivna cirklar och genom att besvara frågan ovan kan π skattas [16].



Figur 3 Mönstret med inskrivna cirklar i kvadrater med sidan $2r$ som avses i Kasta pil-problemet. Genom att beräkna sannolikheten att en pil träffar inuti en cirkel kan π skattas genom upprepade försök.

2.6 Problemen i högre dimensioner

Vad händer om dimensionen går från två till högre dimensioner? Går det att skatta π utifrån modeller i högre dimensioner?

2.6.1 Buffons problem i 3 dimensioner

Vad händer med Buffons problem i en dimension högre? Med två parallella plan med avstånd b , och ponera att istället för en nål med längd ℓ så kastas en cirkelskiva med diameter ℓ . Vad händer då, går π att skattas, och i så fall vad är sannolikheten att cirkelskivan korsar ett plan?

2.6.2 Kasta pil-problemet i 3 dimensioner

Vad händer om kasta-pilproblemet skalas upp till ett tre dimensionellt problem. Den inskrivna cirkeln ersätts av ett klot med radien r som är inskrivet i en kub med sidan $2r$. Analog med problemformuleringen bör problemet i 3 dimensioner formuleras som så att ”Vad är sannolikheten att en dator slumpmässigt genererar fram en koordinat som befinner sig innanför klotets yta?”

Eftersom klotets volym fås som

$$V = \frac{4\pi r^3}{3},$$

kommer sannolikheten för träff kunna beräknas på liknande sätt som i det 2 dimensionella problemet.

2.6.3 Kasta pil-problemet i k dimensioner

Volymen för en boll med radien r i k dimensioner som är centrerad i origo fås som [10].

$$V(0, r) = \frac{\pi^{k/2} r^k}{(k/2)!}.$$

Faktultetsfunktionen är definierad för heltal, och det inses lätt att för varje udda k kommer nämnaren inte bli ett heltal. Genom att använda Gammafunktionen $\Gamma(x)$ kan det problemet hanteras. För alla positiva reella tal x definieras $x! = \Gamma(x + 1)$ och särskilt att $\Gamma(1/2) = \sqrt{\pi}$. Det här betyder att för alla jämna dimensioner kommer volymen för en boll ges som ovan, medan för udda dimensioner kan volymformeln skrivas om till

$$V_{\text{udda}}(0, r) = \frac{2^{(k+1)/2} \pi^{(k-1)/2} r^k}{k (k-2)!!},$$

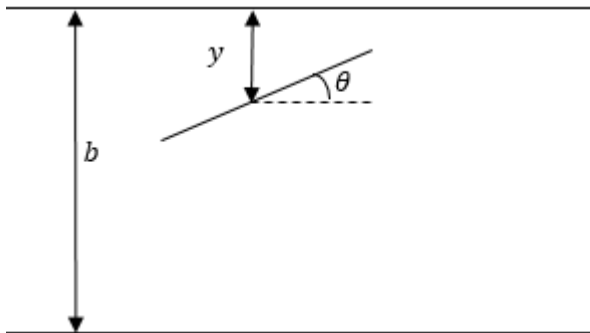
där $!!$ är semifakultet. Motsvarande k -kub i det k -dimensionella rummet kommer ha volymen $V_{\text{kub}} = (2r)^k$.

3 Resultat

I detta kapitel kommer de olika problemen som presenterades i Avsnitt 2.3-2.6 att lösas och sedan avslutas kapitlet med att modellerna i respektive problem jämförs.

3.1 Buffons problem

I Buffons problem fås två stokastiska variabler; Y som anger avståndet y från nålens centrum till närmsta linjen, och y antas vara likformigt fördelad på intervallet $\left[0, \frac{b}{2}\right]$ och Θ som betecknar vinkeln θ mellan nålen och den parallella linjen som löper genom nålens mittpunkt. Även θ är likformigt fördelad, men på intervallet $\left[0, \frac{\pi}{2}\right]$.



Figur 4 Buffons problem. Två parallella linjer med avstånd b och en nål som landat med avstånd y från linjen samt med vinkeln θ mot den streckade linjen som går genom nålens centrum och är parallell mot övriga linjer.

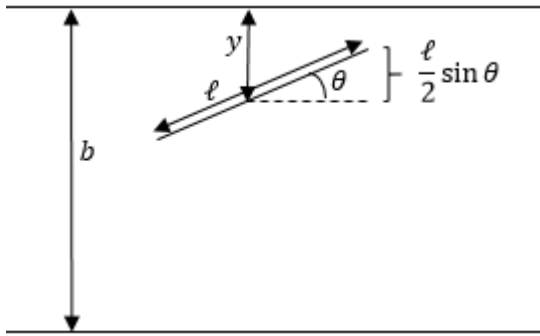
De två stokastiska variablerna Y och Θ kommer ha täthetsfunktionerna som anges nedan, eftersom de är likformigt fördelade på respektive intervall.

$$f_Y(y) = \begin{cases} \frac{2}{b} & \text{då } 0 \leq y \leq \frac{b}{2} \\ 0 & \text{annars} \end{cases}$$
$$f_\Theta(\theta) = \begin{cases} \frac{2}{\pi} & \text{då } 0 \leq \theta \leq \frac{\pi}{2} \\ 0 & \text{annars} \end{cases}$$

När nålen landar kommer nålens position, det vill säga avståndet till närmsta linje och med vilken vinkel nålen har (se figur 4) vara oberoende av varandra. Eftersom Y och Θ är oberoende av varandra kommer deras simultana täthetsfunktion ges som

$$f_{Y,\Theta}(y, \theta) = \begin{cases} \frac{4}{b\pi} & \text{då } 0 \leq y \leq \frac{b}{2}, 0 \leq \theta \leq \frac{\pi}{2} \\ 0 & \text{annars} \end{cases}$$

För att svara på frågan i Buffons problem behövs först frågan ”när korsar nålen en linje?” besvaras. Nålen kommer korsa en linje om avståndet $\frac{\ell}{2} \sin \theta$ är större än y (se figur 5).



Figur 5 Avståndet y som anger avståndet från nålens centrum och närmsta linje samt avståndet $\frac{l}{2} \sin \theta$. Om det senare avståndet är längre än y kommer nålen korsa linjen.

Enligt definition 2.1.2.9 är sannolikheten att nålen träffar

$$\begin{aligned}
 P(\text{träff}) &= \int_{\theta=0}^{\frac{\pi}{2}} \int_{y=0}^{\frac{\pi}{2} \frac{l}{2} \sin \theta} f_{Y,\theta}(y, \theta) dy d\theta \\
 &= \int_{\theta=0}^{\frac{\pi}{2}} \int_{y=0}^{\frac{\pi}{2} \frac{l}{2} \sin \theta} \frac{4}{b\pi} dy d\theta \\
 &= \int_0^{\frac{\pi}{2}} \frac{2l}{b\pi} \sin \theta d\theta \\
 &= \frac{2l}{b\pi} [-\cos \theta]_0^{\frac{\pi}{2}} \\
 &= \frac{2l}{b\pi}.
 \end{aligned}$$

Om försöket med att kasta nålar upprepas tillräckligt många gånger kommer den relativa frekvensen att nålen korsar en linje att gå mot den sanna sannolikheten, enligt Bernoullis sats (korollarium 2.1.2.1). Låt därför n vara totala antalet kast, då är

$$P_R(\text{träff}) = \frac{z_1 + z_2 + \dots + z_n}{n}$$

där $z_i = \begin{cases} 1 & \text{om träff} \\ 0 & \text{annars.} \end{cases}$

Eftersom $P_R(\text{träff})$ går mot den sanna sannolikheten $P(\text{träff})$, kan dessa sannolikheter sätts lika med varandra och genom att lösa ut π , fås en estimator till π som

$$\hat{\pi} = \frac{2ln}{b(z_1 + \dots + z_n)}.$$

Låt Z vara den stokastiska variabeln som anger träff eller inte. Då kommer Z vara binomialfördelad med parametrar

$$Z \sim \text{Bin}\left(n, \frac{2l}{b\pi}\right).$$

Eftersom den stokastiska variabeln Z är i nämnaren så beräknas variansen för $\hat{\pi}$ med hjälp av Deltametoden (sats 2.1.3.5) och sats 2.1.2.1. och då fås variansen för $\hat{\pi}$ som

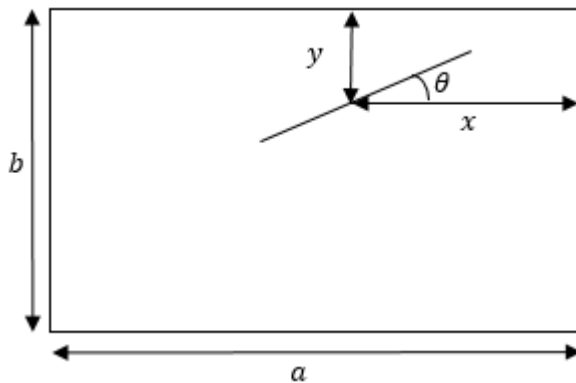
$$\begin{aligned} V(\hat{\pi}) &= \frac{4n^2l^2}{b^2} V\left(\frac{1}{Z}\right) \\ &\approx \frac{4n^2l^2}{b^2} \left(\frac{1}{n \frac{2l}{b\pi}}\right)^4 n \frac{2l}{b\pi} \left(1 - \frac{2l}{b\pi}\right) \\ &= \frac{4n^2l^2}{b^2} \frac{b^3\pi^3}{n^3 8\ell^3} \left(1 - \frac{2l}{b\pi}\right) \\ &= \frac{\pi^2 b\pi}{n 2\ell} \left(1 - \frac{2\ell}{b\pi}\right) \\ &= \frac{\pi^2}{n} \left(\frac{b\pi}{2l} - 1\right) \end{aligned}$$

Det inses att $V(\hat{\pi})$ inte kan bli noll, då $\ell \leq b$ enligt antagande, men genom att låta $n \rightarrow \infty$ minimeras variansen,

$$\lim_{n \rightarrow \infty} \frac{\pi^2}{n} \left(\frac{b\pi}{2l} - 1\right) = 0.$$

3.2 Buffon–Laplaces problem

I Buffon-Laplaces problem tillkommer ytterligare än stokastisk variabel, X , förutom Y och Θ som redan fanns i Buffons problem. Anledningen är att nu har nålen ännu en linje att förhålla sig till när nålens position bestäms. Skillnaden mot Buffons problem är att nu är mönstret nålen kastas mot ett rutnät, parallella linjer både vertikalt och horisontellt. Variabeln Y anger som tidigare avståndet y från nålens mittpunkt till närmsta *horisontella* linje och X anger avståndet x från mittpunkten till närmsta *vertikala* linje (se figur 6). Avståndet mellan de horisontella linjerna är b så y är likformigt fördelad på intervallet $\left[0, \frac{b}{2}\right]$ och avståndet mellan de vertikala linjerna är a , så x är likformigt fördelad på intervallet $\left[0, \frac{a}{2}\right]$. Som tidigare anger Θ vinkeln θ mellan nålen och den parallella linjen som löper genom nålens mittpunkt och som innan är θ likformigt fördelad på intervallet $\left[0, \frac{\pi}{2}\right]$.



Figur 6 Buffons-Laplaces problem. En rektangel med sidan a och b och en nål som landat med avstånd y från den horisontella linjen samt avstånd x från den vertikala linjen. Nålen har landat med vinkeln θ mot den streckade linjen som går genom nålens centrum och är parallell mot de horisontella linjerna.

De stokastiska variablerna Y, X och Θ kommer ha täthetsfunktionerna som anges nedan, eftersom de är likformigt fördelade på respektive intervall

$$f_X(x) = \begin{cases} \frac{2}{a} & \text{då } 0 \leq x \leq \frac{a}{2} \\ 0 & \text{annars} \end{cases}$$

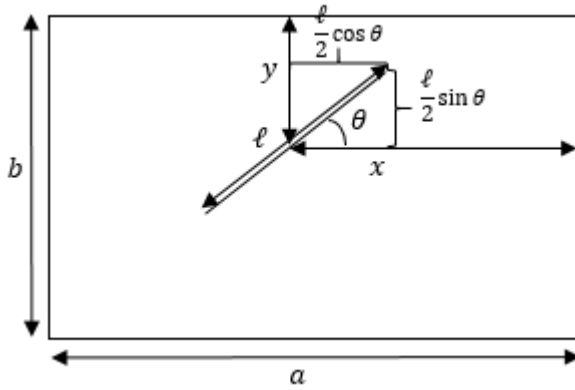
$$f_Y(y) = \begin{cases} \frac{2}{b} & \text{då } 0 \leq y \leq \frac{b}{2} \\ 0 & \text{annars} \end{cases}$$

$$f_\Theta(\theta) = \begin{cases} \frac{2}{\pi} & \text{då } 0 \leq \theta \leq \frac{\pi}{2} \\ 0 & \text{annars} \end{cases}$$

Med samma resonemang som i Buffons problem kommer Y, X och Θ att vara oberoende av varandra, varför också de har den simultana täthetsfunktionen

$$f_{X,Y,\Theta}(x, y, \theta) = \begin{cases} \frac{8}{\pi ab} & \text{då } 0 \leq x \leq \frac{a}{2}, 0 \leq y \leq \frac{b}{2}, 0 \leq \theta \leq \frac{\pi}{2} \\ 0 & \text{annars} \end{cases}$$

När nålen kastas finns det tre möjliga utfall; att nålen inte korsar någon linje eller att nålen korsar en eller två linjer. För att besvara frågan ”vad är sannolikheten att nålen korsar *minst* en av linjerna?” kan Komplementsatsen (sats 2.1.1.2.) tillämpas, där komplementet till att nålen korsar minst en linje är att nålen inte korsar någon linje alls. Nålen kommer inte korsa någon linje om avstånden till respektive linje från nålens mittpunkt är $x \geq \frac{\ell}{2} \cos \theta$ samt att $y \geq \frac{\ell}{2} \sin \theta$ (se figur 7).



Figur 7 Avståndet y som anger avståndet från nålens centrum och närmsta horisontella linje samt avståndet $\frac{\ell}{2} \sin \theta$ och avståndet x som anger avståndet från nålens centrum till närmsta vertikala linje och avståndet $\frac{\ell}{2} \cos \theta$. Om $\frac{\ell}{2} \sin \theta$ och/eller $\frac{\ell}{2} \cos \theta$ är längre än y respektive x kommer nålen korsa linjen.

Avstånden x, y kommer var uppåt begränsade av respektive intervall. Alltså fås sannolikheten att nålen inte korsar någon linje som, enligt definition 2.1.2.9,

$$\begin{aligned}
 P(\text{ingen träff}) &= \int_{\theta=0}^{\frac{\pi}{2}} \int_{y=\frac{\ell}{2} \sin \theta}^{\frac{b}{2}} \int_{x=\frac{\ell}{2} \cos \theta}^{\frac{a}{2}} f_{X,Y,\Theta}(x, y, \theta) dx dy d\theta \\
 &= \int_{\theta=0}^{\frac{\pi}{2}} \int_{y=\frac{\ell}{2} \sin \theta}^{\frac{b}{2}} \int_{x=\frac{\ell}{2} \cos \theta}^{\frac{a}{2}} \frac{8}{\pi ab} dx dy d\theta \\
 &= \frac{8}{ab\pi} \int_0^{\frac{\pi}{2}} \left(\frac{a}{2} - \frac{\ell}{2} \cos \theta \right) \left(\frac{b}{2} - \frac{\ell}{2} \sin \theta \right) d\theta \\
 &= \frac{2}{ab\pi} \int_0^{\frac{\pi}{2}} (ab - a\ell \sin \theta - b\ell \cos \theta + \ell^2 \cos \theta \cdot \sin \theta) d\theta \\
 &= \frac{2}{ab\pi} \left[ab\theta + a\ell \cos \theta - b\ell \sin \theta - \frac{\ell^2 \cos 2\theta}{4} \right]_0^{\frac{\pi}{2}} \\
 &= \frac{2}{ab\pi} \left(\frac{ab\pi}{2} - b\ell + \frac{\ell^2}{4} - \left(a\ell - \frac{\ell^2}{4} \right) \right) \\
 &= \frac{2}{ab\pi} \left(\frac{ab\pi}{2} - \ell(a+b) + \frac{\ell^2}{2} \right) \\
 &= 1 - \frac{2\ell(a+b) - \ell^2}{\pi ab}.
 \end{aligned}$$

Sannolikheten som skulle beräknas var att nålen korsade minst en linje, så enligt Komplementsatsen (sats 2.1.1.2) fås att

$$\begin{aligned} P(\text{träff}) &= 1 - P(\text{ingen träff}) \\ &= 1 - \left(1 - \frac{2\ell(a+b) - \ell^2}{\pi ab}\right) \\ &= \frac{2\ell(a+b) - \ell^2}{\pi ab}. \end{aligned}$$

Som tidigare, om försöket med att kasta nålen upprepas tillräckligt många gånger kommer den relativa frekvensen att nålen korsar en linje gå mot den sanna sannolikheten att nålen korsar en linje, så att en estimator till π fås som

$$\hat{\pi} = \frac{n(2\ell(a+b) - \ell^2)}{ab(z_1 + \dots + z_n)}$$

där n anger totalt antal kast och z_i definieras som tidigare (oberoende om nålen träffar en eller två linjer),

$$Z \sim \text{Bin}\left(n, \frac{2\ell(a+b) - \ell^2}{\pi ab}\right).$$

Variansen för $\hat{\pi}$ fås precis som tidigare, enligt Deltametoden (sats 2.1.3.5) och med sats 2.1.2.1, som

$$\begin{aligned} V(\hat{\pi}) &= \left(\frac{(2\ell(a+b) - \ell^2)n}{ab}\right)^2 V\left(\frac{1}{Z}\right) \\ &\approx \frac{(2\ell(a+b) - \ell^2)^2 n^2}{(ab)^2} \left(\frac{1}{np}\right)^4 np(1-p) \\ &= \frac{(2\ell(a+b) - \ell^2)^2 n^2}{(ab)^2} \frac{1}{n^3 p^3} (1-p) \\ &= \frac{1}{n(ab)^2 \frac{p}{(\pi ab)^2}} (1-p) \\ &= \frac{\pi^2}{n} \left(\frac{1}{p} - 1\right) \\ &= \frac{\pi^2}{n} \left(\frac{\pi ab}{2\ell(a+b) - \ell^2} - 1\right). \end{aligned}$$

Det inses att variansen inte kan bli noll, eftersom

$$\begin{aligned} \frac{\pi ab}{2\ell(a+b) - \ell^2} - 1 &= 0 \\ \pi ab &= 2\ell(a+b) - \ell^2 \\ \ell &= (a+b) \pm \sqrt{(a+b)^2 - \pi ab}. \end{aligned}$$

Om $a = b$ fås att

$$\begin{aligned}\ell &= 2b \pm \sqrt{(2b)^2 - \pi b^2} \\ \ell &= 2b \pm |b|\sqrt{4 - \pi},\end{aligned}$$

där båda lösningarna inte uppfyller kravet att $\ell \leq \min(a, b)$. Så för att minimera variansen måste antal försök ökas.

Anmärkning: Om sidan a väljs oändligt bred så blir sannolikheten för träff

$$\begin{aligned}\lim_{a \rightarrow \infty} P(\text{träff}) &= \lim_{a \rightarrow \infty} \frac{2\ell(a+b) - \ell^2}{\pi ab} \\ &= \lim_{a \rightarrow \infty} \left(\frac{2\ell(a+b)}{\pi ab} - \frac{\ell^2}{\pi ab} \right) \\ &= \lim_{a \rightarrow \infty} \left(\frac{2\ell \left(1 + \frac{b}{a}\right)}{\pi b} - \frac{\ell^2}{\pi ab} \right) \\ &= \frac{2\ell}{\pi b},\end{aligned}$$

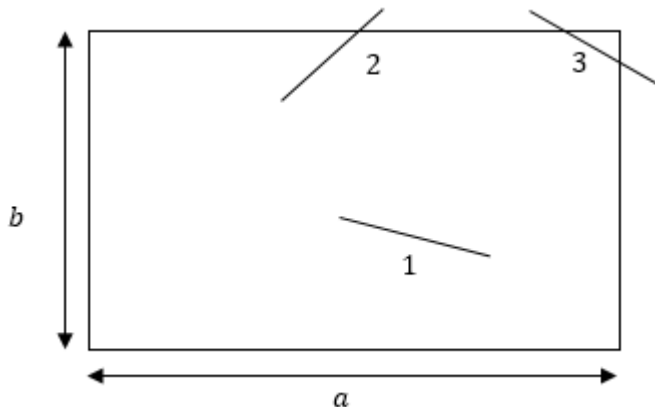
vilket är samma som i Buffons problem, det vill säga Buffons problem är ett specialfall av Buffon-Laplace problem, där sidan a väljs oändligt bred.

3.3 Buffon-Laplace-kast som Buffon-kast

Går det att se ett kast med nålen på rutnätsmönstret (Buffon-Laplace problem) som två stycken kast på mönstret med parallella linjer (Buffons problem)?

Ett kast vid Buffon-Laplace problem har tre möjliga utfall (se figur 8);

1. Nålen träffar *ingen* linje
2. Nålen träffar *en* linje
3. Nålen träffar *två* linjer



Figur 8 De möjliga utfall som finns i Buffon-Laplace problem: nålen korsar noll, en eller två linjer.

Sannolikheten att nålen inte träffar någon linje beräknades i avsnitt 3.2 Buffon-Laplaces problem, och där erhöles sannolikheten

$$P(\text{ingen träff}) = 1 - \frac{2\ell(a+b) - \ell^2}{\pi ab}.$$

Kvar att beräkna är sannolikheten för att nålen korsar en linje samt två linjer. Det räcker att beräkna en av dessa, eftersom de tre utfallen utgör hela utfallsrummet och tillsammans summerar dessa sannolikheter till 1 enligt Kolmogorovs axiom 2 och 3. Eftersom nålen kan korsa en linje på flera sätt (antingen korsar den en vertikal eller en horisontell linje) är det lättare att beräkna sannolikheten att nålen korsar två linjer. Om nålen landar så att $x \leq \frac{\ell}{2} \cos \theta$ och $y \leq \frac{\ell}{2} \sin \theta$, kommer nålen att korsa två linjer. Alltså fås sannolikheten som, enligt definition 2.1.2.9,

$$\begin{aligned} P(\text{två träffar}) &= \int_{\theta=0}^{\frac{\pi}{2}} \int_{y=0}^{\frac{\ell}{2} \sin \theta} \int_{x=0}^{\frac{\ell}{2} \cos \theta} f_{X,Y,\Theta}(x, y, \theta) dx dy d\theta \\ &= \int_{\theta=0}^{\frac{\pi}{2}} \int_{y=0}^{\frac{\ell}{2} \sin \theta} \int_{x=0}^{\frac{\ell}{2} \cos \theta} \frac{8}{\pi ab} dx dy d\theta \\ &= \frac{8}{\pi ab} \int_0^{\frac{\pi}{2}} \frac{\ell}{2} \cos \theta \cdot \frac{\ell}{2} \sin \theta d\theta \\ &= \frac{2}{\pi ab} \left(\int_0^{\frac{\pi}{2}} \frac{\ell^2 \sin 2\theta}{2} d\theta \right) \\ &= \frac{\ell^2}{\pi ab} \left(\left[-\frac{\cos 2\theta}{2} \right]_0^{\frac{\pi}{2}} \right) \\ &= \frac{\ell^2}{\pi ab} \left(\frac{1}{2} + \frac{1}{2} \right) \\ &= \frac{\ell^2}{\pi ab}. \end{aligned}$$

Sannolikheten att nålen träffar en linje fås som

$$\begin{aligned} P(\text{en träff}) &= 1 - P(\text{ingen träff}) - P(\text{två träffar}) \\ &= 1 - \left(1 - \frac{2\ell(a+b) - \ell^2}{\pi ab} \right) - \frac{\ell^2}{\pi ab} \end{aligned}$$

$$\begin{aligned}
&= \frac{2\ell(a+b) - \ell^2}{\pi ab} - \frac{\ell^2}{\pi ab} \\
&= \frac{2\ell(a+b) - 2\ell^2}{\pi ab}.
\end{aligned}$$

Låt ζ_i vara den stokastiska variabeln som anger antalet linjer som nålen korsar vid kast i . Sannolikhetsfunktionerna för ζ för $k = 0, 1, 2$, fås som

$$p_\zeta(0) = P(\zeta = 0) = 1 - \frac{2\ell(a+b) - \ell^2}{\pi ab},$$

$$p_\zeta(1) = P(\zeta = 1) = \frac{2\ell(a+b) - 2\ell^2}{\pi ab},$$

$$p_\zeta(2) = P(\zeta = 2) = \frac{\ell^2}{\pi ab}.$$

Förväntat antal träffar beräknas enligt definition 2.1.2.5. som

$$\begin{aligned}
\mu_\zeta = E(\zeta) &= \sum_{k=0}^2 k \cdot p_\zeta(k) \\
&= 0 \cdot p_\zeta(0) + 1 \cdot p_\zeta(1) + 2 \cdot p_\zeta(2) \\
&= \frac{2\ell(a+b) - 2\ell^2 + 2\ell^2}{\pi ab} \\
&= \frac{2\ell(a+b)}{\pi ab},
\end{aligned}$$

samt med variansen

$$\begin{aligned}
\sigma_\zeta^2 = V(\zeta) &= E(\zeta^2) - (E(\zeta))^2 \\
&= \frac{2\ell(a+b) + 2\ell^2}{\pi ab} - \left(\frac{2\ell(a+b)}{\pi ab}\right)^2.
\end{aligned}$$

Enligt Centrala gränsvärdesatsen (sats 2.1.3.1.) kommer den stokastiska variabeln ζ , som då anger totalt antal träffade linjer, approximativt vara normalfördelad med parametrar $n\mu_\zeta$ och $\sqrt{n}\sigma_\zeta$, alltså

$$\zeta \stackrel{\text{appr}}{\sim} N(n\mu_\zeta, \sqrt{n}\sigma_\zeta).$$

Tidigare beräknades sannolikheten för träff i Buffon-Laplaces problem till

$$P(\text{träff}) = \frac{2\ell(a+b) - \ell^2}{\pi ab}.$$

Då fås en estimator på π som

$$\hat{\pi} = \frac{n(2\ell(a+b) - \ell^2)}{ab(\zeta_1 + \dots + \zeta_n)},$$

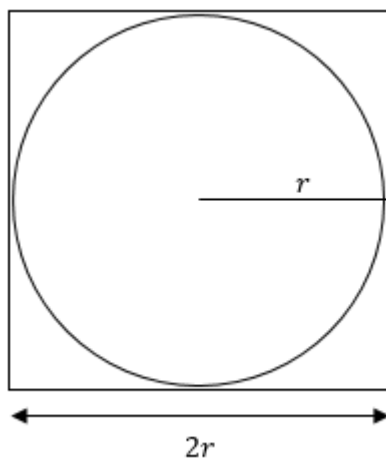
där ζ_i är antalet träffade linjer i kast i .

Variansen på $\hat{\pi}$ blir följande

$$\begin{aligned} V(\hat{\pi}) &= \left(\frac{(2\ell(a+b) - \ell^2)n}{ab} \right)^2 V\left(\frac{1}{\zeta}\right) \\ &\approx \left(\frac{(2\ell(a+b) - \ell^2)n}{ab} \right)^2 \frac{1}{(n\mu_\zeta)^4} n\sigma_\zeta^2 \\ &\approx \left(\frac{(2\ell(a+b) - \ell^2)}{ab} \right)^2 \frac{1}{n} \frac{1}{(\mu_\zeta)^4} \sigma_\zeta^2 \end{aligned}$$

3.4 Kasta pil-problemet

I kasta pil-problemet är förfarandet lite annorlunda från tidigare problem, men väldigt snarligt. Skillnaden ligger i att pilen som träffar tavlan inte beror av någon vinkel som nålen i tidigare problem har gjort. Pilens position kommer att bestämmas av två stokastiska variabler X och Y som anger avståndet x till närmsta vertikala sidan på kvadraten och y som anger avståndet till närmsta horisontella sida. Både x och y är likformigt fördelade på intervallet $[0, r]$. Cirkel har radien r , så att kvadratens sida är $2r$ (se figur 9). Ett försök räknas som "träff" om pilens spets träffar innanför cirkelns rand.



Figur 9 Kasta pil-problemet. En cirkel med radien r är inskriven i en kvadrat med sidan $2r$. Ett kast räknas som träff om pilen träffar innanför cirkelns rand.

Problemet kan lösas enligt den klassiska sannolikhetsdefinitionen (se [11]) som är att genom att beräkna den gynnsamma arean, cirkelns area, och dividera den arean

med den totala arean fås sannolikheten att det gynnsamma ska inträffa (d.v.s. pilen träffar i detta fall),

$$P(\text{träff}) = \frac{A_{\text{cirkel}}}{A_{\text{total}}} = \frac{\pi r^2}{(2r)^2} = \frac{\pi}{4}.$$

Förutom att använda den klassiska sannolikhetsdefinitionen kan sannolikheten att pilen träffar innanför cirkeln beräknas på samma sätt som i de tidigare problemen – med de två stokastiska variablernas simultana täthetsfunktion, se nedan.

Anta att cirkeln har centrum i origo. Då fås cirkelns ekvation som $x^2 + y^2 = r^2$. De två variablerna (x, y) anger pilens position i första kvadranten (p.g.a. symmetri räcker det med en kvadrant). Variablerna x, y fördelas likformigt samt oberoende av varandra på intervallen $x, y \in [0, r]$.

Som tidigare kommer de stokastiska variablerna Y, X kommer att ha täthetsfunktionerna som anges nedan, eftersom de är likformigt fördelade,

$$f_X(x) = \begin{cases} \frac{1}{r} & \text{då } 0 \leq x \leq r \\ 0 & \text{annars} \end{cases}$$
$$f_Y(y) = \begin{cases} \frac{1}{r} & \text{då } 0 \leq y \leq r \\ 0 & \text{annars} \end{cases}$$

så att den simultana täthetsfunktionen fås som produkten $f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y)$, på grund av att x och y är oberoende,

$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{r^2} & \text{då } 0 \leq x \leq r \text{ } 0 \leq y \leq r \\ 0 & \text{annars} \end{cases}$$

Kastet räknas som träff om $x^2 + y^2 \leq r^2$, för att pilen ska träffa innanför cirkeln. Genom ett byte till polära koordinater fås att

$$x = R \cos \theta$$
$$y = R \sin \theta$$

där $0 \leq R \leq r, 0 \leq \theta \leq \frac{\pi}{2}$. Så att sannolikheten att pilen träffar innanför cirkeln beräknas som

$$P(\text{träff}) = \int_{y=0}^r \int_{x=0}^{\sqrt{r^2-y^2}} f_{X,Y}(x, y) dx dy$$
$$= \int_{\theta=0}^{\pi/2} \int_{R=0}^r \frac{1}{r^2} R dR d\theta$$

$$\begin{aligned}
&= \int_{\theta=0}^{\pi/2} \left(\frac{1}{r^2} \left[\frac{R^2}{2} \right]_0^r \right) d\theta \\
&= \int_{\theta=0}^{\pi/2} \frac{1}{2} d\theta \\
&= \frac{\pi}{4}.
\end{aligned}$$

Med samma resonemang som innan kommer den relativa frekvensen för träff att närma sig den verkliga sannolikheten om försöket upprepas många gånger. En estimator till π fås som,

$$\hat{\pi} = \frac{4(z_1 + \dots + z_n)}{n},$$

där n anger totalt antal kast och z_i definieras som tidigare.

$$Z \sim \text{Bin}\left(n, \frac{\pi}{4}\right).$$

Det är en liten, men viktig, detalj som skiljer estimatorn för $\hat{\pi}$ i kasta pil-problemet mot estimatorerna i Buffons olika problem, den stokastiska variabeln Z återfinns nu i täljaren och inte i nämnaren som tidigare, vilket förenklar variansberäkningarna avsevärt. Variansen för $\hat{\pi}$ fås som

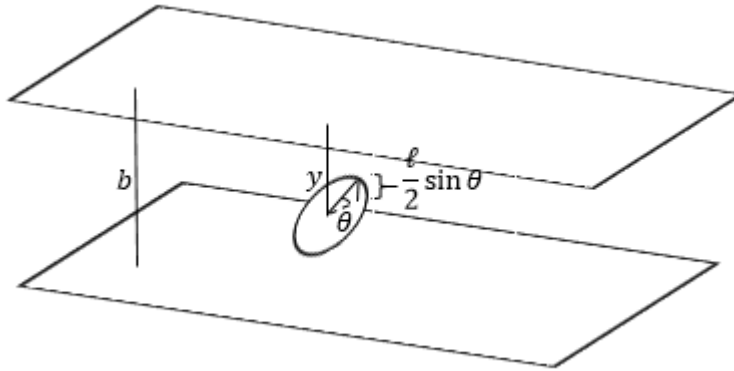
$$\begin{aligned}
V(\hat{\pi}) &= \frac{16}{n^2} V(Z) \\
&= \frac{16}{n^2} n \frac{\pi}{4} \left(1 - \frac{\pi}{4}\right) \\
&= \frac{4\pi}{n} \left(1 - \frac{\pi}{4}\right)
\end{aligned}$$

det inses att variansen aldrig kan bli noll, men genom att låta $n \rightarrow \infty$ minimeras variansen,

$$\lim_{n \rightarrow \infty} \frac{4\pi}{n} \left(1 - \frac{\pi}{4}\right) = 0.$$

3.5 Buffons problem i 3D

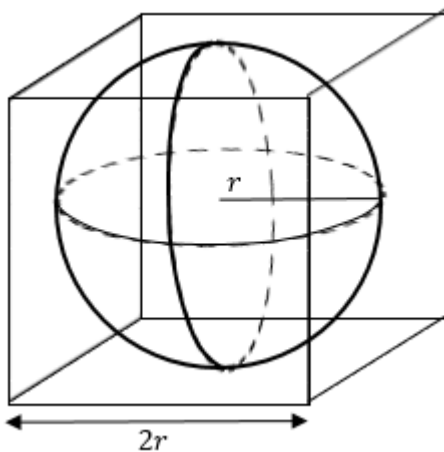
I Buffons problem i 3D kommer det finnas två stokastiska variabler som är av intresse, Y som anger avståndet y från nålen centrum till närmsta plan och Θ som anger vinkeln θ , där båda variablerna kommer vara likformigt fördelade på intervall $\left[0, \frac{b}{2}\right]$ samt $\left[0, \frac{\pi}{2}\right]$. Anta att avståndet mellan planen är b och cirkelskivan har diameter ℓ . Cirkelskivan kommer korsa ett plan om $\frac{\ell}{2} \sin \theta \geq y$ (se figur 10), och då inses det att det är samma förfarande som i det två dimensionella fallet, se avsnitt 3.1 för motsvarande lösning av ursprungsproblemet.



Figur 10 Buffons problem i 2D. Två parallella plan med avstånd b . En cirkelskiva med diameter "kastas" och kommer då landa med ett visst avstånd y från närmsta plan och med en vinkel θ .

3.6 Kasta pil-problemet i 3 dimensioner

Anta att klotet har radien r , så att kuben har sidan $2r$, där klotets centrum antas vara origo (se figur 11). En koordinat väljs slumpmässigt ut av ett dataprogram och om punkten befinner sig innanför klotets rand räknas det som en träff. Punkten beskrivs av tre koordinater (x, y, z) , som var för sig väljs likformigt oberoende av varandra.



Figur 11 Ett inskrivet klot i en kub, klot tangerar kubens insidor. Klotets radie är r , så att kubens sida är $2r$. Om en punkt är innanför klotets rand räknas den som träff.

Sannolikheten för träff fås som, enligt den klassiska sannolikhetsdefinitionen (se [8]),

$$P(\text{träff}) = \frac{V_{\text{Klot}}}{V_{\text{Kub}}} = \frac{4\pi r^3}{3(2r)^3} = \frac{\pi}{6}.$$

Samma svar fås om täthetsfunktioner används. Tack vare symmetrin kan problemet begränsas till första oktanten¹. Sfärens yta kan beskrivas med ekvationen $x^2 + y^2 + z^2 = r^2$. Punktens koordinater kommer vara likformigt

¹ Inom matematiken är en **oktant** en av åtta indelningar av ett vanligtvis tredimensionellt rum.

fördelad på intervallet $[0, r]$ respektive. Så att täthetsfunktionerna för respektive variabel fås som

$$f_X(x) = \begin{cases} \frac{1}{r} & \text{då } 0 \leq x \leq r \\ 0 & \text{annars} \end{cases}$$

$$f_Y(y) = \begin{cases} \frac{1}{r} & \text{då } 0 \leq y \leq r \\ 0 & \text{annars} \end{cases}$$

$$f_Z(z) = \begin{cases} \frac{1}{r} & \text{då } 0 \leq z \leq r \\ 0 & \text{annars} \end{cases}$$

och eftersom variablerna är oberoende av varandra fås den simultana täthetsfunktionen som

$$f_{X,Y,Z}(x, y, z) = \begin{cases} \frac{1}{r^3} & \text{då } 0 \leq x \leq r, 0 \leq y \leq r, 0 \leq z \leq r \\ 0 & \text{annars.} \end{cases}$$

Genom byte till sfäriska koordinater fås att

$$x = R \sin \varphi \cos \theta$$

$$y = R \sin \varphi \sin \theta$$

$$z = R \cos \varphi,$$

där $0 \leq R \leq r, 0 \leq \theta \leq \frac{\pi}{2}, 0 \leq \varphi \leq \frac{\pi}{2}$. Alltså ska följande integral beräknas, för att beräkna sannolikheten för träff,

$$\begin{aligned} P(\text{träff}) &= \iiint_{R^2=x^2+y^2+z^2} \frac{1}{r^3} dx dy dz \\ &= \int_{\varphi=0}^{\frac{\pi}{2}} \int_{\theta=0}^{\frac{\pi}{2}} \int_{R=0}^r \frac{1}{r^3} R^2 \sin \theta dR d\theta d\varphi \\ &= \frac{\pi}{6} \int_{\theta=0}^{\frac{\pi}{2}} \sin \theta d\theta \\ &= \frac{\pi}{6} [-\cos \theta]_0^{\frac{\pi}{2}} \\ &= \frac{\pi}{6} (0 - (-1)) \\ &= \frac{\pi}{6}. \end{aligned}$$

Genom att upprepa försöket många gånger kan, enligt Bernoullis sats (korollarium 2.1.2.1.) en estimator till π fås som

$$\hat{\pi} = \frac{6(z_1 + \dots + z_n)}{n},$$

där n anger totalt antal kast och z_i definieras som tidigare.

$$Z \sim \text{Bin}\left(n, \frac{\pi}{6}\right).$$

Precis som i kasta pil-problemet beräknas variansen för $\hat{\pi}$ utan problem,

$$\begin{aligned} V(\hat{\pi}) &= \frac{36}{n^2} V(Z) \\ &= \frac{36}{n^2} n \frac{\pi}{6} \left(1 - \frac{\pi}{6}\right) \\ &= \frac{6\pi}{n} \left(1 - \frac{\pi}{6}\right). \end{aligned}$$

3.7 Kasta pil-problemet i k dimensioner

Anta att en boll i \mathbb{R}^k , där k är ett jämnt tal, har radie r och centrum i origo. Volymen fås som, (se avsnitt 2.6.3),

$$V(0, r) = \frac{\pi^{k/2} r^k}{(k/2)!}.$$

Bollen är inskriven i en kub, med sidan $2r$, som har volymen $V_{k\text{-kub}} = (2r)^k$. En koordinat i kuben väljs slumpmässigt ut, där alla punkter har samma sannolikhet att väljas (likformigt). Då kan sannolikheten för att punkten ligger inuti bollen beräknas som

$$P(\text{träff}) = \frac{V(0, r)}{V_{k\text{-kub}}} = \frac{\pi^{k/2} r^k}{(2r)^k (k/2)!} = \frac{\pi^{k/2}}{2^k (k/2)!}.$$

I enlighet med Bernoullis sats (korollarium 2.1.2.1.) kan sannolikheten för träff beräknas som

$$P_R(\text{träff}) = \frac{z_1 + z_2 + \dots + z_n}{n}$$

om n är tillräckligt stor, och z_i definieras som tidigare, så att π kan skattas som

$$\hat{\pi}_{\text{jämn}} = \left(\frac{2^k \left(\frac{k}{2}\right)! (z_1 + z_2 + \dots + z_n)}{n} \right)^{\frac{2}{k}}.$$

Om dimensionen istället är udda fås att för en boll i \mathbb{R}^k , där k är ett *udda* tal, som har radie r och centrum i origo. Blir sannolikheten för träff, efter omskrivning av volymen för bollen (se avsnitt 2.6.3),

$$P(\text{träff}) = \frac{V_{\text{udda}}(0, r)}{V_{k\text{-kub}}} = \frac{2^{(k+1)/2} \pi^{(k-1)/2} r^k}{(2r)^k k (n-2)!!} = \frac{\pi^{(k-1)/2}}{2^{(k-1)/2} k (k-2)!!}$$

Så att en skattning av π , enligt Bernollis sats då sannolikheten vid tillräckligt många försök kommer närma sig den teoretiska sannolikheten för träff, istället fås som

$$\hat{\pi}_{\text{udda}} = \left(\frac{2^{(k-1)/2} k (k-2)!! (z_1 + z_2 + \dots + z_k)}{n} \right)^{\frac{2}{k-1}}.$$

Vad har respektive skattning för varians? Enligt deltametoden (sats 2.1.3.5.) kommer

$$V(\hat{\pi}_{\text{jämn}}) \approx \left(\frac{2^k \left(\frac{k}{2}\right)!}{n} \right)^{\frac{4}{k}} \cdot \left(\frac{4}{k^2} (np)^{\frac{4}{k}-1} \right)$$

eftersom $Z \sim \text{Bin}(n, p_{\text{jämn}})$ där $p_{\text{jämn}}$ är sannolikheten för träff, och $g(Z) = Z^{2/k}$.

$$V(\hat{\pi}_{\text{udda}}) \approx \left(\frac{2^{(k-1)/2} k (k-2)!!}{n} \right)^{\frac{4}{k-1}} \cdot \left(\frac{4}{(k-1)^2} (np)^{\frac{4}{k-1}-1} \right),$$

Som ovan, $Z \sim \text{Bin}(n, p_{\text{udda}})$ där p_{udda} är sannolikheten för träff, och $g(Z) = Z^{2/(k-1)}$.

3.8 Jämförelse mellan modellerna

Hur förhåller sig de olika modellerna till varandra? En ”bra” estimator är väntevärdesriktig och har liten spridning på skattningarna, alltså liten varians.

Nedan är en tabell (tabell 1) med olika värden på sidorna a , b och nålens längd ℓ i en jämförelse mellan Buffon, Buffon-Laplace problem och fallet när Buffon-Laplace-kast ses som Buffon-kast (namngett som ”Dubbla Buffon” i tabell 1).

Tabell 1 Variansberäkning för de olika modellerna med olika länder på a , b samt ℓ .

Modell:	Buffon	Buffon-Laplace	”Dubbla Buffon”
$a = b = 10, \ell = 10$	$\frac{5,64}{n}$	$\frac{0,47}{n}$	$\frac{0,99}{n}$
$a = b = 15, \ell = 10$	$\frac{5,63}{n}$	$\frac{4,08}{n}$	$\frac{3,91}{n}$
$a = b = 20, \ell = 10$	$\frac{21,14}{n}$	$\frac{7,84}{n}$	$\frac{7,28}{n}$
$a = b = 30, \ell = 10$	$\frac{36,64}{n}$	$\frac{15,50}{n}$	$\frac{14,50}{n}$

Från tabellen (tabell 1) fås att Buffons originalproblem har en avvikande hög varians jämfört med Buffon-Laplace problem, om inte sidan a , b är nära nålen längd ℓ , så att Buffon-Laplace problem är att föredra som modell för att skatta π eftersom dess estimator är mer effektiv än estimatorn i Buffons problem.

Att utnyttja Buffon-Laplaces-kast för att skatta π med Buffons problem minimerar variansen litegrann i alla fall utom då nålen är lika lång som avståndet mellan linjerna, så ja, det är mer effektivt att kasta nålar på rutnätsmönstret och notera om nålen korsar ingen, en eller två linjer i varje kast för att skatta π jämför med Buffons originalidé.

I tabell 2 nedan ges variansen för kasta pil-problemen i dimension 2, 3, 5, 10 och 20. Till skillnad från Buffons olika problem beror inte sannolikheten för träff på vad radien är vilket gör att variansen kommer vara oberoende av r .

Tabell 2 Variansberäkning för kasta pil-problemet i olika dimensioner.

Modell kasta pil (dimension):	Varians:
Kasta pil	$\frac{2,70}{n}$
Kasta pil 3D	$\frac{8,98}{n}$
$k = 5$	$\frac{15}{n}$
$k = 10$	$\frac{158,5}{n}$

$k = 20$	$\frac{1,8 \cdot 10^{29}}{n}$
----------	-------------------------------

Den stokastiska variabeln Z som anger om pilen träffar innanför cirkelns rand eller inte, är som tidigare nämnt, binomialfördelad och variansen för Z är $V(Z) = np(1 - p)$. Läsaren bör notera att variansen är som störst då sannolikheten för träff/inte träff är lika stora (då $p = 1/2$), men kommer minska när skillnaden mellan händelserna ökar, vilket är anledningen till att kasta pil-problemet i 3D har högre varians än vad ursprungsproblemet har. För i båda de fallen beror variansen för $\hat{\pi}$ på just Z (och en konstant som naturligtvis kommer vara med och påverka storleken på variansen för $\hat{\pi}$, men i jämförelse med varandra kommer det vara sannolikheten för träff som avgör vilken varians som blir minst), och inte en funktion $g(Z)$ som i alla andra problem. Sannolikheten för träff i kasta pil-problemet är $P_{2D}(\text{träff}) \approx 0.79$ medan för 3D problemet $P_{3D}(\text{träff}) \approx 0.52$, vilket förklarar resultatet.

När dimensionen ökar i kasta pil-problemet ökar också variansen för skattningen på π . Det är rimligt, för när dimensionen ökas kommer k -kubens volym öka mer än vad k -bollens volym gör. I oändligheten kommer sannolikheten för träff att vara obefintlig, och eftersom hela idén bakom problemet är att det ska finnas två möjliga utfall, kommer det inte gå att skatta π (eftersom variansen är så stor). Så slutsatsen blir att det *går* att generalisera kasta pil-problemet till k dimensioner, men att variansen ökas så kraftigt redan efter ett par dimensioner så att skattningarna inte längre kommer vara tillförlitliga och därmed oanvändbara.

Slutligen, i detta arbete har fokus varit på π med utgångspunkt i Buffons ursprungsidé att hitta ett slumpförsök som kan användas för att skatta just π . De olika problemen som har behandlats här har samma grundtanke och ger liknande, sånär på hur precisa skattningar som fås, resultat. Att det finns andra modeller för att skatta π är redan känt, men en naturlig fortsättning som fördjupning på arbetet är undersöka möjligheten att skatta andra matematiska konstanter på liknande sätt.

Referenslista

- [1] "Buffon's Needle Problem", *Mathworld*, [online]. Tillgänglig: <http://mathworld.wolfram.com/BufonsNeedleProblem.html> [Hämtad: 2018-05-24]
- [2] "Clean Tile Problem", *Mathworld*, [online]. Tillgänglig: <http://mathworld.wolfram.com/CleanTileProblem.html> [Hämtad: 2018-05-24]
- [3] "Buffon-Laplace Needle Problem", *Mathworld*, [online]. Tillgänglig: <http://mathworld.wolfram.com/Bufon-LaplaceNeedleProblem.html> [Hämtad: 2018-05-24]
- [4] S. E. Alm och T. Britton, *Stokastik: Sannolikhetsteori med statistikteori och tillämpningar*. Stockholm: Liber AB, 2008.
- [5] "Pi", *Mathworld*, [online]. Tillgänglig: <http://mathworld.wolfram.com/Pi.html> [Hämtad: 2018-05-24]
- [6] "Pi formulas" *Mathworld*, [online]. Tillgänglig: <http://mathworld.wolfram.com/PiFormulas.html> [Hämtad: 2018-05-24].
- [7] J. Arndt och C. Haenel, *Pi unleashed*. New York: Springer, 2000.
- [8] "Celebrate pi day with 9 trillion more digits more than ever before", *New scientist*, [online]. Tillgänglig: <https://www.newscientist.com/article/2124418-celebrate-pi-day-with-9-trillion-more-digits-than-ever-before/> [Hämtad: 2018-06-15].
- [9] "Chudnovsky algorithm", *Wikipedia*, [online]. Tillgänglig: https://en.wikipedia.org/wiki/Chudnovsky_algorithm [Hämtad: 2018-06-15].
- [10] F. Morgan, *Real Analysis*. USA: American Mathematical Society, 2005.
- [11] G. Blom, J. Enger, G. Englund, J. Grandell och L. Holst, *Sannolikhetsteori och Statistikteori: med tillämpningar*. Uppl 7. Lund: Studentlitteratur AB, 2017.
- [12] G. Casella och R. L. Berger, *Statistical Inference*. Uppl 2. USA: Duxbury, 2002.
- [13] "Slutskys theorem" *Wikipedia*, [online]. Tillgänglig: https://en.wikipedia.org/wiki/Slutsky%27s_theorem [Hämtad: 2018-05-24]
- [14] A. Person och L-C. Böiers, *Analys i en variabel*. Upplaga 3. Lund: Studentlitteratur AB, 2010.
- [15] J.V Uspensky, *Introduction to Mathematical Probability*. New York: McGraw-Hill Book Company, Inc, 1937.
- [16] "Estimating pi: throwing darts", *Google Sites*. [online]. Tillgänglig: <https://sites.google.com/site/appelacad/estimating-pi-throwing-darts> [Hämtad: 2018-05-24].