Универзитет у Београду, Математички факултет

Маја Теофиловић

Приближно израчунавање интеграла - детерминистички и стохастички приступ

МАСТЕР РАД

Београд, 2013.

Аутор: Маја Теофиловић
Ментор:
Др Весна Јевремовић
Чланови комисије:
Мр Марко Обрадовић
Др Зоран Станић

Садржај:

1. Увод

Из чињенице да се $\int_a^b f(x)dx$ не може увек елементарно израчунати произилази потреба за приближним израчунавањима. На пример, интеграл нормалне, Бета и Гама расподеле, се не могу на тај начин израчунати па ће у овом раду бити објашњене неке могућности израчунавања интеграла чије израчунавање није могуће елементарним путем или је добијена формула веома компликована. Постоје две могућности за приближно рачунање одређених интеграла.

Прва се заснива на детерминистичком приступу и ту спадају методе приближне интеграције развијене у оквиру нумеричке анализе. Други приступ је стохастички када се Монте Карло алгоритмом израчунавају одређени интеграли. У стохастички приступ спадају метод оцене средње вредности и погодак - промашај метод. Монте Карло алгоритми су једноставни и практични за употребу израчунавања вишедимензионалних интеграла. Решавање оваквих проблема заснива се на генерисању случајних бројева.

У поглављу нумеричка интеграција говори се о апроксимирању подинтегралне функције интерполационим полиномом што би значило да се одређени интеграл $\int_a^b f(x)dx$ приближно израчунава.

У поглављу основне интеграционе формуле помињу се Трапезна и Симпсонова интеграциона формула као један од начина нумеричке интеграције над одређеним интервалом $[a, b]$. Такође, говори се о гре7и нумеричке интеграције која се јавља приликом примене поменутих формула на одређеном интегралу. Биће описана примена интеграционих формула кроз примере као и процена грешке која се приликом тог израчунавања јавља.

У поглављу Монте Карло говори се о пореклу самог назива методе као и одакле метод потиче. Овај метод спада у стохастички приступ израчунавања одређеног интеграла који је ефикаснији од нумеричке интеграције јер је бржи приликом израчунавања одређеног интеграла.

На крају, у петом поглављу биће наведено упоређивање оба приступа, пример израчунавања одређеног интеграла детерминистичким и стохастичким приступом као и израчунавање вишеструких интеграла методом Монте Карло.

2. Нумеричка интеграција

Нумеричка интеграција је поступак израчунавања приближне вредности одређеног интеграла

$$I = \int_a^b f(x)dx \ \ (a < b) \tag{1}$$

из вредности подинтегралне функције датих уређеном табелом $(x_i, f(x_i))$, $i = 0, 1, ..., n$, при чему претпостављамо да је $x_0 \le a$ и $b \le x_n$. Слично као код нумеричког диференцирања, на уоченом одсечку $[a, b]$ функција $f(x)$ се апроксимира интерполационим полиномом $P_n(x)$, па следи да је

$$f'(x) \approx P_n'(x), \qquad a \le x \le b$$

Ако је функција довољно глатка и ако су растојања између чворова довољно мала, може се очекивати да се $f'(x)$ и $P_n'(x)$ не разликују много на одсечку $[a, b]$. Нумеричка интеграција се односи, као и нумеричко диференцирање, на апроксимацију подинтегралне функције интерполационим полиномом (ИП):

$$f(x) \approx P_m(x), \quad m \le n \qquad \int_a^b f(x)dx \approx \int_a^b P_m(x)dx \tag{2}$$

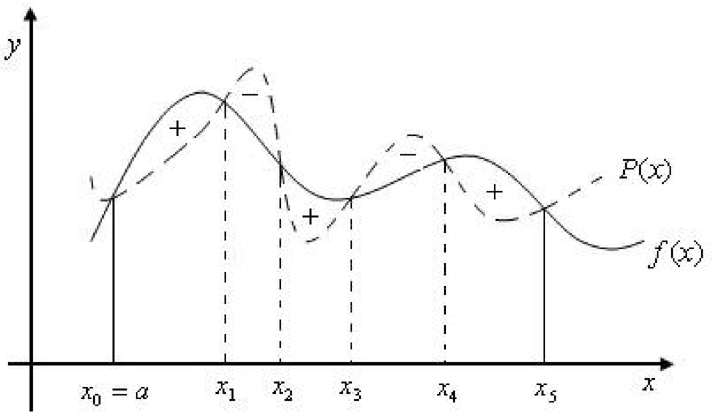
Ако претпоставимо да интерполациони полином пролази кроз све тачке табеле $\big(x_i, f(x_i)\big)$, $i = 0, 1, ..., n$ и да се границе интеграције поклапају са првом и последњом вредношћу независно промењиве у табели, тражени интеграл I рачунамо приближно као:

$$I = \int_{a=x_0}^{b=x_n} f(x)dx \approx \int_{x_0}^{x_n} P_n(x)dx = I_n \tag{3}$$

а грешка методе E_n је једнака интегралу грешке интерполације R_n:

$$E_n = |I - I_n| = \int_{x_0}^{x_n} f(x)dx - \int_{x_0}^{x_n} P_n(x)dx = \int_{x_0}^{x_n} R_n(x)dx \tag{4}$$

На слици 1.1 илустрована је нумеричка интеграција функције $f(x)$, интеграцијом интерполационог полинома са еквидистантним интерполационим чворовима, степена $n = 5$. Тачна вредност интеграла I једнака је површини испод криве $f(x)$ а приближна или нумеричка вредност I_n површини испод криве полинома $P_n(x)$ над интервалом интеграције $[a, b]$. Грешка нумеричке интеграције једнака је збиру грешака интеграције на појединачним подинтервалима ширине h.



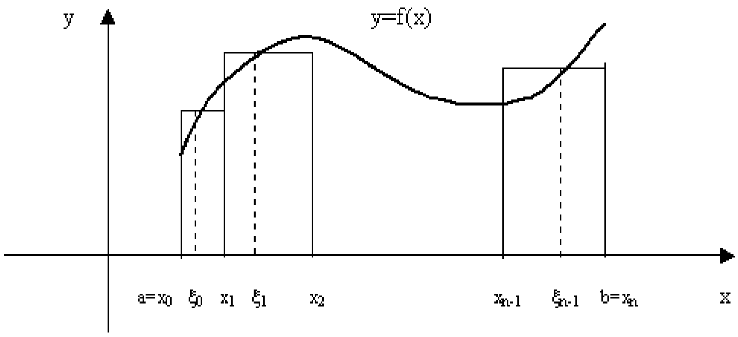
Слика 1.1: Нумеричка интеграција функције $f(x)$.

Грешка методе у неком подинтервалу ширине h између два суседна интерполациона чвора по апсолутној вредности је једнака површини између кривих $f(x)$ и $P(x)$. Видимо да грешке интеграције у појединим подинтервалима имају различите знаке па се у збиру делимично поништавају: долази до **међусобне компензације** грешака. Тако се за **нумеричку интеграцију** може рећи да је:

- нумеричка метода, **прецизнија од интерполације полинома**

- **стабилан** или добро услољен **рачунски процес** (грешке услед губитка значајних цифара на појединим подинтервалима се такође међусобно компензују)

Класичне методе нумеричке интеграције засноване су на геометријској интерпретацији интеграла (1) као области под функцијом $f(x)$ на интервалу $[a,b]$ (видети слику 1.2).



Слика 1.2: Интеграл I је једнак области испод криве $f(x)$, $f(x) > 0$, $x \in [a,b]$

Код ових метода x-оса је подељена на n једнаких интервала ширине Δx, где је

$$\Delta x = \frac{b-a}{n}$$

и

$$x_n = x_0 + n\Delta x$$

У претходним једнакостима $x_0 = a$ и $x_n = b$. Најједноставније приближно израчунавање области под функцијом $f(x)$ је познато као **правоугаона апроксимација**. Код правоугаоне апроксимације, $f(x)$ је оцењена на почетку интервала и апроксимација F_n интеграла је дата са

$$F_n = \sum_{i=0}^{n-1} f(x_i)\Delta x \qquad \textbf{правоугаона апроксимација}$$

Постоје различити начини за одређивање приближне вредности интеграла над одређеним доменом. Такође, постоје различити разлози због којих апроксимација може бити корисна. На пример, не може се свака функција интегралити. Чак, иако постоји интеграциона формула и даље то неће бити најпрецизнији начин за израчунавање интеграла.

У циљу да се добије неки увид нумеричке интеграције, природно је подсетити се Риманове интеграције, која може бити виђена као један од приступа за апроксимирање интеграла.
Претпостављамо да је $f(x)$ ограничена функција дефинисана на интервалу $[a,b]$ и да је $x_0,...,x_n$ подела P тог интервала. За свако i следи да је

$$M_i(f) = \sup_{x\in[x_{i-1},x_i]} f(x)$$

и

$$m_i(f) = \inf_{x\in[x_{i-1},x_i]} f(x)$$

Нека је $\Delta x_i = x_i - x_{i-1}$. Горња Дарбуова сума функције $f(x)$ која одговара подели P је дефинисана са

$$U(f,P) = \sum_{i=1}^{n} M_i \Delta x_i$$

док је доња Дарбуова сума функције $f(x)$ која се односи на поделу P дефинисана са

$$L(f,P) = \sum_{i=1}^{n} m_i \Delta x_i$$

Горњи интеграл функције $f(x)$ на интервалу $[a,b]$ је дефинисан као

$$U(f) = inf(U(f,P))$$

и доњи интеграл функције $f(x)$ је дефинисан као

$$L(f) = sup(L(f,P))$$

где инфимум и супремум узимају вредности над свим могућим поделама P интервала $[a,b]$. Ако су вредности горњег и доњег интерграла функције $f(x)$ једнаке уводи се заједничка ознака, интеграл $\int_a^b f(x)dx$, и при том се назива **Риманов интеграл функције** $f(x)$.

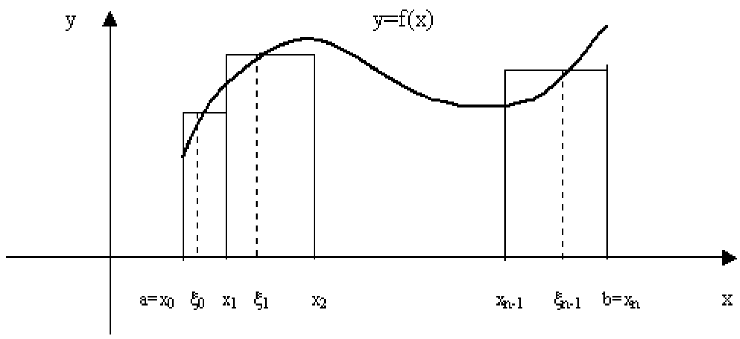
1. Појам одређеног интеграла

Уочимо интервал $[a, b] \in R$. Поделом сегмента $[a, b]$ називамо коначан скуп P тачака $P = \{x_0, x_1, ..., x_n\}$, тако да је $a = x_0 < x_1 < ... < x_n = b$. Скуп $P[a, b]$ свих подела сегмента $[a, b]$ уредићемо релацијом $\subseteq$. Ако је $P' \subseteq P$, рећићемо да је подела P финија од поделе P', односно да је подела P' грубља од поделе P. Са Δx_i означићемо $\Delta x_i = x_i - x_{i-1}, i = 1, 2, 3..., n$. Под параметром поделе подразумевамо $max\Delta x_i = \lambda(P)$.
Тачку ξ_i изаберемо на сваком сегменту $[x_{i-1}, x_i], i = 1, 2, 3, ..., n$. Скуп свих тачака означавамо са $\xi = \{\xi_1, \xi_2, ..., \xi_n\}$. На тај начин добија се подела са истакнутим тачкама (P, ξ) сегмента $[a, b]$.

Дефиниција: Нека је $f : [a, b] \to R$ и нека је (P, ξ) подела са истакнутим тачкама сегмента $[a, b]$. Збир

$$\sigma(f, P, \xi) = \sum_{i=1}^{n} f(\xi_i)\Delta x_i$$

назива се интегралном сумом функције f за дату поделу.



Слика 1.4: Подела (P, ξ) сегмента $[a, b]$, где је $\Delta x_i = x_i - x_{i-1}$, $i = 1, 2, ..., n$

Збир правоугаоника са основама $[x_{i-1}, x_i]$ и висинама $f(x_i)$ једнак је датој интегралној суми. Уочимо делове равни који садрже тачке које се налазе у правоугаоницима а не налазе се у криволинијском трапезу, као и тачке које се налазе у криволинијском трапезу, а имајући у виду интуитиван појам површине, можемо закључити да укупне површине поменутих делова могу бити произвољно мале, а да је површина криволинијског трапеза приближно једнака интегралној суми.

Дефиниција: За број I кажемо да је лимес (гранична вредност) интегралних сума $\sigma(f, P, \xi)$, функције $f : [a, b] \to R$ кад параметар поделе тежи нули и пишемо

$$\lim_{P(\xi)\to 0} \sigma(f, P, \xi) = I$$

ако за свако $\epsilon < 0$ постоји δ, тако да за сваку поделу са истакнутим тачкама $(P, \xi) \in P[a, b]$ важи неједнакост

$$|\sigma(f, P, \xi)| < \epsilon$$

кад год је параметар поделе $P(\xi)$ мањи од ϵ.

Ако лимес постоји и коначан је онда се каже да је функција f интеграбилна у Римановом смислу на сегменту $[a, b]$. Број

$$I = \lim_{P(\lambda)\to 0} \sigma(f, P, \xi)$$

назива се Римановим интегралом функције f на сегменту $[a, b]$ и пише се

$$I = \int_a^b f(x)dx$$

При том се:

- a и b називају доњом и горњом границом интеграла
- функција f назива се подинтегралном функцијом (интеграндом)
- израз $f(x)dx$ је подинтегрални израз

Промењљива x је интеграциона промењљива. Скуп свих функција, интеграбилних на сегменту $[a, b]$, означавамо са $R[a, b]$ и надаље такве функције ћемо називати интеграбилним функцијама. Риманов интеграл кратко ћемо назвати интегралом.
Приметимо да се интеграл I често назива и одређеним интегралом за разлику од скупа примитивних функција [1] неке функције, који смо називали неодређеним интегралом.

[1]Примитивном функцијом функције $f(x)$ називамо функцију $\varphi(x)$, $x \in (a, b)$, ако је она диференцијабилна и ако задовољава једнакост $\varphi'(x) = f(x)$, $x \in (a, b)$

2. Њутн - Котесове квадратурне формуле

Нека треба израчунати интеграл

$$I = \int_a^b f(x)dx$$

Налажење вредности интеграла назива се механичка квадратура или, краће, квадратура. Ако су у питању двоструки интеграли онда се назива механичка кубатура или само кубатура. Израчунавање одредјеног интеграла на основу низа познатих, израчунатих вредности $y_i = f(x_i)$, где су чворови $x_i \in [a, b]$, $i = 0, ..., n$ подинтегралне функције или интегранда $f(x)$, назива се **приближна** или **нумеричка интеграција**.

Могу се конструисати различите формуле за приближно израчунавање интеграла - **квадратурне формуле**. Често се те формуле добијају на следећи начин. Функција $f(x)$ се замени на одсечку $[a, b]$ или на његовим деловима другом, од ње једноставнијом функцијом $F(x)$, која је у неком смислу блиска функцији $f(x)$. На пример, тражи се да се функције $F(x)$ и $f(x)$ поклапају у чворовима тј. захтева се да је $F(x_i) = f(x_i)$, $i = 0, ..., n$. У својству функције узимају се алгебарски полиноми или тригонометријски полиноми или рационалне функције итд. Ако је интервал интеграције коначан и $f(x)$ на њему нема сингуларитета, онда се може постићи висока тачност са полиномом релатинво ниског степена.

Квадратурне формуле су често следећег облика

$$I = \int_a^b f(x)dx \approx \sum_{k=0}^{n} A_k f(x_k), \qquad x_k \in [a, b] \tag{5}$$

где се A_k називају коефицијентима а x_k чворовима квадратурне формуле. Смисао параметра n је очигледан: што је веће n, то се по правилу може достићи већа тачност одговарајућим избором A_k и x_k. Према томе, сматраћемо да је n фиксирано и размотрити задатак избора A_k и x_k пошто некад ни x_k не можемо бирати. Подинтегрална функција је дата таблично па су онда x_k фиксирани, па се могу бирати само A_k.

Чворови су изабрани на било који начин. Треба наћи коефицијенте A_k.

Конструишимо интерполациони полином $L_n(x)$ за функцију $f(x)$ са чворовима $x_k,\ k = 0, ..., n$. Дакле,

$$L_n(x) = \sum_{k=0}^{n} \frac{\Pi_{n+1}(x)}{(x - x_k)\Pi'_{n+1}(x_k)} f(x_k)$$

где је $\Pi_{n+1}(x) = (x - x_0)(x - x_1)...(x - x_n)$.

Сада имамо

$$f(x) = L_n(x) + R_n(x)$$

Заменом претходних једнакости добија се

$$I = \int_a^b f(x)dx = \int_a^b L_n(x)dx + \int_a^b R_n(x)dx$$

Имајући у виду једнакост (5) следи

$$I = \int_a^b f(x)dx \approx \int_a^b L_n(x)dx = \sum_{k=0}^{n} f(x_k) \int_a^b \frac{\Pi_{n+1}(x)}{(x - x_k)\Pi'_{n+1}(x_k)} dx$$

Дакле,

$$\int_a^b f(x)dx \approx \sum_{k=0}^{n} A_k f(x_k) \tag{6}$$

где се коефицијенти квадратурне формуле рачунају по формули

$$A_k = \int_a^b \frac{\Pi_{n+1}(x)}{(x - x_k)\Pi'_{n+1}(x_k)} dx \qquad k = 0, ..., n \tag{7}$$

Формуле (6) и (7) се називају **Њутн Котесове квадратурне формуле**. Приметимо да за дати распоред чворова коефицијенти A_k не зависе од функције $f(x)$.

Грешка квадратурне формуле је

$$R_n(f) = \int_a^b R_n(x)dx = \int_a^b \frac{f^{n+1}(\psi)}{(n+1)!} \Pi_{n+1}(x)dx$$

одавде је

$$|R_n(f)| \leq \frac{M_{n+1}}{(n+1)!} \int_a^b |\Pi_{n+1}(x)|dx$$

где је $|f^{n+1}(x)| \leq M_{n+1} = const.,\ x \in [a, b]$. Ако је $f(x) = P_m(x)$ и $m \leq n$, онда је $R_n(x) = 0$. Ако су границе интеграције a и b уједно и чворови квадратурне формуле, онда је квадратурна формула **затвореног типа** а у супротном случају је **отвореног типа**.

Чворови су еквидистантно распоређени.

Нека треба израчунати интеграл

$$I = \int_a^b f(x)dx$$

Нека су чворови $x_0, x_1, ..., x_n$ еквидистантни тј. нека је интервал интеграције подељен на n једнаких делова дужине

$$h = \frac{b-a}{n}$$

односно, нека су тачке $x_k = a + kh$, $k = 0, ..., n$ чворови интерполације. Тада квадратурна формула

$$I = \int_a^b f(x)dx \approx \sum_{k=0}^{n} A_k f(x_k)$$

следећег облика

$$I = \int_a^b f(x)dx \approx (b-a) \sum_{k=0}^{n} C_k^n f(a+kh)$$

где је,

$$C_k^n = \frac{A_k}{b-a} = \frac{1}{b-a} \int_a^b \frac{\Pi_{n+1}(x)}{(x-a-kh)\Pi'_{n+1}(a+kh)} dx$$

а

$$\Pi_{n+1}(x) = (x-a)(x-a-h)(x-a-2h)...(x-a-nh).$$

Ако уведемо смену $x = a + th (0 \leq t \leq n)$ онда се добија:

$$\Pi_{n+1}(x) = \Pi_{n+1}(a+th) = h^{n+1} t(t-1)...(t-n)$$

$$x - a - kh = th - kh = h(t-k)$$

$$\Pi'_{n+1}(a+th) = (x_k - x_0)...(x_k - x_{k-1})(x_k - x_{k+1})...(x_k - x_n) =$$

$$= (-1)^{n-k} h^n k!(n-k)!$$

па је

$$C_n^k = \frac{1}{b-a} \int_0^n \frac{h^{n+1} t(t-1)...(t-n)}{h(t-k)(-1)^{n-k} h^k k!(n-k)!} hdt$$

или

$$C_n^k = \frac{h}{b-a} \frac{(-1)^{n-k}}{k!(n-k)!} \int_0^n \frac{t(t-1)...(t-n)}{t-k} dt$$

На тај начин добијамо квадратурне формуле

$$\int_a^b f(x)dx \approx (b-a)\sum_{k=0}^{n} C_k^n f(a+kh)$$

са чворовима: $a = x_0, x_1 = x_0 + h, ..., x_k = x_0 + kh, ..., x_n = b$ и коефицијентима

$$C_k^n = \frac{1}{n}\frac{(-1)^{n-k}}{k!(n-k)!}\int_0^n \frac{t(t-1)...(t-n)}{t-k}dt$$

Пример: Извести квадратурне формуле облика

$$\int_0^1 f(x)dx = A_1 f(\frac{1}{4}) + A_2 f(\frac{1}{2}) + A_3 f(\frac{3}{4})$$

Решење: За дате чворове, горе наведене квадратурне формуле, треба одредити непознате коефицијенте A_1, A_2, A_3. Узимајући за подинтегралну функцију $f(x) = x^i, i = 0, 1, 2$ добија се систем

$$\int_0^1 dx = A_1 1 + A_2 1 + A_3 1$$

$$\int_0^1 xdx = A_1\frac{1}{4} + A_2\frac{1}{2} + A_3\frac{3}{4}$$

$$\int_0^1 x^2 dx = A_1(\frac{1}{4})^2 + A_2(\frac{1}{2})^2 + A_3(\frac{3}{4})^2$$

Након израчунавања одређених интеграла добија се систем

$$A_1 + A_2 + A_3 = 1$$

$$\frac{1}{4}A_1 + \frac{1}{2}A_2 + \frac{3}{4}A_3 = \frac{1}{2}$$

$$\frac{1}{16}A_1 + \frac{1}{4}A_2 + \frac{9}{16}A_3 = \frac{1}{3}$$

Одавде је решење $A_1 = \frac{2}{3}, A_2 = -\frac{1}{3}, A_3 = \frac{2}{3}$ па квадратурна формула гласи

$$\int_0^1 f(x)dx \approx \frac{2}{3}f(\frac{1}{4}) - \frac{1}{3}f(\frac{1}{2}) + \frac{2}{3}f(\frac{3}{4})$$

Приметимо да су добијени коефицијенти независни од подинтегралне функције и да је формула тачна за све полиноме чији је степен није већи од два.

3. Грешка оцењивања нумеричке интеграције

Добијамо од чега зависи процена грешке оцењивања одређеног интеграла са одређеним бројем интервала која се односи на метод нумеричке интеграције разматране у поглављу 2. Ове оцене се заснивају на претпостављеним одговарајућим Тејлоровим серијама интегранда $f(x)$

$$f(x) = f(x_i) + f'(x_i)(x - x_i) + \frac{1}{2}f''(x_i)(x - x_i)^2 + ... \tag{8}$$

и интеграције на интервалу $x_i \leq x \leq x_{i+1}$

$$\int_{x_i}^{x_{i+1}} f(x)dx = f(x_i)\Delta x + \frac{1}{2}f'(x_i)(\Delta x)^2 + \frac{1}{6}f''(x_i)(\Delta x)^3 + ... \tag{9}$$

Прво оцењујемо грешку која се односи на правоугаону апроксимацију са оцењеним интеграндом $f(x)$ на левом делу сваког интервала. Грешка Δ_i на интервалу $[x_i, x_{i+1}]$ је различита од претходне релације и оцене $f(x_i)\Delta x$:

$$\Delta_i = [\int_{x_i}^{x_{i+1}} f(x)dx] - f(x_i)\Delta x \approx \frac{1}{2}f'(x_i)(\Delta x)^2$$

Грешка на сваком интервалу је реда $(\Delta x)^2$. Пошто постоји укупан број n интервала и $\Delta x = (b - a)/n$, укупна грешка правоугаоне апроксимације је $n\Delta_i \sim n(\Delta x)^2 \sim n^{-1}$.
Оцењена грешка трапезоидне апроксимације може бити нађена на исти начин. Грешка на интервалу $[x_i, x_{i+1}]$ је различита од тачне оцене интеграла и оцене $\frac{1}{2}[f(x_i) + f(x_{i+1})]\Delta x$

$$\Delta_i = [\int_{x_i}^{x_{i+1}} f(x)dx] - \frac{1}{2}[f(x_i) + f(x_{i+1})]\Delta x \tag{10}$$

Користимо релацију (8) да оценимо интеграл и тако оцењено $f(x_{i+1})$ даје услов који није пропорционалан са f' и грешка реда $(\Delta x)^3$ је придружена једном интервалу. Укупна грешка на интервалу $[a, b]$ која се односи на трапезоидну апроксимацију је реда n^{-2}.
Пошто се Симпсоново правило заснива на погодном интегранду $f(x)$ на интервалу $[x_{i-1}, x_i]$ параболе, грешка је сразмерно пропорционална отказаном $f''(x)$. Може се очекивати да услов грешке у редоследу $f'''(x_i)(\Delta x)^4$ доприноси, али ови услови су отказани због њихове симетричности.

Отуда је $(\Delta x)^4$ услов Тејлоровог израза функције $f(x)$ адекватан Симпсоновом правилу. Ако се задржи услов $(\Delta x)^4$ у Тејлоровој серији, може се наћи грешка на интервалу $[x_i, x_{i+1}]$ која је реда $f''''(x_i)(\Delta x)^5$ па је онда укупна грешка на интервалу $[a,b]$ која се односи на Симпсоново правило реда $O(n^{-4})$.
Оцене грешке могу бити изведене у две димензије на сличан начин. Дводимензионални интеграл функције $f(x,y)$ је опсега под површином одређеном са $f(x,y)$. Код **правоугаоне** апроксимације, интеграл је записан као сума паралелограма са попречним пресеком $\Delta x \Delta y$ и висином одређеном са $f(x,y)$. Да би се утврдила грешка, проширује се $f(x,y)$ у Тејлоровој серији на

$$f(x,y) = f(x_i, y_i) + \frac{\partial f(x_i, y_i)}{\partial x}(x - x_i) + \frac{\partial f(x_i, y_i)}{\partial y}(y - y_i) + ...$$

и грешка се записује као

$$\Delta_i = [\iint f(x,y)\, dx\, dy] - f(x_i, y_i)\Delta x \Delta y$$

Ако из претходне две једнакости прву једнакост заменимо у другу налази се да је услов отказан који је пропорционалан са f и интеграл по $(x - x_i)dx$ је $\frac{1}{2}(\Delta x)^2$. Интеграл под условом са поштовањем dy даје фактор Δy. Интеграл који је пропорционалан са $y - y_i$ даје сличан допринос. Пошто је Δy реда Δx, грешка која се односи на интервале $[x_i, x_{i-1}]$ и $[y_i, y_{i+1}]$ је водећег реда у Δx:

$$\Delta_i \approx \frac{1}{2}[f'_x(x_i, y_i) + f'_y(x_i, y_i)](\Delta x)^3$$

Види се да је грешка која се односи на један паралелограм реда $(\Delta x)^3$. Пошто постоји n паралелограма, укупна грешка је реда $n(\Delta x)^3$. Међутим у две димензије, $n = A/(\Delta x)^2$ па је укупна грешка реда $n^{-1/2}$. Укупна гешка у једној димензији је реда n^{-1}, као што се могло пре видети.
Одговарајуће оцењене грешке у две димензије генерализујући трапезоидну апроксимацију и Симпсоново правило су редом реда n^{-1} и n^{-2}. У општем случају, ако се грешка креће до реда n^{-a} у једној димензији, онда грешка у d димензија се креће до $n^{-a/d}$. Монте Карло грешке се крећу до реда $n^{-1/2}$ независно од d. За довољно велико d, Монте Карло метод интеграције ће водити смањавању грешке за исти избор покушаја n.

3. Основне интеграционе формуле

Основне формуле за нумеричку интеграцију се добијају из једначине (3) за мале степене ИП $(n = 1, 2, 3)$ са еквидистантним интерполационим чворовима. Овако добијене основне интеграционе формуле су познате под називом **Њутн - Котесове формуле**. Оне се примењују на малим подинтервалима, ширине неколико корака h а интеграл на целом интервалу $[a, b]$ се онда добија као збир интеграла добијених применом основних формула (поглавље Трапезна и Симпсонова интеграциона формула). Корак интерполационих чворова h се у контексту нумеричке интеграције назива интеграциони корак. Њутн - Котесову формулу која се базира на ИП n-тог степена добијамо према једнакости (3) интеграцијом Њутновог ИП 1. степена уз смену интеграционе променљиве:

$$x = x_0 + \alpha h,\ dx = h d\alpha$$

Тако изводимо,

$$I_n = \int_{x_0}^{x_n} P_n(x)dx = h\int_0^n (y_0 + \alpha\Delta y_0 + \frac{\alpha(\alpha-1)}{2}\Delta^2 y_0 + ... + \frac{\alpha(\alpha-1)...(\alpha-n+1)}{n!}\Delta^n y_0)d\alpha \tag{11}$$

$$I_n = h[ny_0 + \tfrac{n^2}{2}\Delta y_0 + \tfrac{1}{2}(\tfrac{n^3}{3} - \tfrac{n^2}{2})\Delta^2 y_0 + (\tfrac{n^4}{24} - \tfrac{n^3}{6} + \tfrac{n^2}{6})\Delta^3 y_0 + ...],\ (n > 3)$$

Из последње једнакости добијамо **три основне интеграционе формуле**:

$$I_n = h[ny_0 + \tfrac{n^2}{2}\Delta y_0],\ \text{за}\ n = 1\ \ (11.\text{а})$$

$$I_n = h[ny_0 + \tfrac{n^2}{2}\Delta y_0 + \tfrac{1}{2}(\tfrac{n^3}{3} - \tfrac{n^2}{2})\Delta^2 y_0],\ \text{за}\ n = 2\ \ (11.\text{б})$$

$$I_n = h[ny_0 + \tfrac{n^2}{2}\Delta y_0 + \tfrac{1}{2}(\tfrac{n^3}{3} - \tfrac{n^2}{2})\Delta^2 y_0 + (\tfrac{n^4}{24} - \tfrac{n^3}{6} + \tfrac{n^2}{6})\Delta^3 y_0],\ \text{за}\ n = 3\ \ (11.\text{ц})$$

Грешка Њутн - Котесове формуле (12) се према (4) добија интеграцијом грешке интерполације што након смене интеграционе променљиве даје општи израз:

$$E_n = \frac{h^{n+2}}{(n+1)!}\int_0^n \alpha(\alpha-1)...(\alpha-n)f^{n+1}(\xi)d\alpha \tag{12}$$

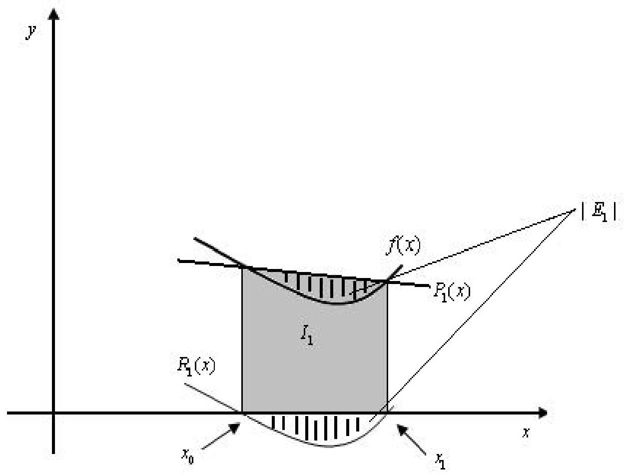
за $\xi \in (x_0, x_n)$.

1. Трапезно правило

Најједноставнија интеграциона формула је основна Трапезна формула која је позната и под називом трапезно правило. Добијамо га сменом за $n = 1$, $\Delta y_0 = y_1 - y_0$ у формули $(11.a)$

$$I_1 = \frac{h}{2}(y_0 + y_1) \tag{13}$$

Геометријска интерпретација формуле је површина трапеза са основама y_0 и y_1 и висином једнаком кораку интеграције h.



Слика 1.4: Геометријска интерпретација трапезног правила

Код трапезоидне апроксимације или правила, примена опште формуле (12) за грешку Трапезног правила даје:

$$E_1 = \tfrac{h^3}{12} f'(\xi),\ \xi \in (x_0, x_1)$$

2. Симпсоново прво и друго правило

Подинтегралну функцију на интервалу интеграције $[x_0, x_2]$ замењујемо квадратним ИП који пролази кроз еквидистантне чворове x_0, x_1, x_2. Резултат је формула (11.б) која након замене $n = 2$ и израза за коначне разлике даје **прво Симпсоново правило**:

$$I_2 = \frac{h}{3}(y_0 + 4y_1 + y_2) \tag{14}$$

Друго Симпсоново правило се базира на замени подинтегралне функције на интервалу интеграције $[x_0, x_3]$ кубном параболом и добијамо га са $n = 3$ из (11.ц):

$$I_3 = \frac{3h}{8}(y_0 + 3y_1 + 3y_2 + y_3) \tag{15}$$

Грешка Симпсонових правила

Јасно је да друго Симпсоново правило даје тачну вредност интеграла ако је функција $f(x)$ квадратни или кубни полином. Тада су сви чланови 4. и вишег реда (они који садрже коначне разлике 4. и вишег реда) у општој интеграционој формули (12) која се базира на ИП n-тог степена, једнаки нули.

Међутим,**и прво Симпсоново правило** даје **тачну вредност** интеграла ако је подинтегрална функција **полином 3. степена**, захваљујући томе што се фактор

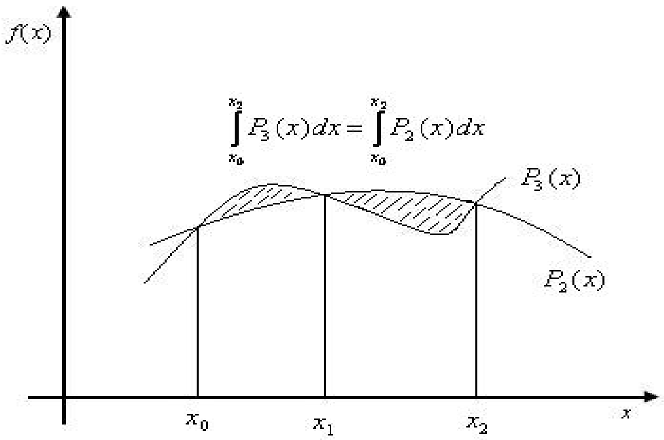
$$\tfrac{n^4}{24} - \tfrac{n^3}{6} + \tfrac{n^2}{6}$$

уз $\Delta^3 y$ у једначини за $n = 2$ анулира. Геометријски, то значи да долази до поништавања грешака при нумеричкој интеграцији полинома трећег степена на интервалу $[x_0, x_2]$ првом Симпсоновом формулом.

Закључујемо да су Симпсонова правила приближно једнако тачна, иако је Симпсоново друго правило сложеније. Према општој формули за грешку методе (12), грешка Симпсоновог другог правила, E_3 је пропорционална са $h^{n+2} = h^5$. Може се показати да је и грешка првог Симпсоновог правила E_2 такође пропорционална петом степену интеграционог корака, мада би се на основу (11) очекивало да она буде пропорционална са h^4. Тако се изводи:

$$E_2 = \tfrac{1}{90} h^5 f^4(\xi)\ ,\ \xi \in (x_0, x_2)$$

$$E_3 = \tfrac{3}{80} h^5 f^4(\xi),\ \xi \in (x_0, x_3)$$



Слика 1.5: Илустрација тачности првог Симпсоновог правила, ако је $f(x) = P_3(x)$

Каже се да су обе формуле истог реда тачности и зато прво Симпсоново правило које је једноставније а приближно тачно као и друго има далеко већу примену у пракси. Уопште, показује се да је Њутн - Котесова формула (11.а, 11.б, 11.ц), која се базира на ИП парног степена (n паран број) истог реда тачности као следећа по сложености формула базирана на ИП степена $(n+1)$.

3. Трапезна и Симпсонова интеграциона формула

Уместо да се Трапезно или Симпсоново прво правило примене на цео интервал интеграције $[a, b]$, што би дало лоше процене интеграла (због великог интеграционог корака h - грешка су пропорционалне степену корака), интервал интеграције се дели на више мањих подинтервала. Интеграл се добија као збир интеграла на појединим подинтервалима, који се рачунају применом једног или другог правила. Тако се добијају сложене интеграционе формуле или једноставно интеграционе формуле.

Трапезна формула

Делимо интервал интеграције $[a, b]$ на n једнаких подинтрвала ширине,

$$h = \frac{b-a}{n}$$

тачкама,

$$x_i = x_0 + ih,\ i = 0, 1, ..., n\ \ (x_0 = a, x_n = b)$$

и приближну вредност траженог одређеног интервала функције $f(x)$ добијамо као:

$$\int_{a=x_0}^{b=x_n} f(x)dx = I = \int_{x_0}^{x_n} f(x)dx + \int_{x_1}^{x_2} f(x)dx + ... + \int_{x_{n-1}}^{x_n} f(x)dx \approx \tag{16}$$

$$\approx \frac{h}{2}(y_0 + y_1) + \frac{h}{2}(y_1 + y_2) + ... + \frac{h}{2}(y_{n-1} + y_n) = I_n$$

где индекс n означава број интеграционих корака на интервалу интеграције $[a, b]$. Тако, Трапезна формула гласи:

$$I_n = \frac{h}{2}(y_0 + y_n + 2\sum_{i=1}^{n-1} y_i) \tag{17}$$

Грешка Трапезне формуле E_n једнака је збиру грешака трапезних правила на појединим подинтервалима:

$$E_n = \sum_{i=1}^{n} E_i = -\frac{h^3}{12}\sum_{i=1}^{n} f''(\xi_i),\ \xi_i \in (x_{i-1}, x_i)$$

Ако је $f''(x)$ на интервалу интеграције непрекидна функција, онда постоји ξ за које важи:

$$f''(\xi) = \frac{\sum_{i=1}^{n} f''(\xi_i)}{n},\ \xi \in (a, b)$$

па се сума других извода може приказати као:

$$\sum_{i=1}^{n} f''(\xi_i) = n,\ f''(\xi) = \frac{b-a}{h} f'(\xi)$$

и за грешку Трапезне формуле добијамо:

$$E_n = -\frac{(b-a)}{12} h^2 f''(\xi),\ \xi \in (a, b)$$

Трапезоидна апроксимација или правило даје приближну вредност површине под трапезом где функција $f(x)$ узима вредности у почетној и крајњој тачки интервала $[a, b]$.

Пошто је апроксимација области под кривом у тачкама x_i и x_{i+1} дата са $\frac{1}{2}[f(x_{i+1}) + f(x_i)]\Delta x$, онда је приближна вредност укупне области F_n једнака

$$F_n = [\frac{1}{2} f(x_0) + \sum_{i=1}^{n-1} f(x_i) + \frac{1}{2} f(x_n)]\Delta x \qquad \textbf{трапезно правило} \tag{18}$$

Симпсонова формула

Пошто се Симпсоновим првим правилом нумерички рачуна интеграл на интервалу ширине $2h$ (три тачке), интервал интеграције $[a, b]$ делимо на m једнаких подинтервала од којих сваки обухвата по два подинтервала ширине h тј. по три тачке. Очигледно је поступак применљив само ако је укупан број корака n на интервалу $[a, b]$ паран број (тј. укупан број тачака непаран) и тада имамо:

$$n = 2m = \frac{b-a}{h}$$

Тада, приближну вредност траженог интеграла добијамо као збир приближних вредности интеграла на појединим подинтервалима, ширине $2h$, добијених Симпсоновим првим правилом:

$$I = \int_a^b f(x)dx \approx \frac{h}{3}(y_0 + 4y_1 + y_2) + \frac{h}{3}(y_2 + 4y_3 + y_4) + ... + \frac{h}{3}(y_{n-2} + 4y_{n-1} + y_n) = I_n \tag{19}$$

Након сређивања добијене суме, долазимо до **Симпсонове интеграционе формуле**

$$I_n = \frac{h}{3}(y_0 + y_n + 4 \sum_{i=1,\Delta i=2}^{n-1} y_i + 2 \sum_{i=2,\Delta i=2}^{n-2} y_i) \qquad (20)$$

где индекс n означава укупан број интеграционих корака на интервалу $[a, b]$. Прва сума у загради представља збир свих "непарних" координата од 1. до $(n-1)$ а друга сума збир свих "парних" координата од 2. до $(n-2)$:

$$\sum_{i=1,\Delta i=2}^{n-1} y_i = y_1 + y_3 + ... + y_{n-1}, \sum_{i=2,\Delta i=2}^{n-2} y_i = y_2 + y_4 + ... + y_{n-2} \qquad (21)$$

Да поновимо да је ограничење за примену Симпсонове формуле паран број подинтервала, односно непаран број тачака на интервалу интеграције $[a, b]$.
Ако се позовемо на пример полинома другог реда, укупна површина под свим сегментима параболе даје апроксимацију укупне површине (**Симпсоново правило**):

$$F_n = \tfrac{1}{3}[f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + ...$$

$$+2f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)]\Delta x$$

Симпсоново правило је погодно за функцију $f(x)$ која може бити полиномна. Ако је функција $f(x)$ таква, можемо оценити површину испод ње за дати број n интервала па дуплирати број интервала и поновити процену. Ако су две процене приближне једна другој, можемо стати са повећавањем броја интервала. У супротном, поново ћемо дуплирати n док не достигнемо жељену тачност. Наравно, стратегија не важи ако $f(x)$ дивергира. Пример функције која дивергира је $f(x) = x^{-\frac{1}{3}}$ за $x = 0$.
Грешка Симпсонове интеграционе формуле се добија као збир грешака Симпсоновог првог правила на појединим подинтервалима ширине $2h$.

$$E_n = -\tfrac{b-a}{180} h^4 f^4(\xi),\ \xi \in (a, b)$$

4. Грешка и ред интеграционе формуле

Било која интеграциона формула изведена из неког од Њутн - Котесових правила има облик:

$$I_n = h \sum_{i=0}^{n} w_i y_i \tag{22}$$

где су $w_i, i = 0, 1, ..., n$ константе које се зову пондери или тежине а сума се зове пондерисана сума. При томе, сума свих тежина је једнака n:

$$\textstyle\sum_{i=0}^{n} w_i = n$$

На пример за Трапезну формулу

$$w_0 = w_n = \frac{1}{2}, w_i = 1, i = 1, 2, ..., n-1$$

па је

$$\sum_{i=0}^{n} w_i = w_0 + w_1 + ... + w_{n-1} + w_n = \frac{1}{2} + 1 + 1 + ... + 1 + \frac{1}{2} = 1 + n - 1 = n$$

Може се извести да **грешка опште интеграционе формуле** има облик

$$E_n = c(b-a)h^N f^N(\xi), N > 1, \xi \in (a, b) \tag{23}$$

где је c нека константа која је на пример за Симпсонову формулу једнака

$$c = -\frac{1}{180}$$

На основу израза закључујемо

- **грешка** неке **методе опада** са **смањивањем** интеграционог **корака** h
- што је **експонент** N у изразу за грешку неке методе **већи** утолико је при датом интеграционом кораку та **метода тачнија**

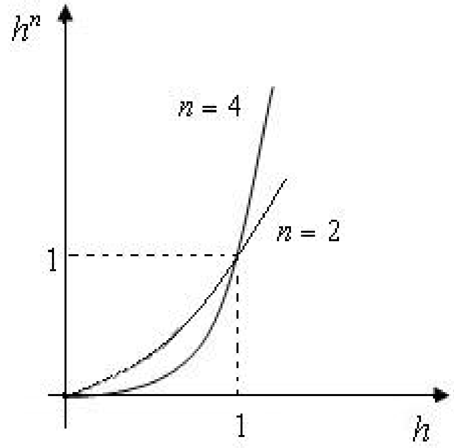
Други закључак је очигледан за интеграционе кораке мање од 1, јер је:

$$h^{N_1} < h^{N_2},\ (h < 1),\ \text{ако је } N_1 > N_2$$

али пошто је

$$h^{N_1} > h^{N_2},\ (h > 1),\ \text{ако је } N_1 > N_2$$

онда, ако је корак h већи од 1, важи обрнуто: уколико је експонент N већи, метода је мање тачна.



Слика 1.6:Понашање функције h^n

Уочили смо дакле да је грешка неке интеграционе формуле пропорционална са h^N , па када h тежи нули постаје бесконачно мала величина истог реда као бесконачно мала величина h^N, под условом да је N-ти извод ограничен у интервалу интеграције. То значи:

$$\lim_{h\to 0}\frac{E_n}{h^N} = \lim_{h\to 0}\frac{c(b-a)(\frac{h}{b-a})^N {f_z}^N(\zeta)}{h^N} = \tag{24}$$

$$= \lim_{h\to 0}\frac{kh^N}{h^N} = k \neq 0, \propto \tag{25}$$

где је $k = c(b-a)^{N-1} {f_z}^N(\zeta)$ и пишемо:

$$E_n = O(h^N)$$

За неку интеграциону формулу кажемо да је интеграциона формула реда N, ако је њена грешка пропорционална са h^N. Уколико је ред формуле већи, она је тачнија. Док је Трапезна формула другог реда, следећа по сложености, Симпсонова формула је четвртог реда, дакле знатно тачнија од Трапезне формуле.

Пример 1: Проценити нумерички вредност интеграла

$$\int_0^1 \frac{dx}{1+x^2} (= \frac{\pi}{4})$$

помоћу:
а) Трапезне формуле
б) Симпсонове формуле
са кораком интеграције 0.1. За обе методе проценити грешку добијене вредности интеграла користећи формуле. Вредности подинтегралне функције рачунати са 5 сигурних цифара.

Решење:

Трапезна формула $I_T = \frac{h}{2}(y_0 + y_n + 2\sum_{i=1}^{n-1} y_i)$

Симпсонова формула $I_S = \frac{h}{3}(y_0 + y_n + 4\sum_{i=1,\Delta i=2}^{n-1} y_i + 2\sum_{i=2,\Delta i=2}^{n-2} y_i)$

Корак интеграције: $h = 0.1$
Границе интеграције: $a = 0$ и $b = 1$ [$x_0 = 0$ и $x_n = 1$]
Број подинтервала: $n = 10$
Тачна вредност интеграла: $I = 0.78539$
Вредности подинтегралне функције са 5 сигурних цифара:

i	x_i	y_i
1	0.1	0.99001
2	0.2	0.96154
3	0.3	0.91743
4	0.4	0.86207
5	0.5	0.8
6	0.6	0.73529
7	0.7	0.67114
8	0.8	0.60976
9	0.9	0.55249

Процена интеграла на 5 значајних цифара и њихове грешке:

$I_T = 0.78498 \qquad E_T = I - I_T = 0.000416$
$I_S = 0.78542 \qquad E_S = I - I_S = 0.00003$
Пошто је грешка $< 0.5 \times 10^{-3}$, трапезна формула је дала процену са 3 сигурне цифре а Симпсонова формула је дала тачан резултат на свих 5 значајних цифара.

Процена границе апсолутне грешке Трапезне формуле на основу $|E| < \frac{b-a}{12}h^2 M, \qquad M = |f''(\xi)|_{max}, \qquad \xi \in (a,b)$

Потребно је наћи највећу вредност другог извода $f''(x) = \frac{8}{(1+x^2)^3}x^2 - \frac{2}{(1+x^2)^2}$ на интервалу $[a,b]$.
Други извод криве $f(x) = \frac{1}{1+x^2}$ је монотона функција и има највећу апсолутну вредност на доњој граници $M = |f''(x)| = 2$ за $x = 0$.

Процена границе апсолутне грешке Симпсонове формуле помоћу

$$|E| < \frac{b-a}{180}h^4 M,\ M = |f^4(\xi)|_{max},\ \xi \in (a,b).$$

Потребна нам је највећа вредност четвртог извода функције
$f^4(x) = \frac{384}{(1+x^2)^5}x^4 - \frac{288}{(1-x^2)^4}x^2 + \frac{24}{(1+x^2)^3}$на интервалу $[0,1]$.
Функција има највећу апсолутну вредност на доњој граници $M = |f^4(a)| = 24$ за $x = a$.

5. Процена грешке методе - Ричардсонова екстраполација

Процењивање грешака нумеричке интеграције помоћу формуле (17) није једноставно јер захтева познавање N-тог извода подинтегралне функције и његовог екстрема, а као резултат даје често знатно **прецењене грешке**. Ричардсон је предложио следећи поступак за процену грешке, који се базира на претпоставци да се вредност N-тог извода у формули (17) **не мења много** на интервалу интеграције. Нека су:

- I_n, E_n - процена интеграла добијена интеграционом формулом (17) са бројем подинтервала n, односно са кораком $h - (b-a)/n$ и њена грешка

- I_{2n}, E_{2n} - процена интеграла са дупло мањим кораком $h = h/2$, односно са дупло већим бројем подинтервала $2n$ и грешка те процене.

Следи да је тачна вредности интеграла I

$$I = I_n + E_n = I_{2n} + E_{2n}$$

а одатле

$$E_n - E_{2n} = I_{2n} - I_n \tag{26}$$

Однос грешака те две процене је уз дату претпоставку једнак:

$$\frac{E_{2n}}{E_n} = \frac{c(b-a)(\frac{h}{2})^N f^N(\xi_1)}{c(b-a)h^N f^N(\xi_2)} = \frac{1}{2^N} \tag{27}$$

и имамо две једначине из којих можемо да нађемо грешке. Из претходне једнакости следи да је $E_{2n} = \frac{E_n}{2^N}$ и смена тог израза у (26) даје једначину по E_n из које добијамо:

$$E_n = \frac{I_{2n} - I_n}{1 - 1/2^N}$$

и

$$E_{2n} = \frac{I_{2n} - I_n}{2^N - 1}$$

Тако можемо да проценимо тачну вредност интеграла:

$$I = I_{2n} - E_{2n} = I_{2n} + \frac{I_{2n} - I_n}{2^N - 1} \tag{28}$$

За **Трапезну формулу** која је другог реда добијамо изразе за грешку и процену тачне вредности интеграла сменом $N = 2$ у опште изразе:

$$E_{2n} = \frac{1}{3}(I_{2n} - I_n), \qquad I = I_{2n} + E_{2n} = \frac{4}{3}I_{2n} - \frac{1}{3}I_n$$

За **Симпсонову формулу** сменом $N = 4$ добијамо

$$E_{2n} = \frac{1}{15}(I_{2n} - I_n), \qquad I = I_{2n} + E_{2n} = \frac{16}{15}I_{2n} - \frac{1}{15}I_n$$

Пример 2: Проценити грешку интеграла израчунатог Трапезном формулом из примера 1. Ричардсоновом екстраполацијом.

Вредност подинтегралне функције израчунате у примеру 1. користимо у овом примеру, где је корак интеграције $h = 0,1$. Добијена вредност интеграла Трапезном формулом је $I_T = 0.78498$. Потребна нам је још једна вредност интеграла добијена са дупло мањим или дупло већим интеграционим кораком.
Једноставније је рачунање интеграла са дупло већим кораком јер не захтева нове вредности подинтегралне функције. Тако је

$$I_{2n} = I_T$$

Следи рачунање интеграла са дупло мањим бројем подинтервала, $n = 5$. У суму улази свака друга ордината тј. само парне ординате $(y_0, y_2, ..., y_{10})$.

i	x_i	y_i
1	0.2	0.96154
2	0.4	0.86207
3	0.6	0.73529
4	0.8	0.60976

$$h_n = 2h \qquad I_n = \frac{h_n}{2}\left(y_0 + y_n + 2\sum_{i=1}^{\frac{n-2}{2}} y_{2i}\right) = 0.78373$$

По Трапезној формули вредност интеграла је $I_T = 0.78498$ па би онда процена грешке Трапезне формуле била $E_{2n} = \frac{I_{2n} - I_n}{3} = 4.2 \times 10^{-4}$ где је добијена реална процена грешке на свих 5 децимала.

4. Монте Карло

"Нумеричку методу Монте Карло карактерише решавање математичких проблема помоћу моделирања случајних променљивих и статиситичког оцењивања карактеристика тих променљивих. Назив методе потиче из чланка *"The Monte Carlo method", N. Metropolis, S.M.Ulam, 44, No.247, 335-341*, из 1949. године. Сматра се да је метода Монте Карло коришћена и раније. Развој рачунарске технике допринео је значајном повећању коришћење методе Монте Карло. Са развојем рачунара метода Монте Карло је постала применљивија јер је иначе "ручно" моделирање вредности случајних променљивих веома спор процес.

Област примене методе Монте Карло се пуно проширила последњих година. Неке од тих области примене су: биологија, екологија, хидрологија, атомска физика, статистичка физика, статистика, системи масовног опслуживања (планирање капацитета складишта, лука, паркинга...) , теорија поузданости система (планирање електро енергецке мреже)...

Проблеми који се срећу у разним областима се могу "превести" или свести на математичке проблеме: решавање система линеарних једначина или неједначина, решавање интеграла (једноструких или вишеструких), решавање диференцијалних једначина, решавање парцијалних диференцијалних једначина итд. Сваки од наведених математичких задатака се може решити и методом Монте Карло, што се посебно користи кад је теоријско решење сувише компликовано или не може да се одреди, иако знамо да постоји. Такође је значајно да су алгоритми Монте Карло по правилу једноставни и лаки за програмирање. Важно је нагласити да се методом Монте Карло може послужити да се нека "теорија" оповргне, али наравно не може представљати доказ да је нека "теорија" исправна (добијени резултати могу само бити индиција истраживачу да је на добром путу)."(из [1])

Монте Карло методе су класа алгоритама за израчунавања која се ослањају на поновно случајно бирање узорка за израчунавање резултата. Често се користе за симулирање математичких и физичких система. Монте Карло симулација одсликава стохастичке процесе код којих време не игра улогу. Она се означава и као метода поновљених покушаја. Нумеричке методе, познате као Монте Карло методе, могу се слободно дефинисати као статистички симулациони методи, код којих се употребљавају низови случајних бројева за извршење симулације. Метода Монте Карло последњих неколико деценија добија статус потпуно заокружене нумеричке методе способне за решавање најкомплекснијих захтева.

Назив Монте Карло произилази из назива чувеног казина у Монаку. Коришћење случајности и процеса понављања у методи је аналогно активностима који се догађају у казину.

Монте Карло метода се примењује у разним симулацијама које користе случајне бројеве. Најчешће се метода користи само за статичке типове симулација код којих се у решавању проблема користи стварање узорака из расподела случајних промењљивих. При томе, проблеми могу бити било детерминистичког, било стохастичког карактера. Разликујемо следеће примене Монте Карло симулације:

- Детерминистички проблеми које је тешко или скупо решавати. Типичан пример је коришћење овог метода за израчунавање одређених интеграла који се не могу решити аналитички. Монте Карло метода се у овом случају може користи тако што се генерише низ случајних тачака (x_j, y_j) са једнаким вероватноћама унутар одредјеног правоугаоника. Затим испитује колико је генерисаних тачака садржано у површини која одговара интегралу и вредност интеграла се оцењује одговарајућом релативном фрекфенцијом.

- Сложени феномени који нису довољно познати. Ово је друга класа проблема које се решавају уз помоћ Монте Карло метода. Вероватноће се користе за извођење низа екперимената који дају узорке могућих стања зависних промењљивих. Статистичком обрадом резултата добија се расподела вероватноћа зависних промењљивих које су од интереса. Друштвени и економски проблеми се решавају на овај начин.

- Статистички проблеми који немају аналитичка решења. Статистички проблеми без аналитичког решења (на пр. процене критичних вредности или тестирање нових хипотеза) су једна специфична класа проблема који се решавају Монте Карло симулацијом. Решавање таквих проблема такође се заснива на генерисању случајних бројева и промењљивих."

1. Погодак-промашај метод

Сада ће бити истражен другачији метод оцењивања интеграла $\int_a^b f(x)dx$. Изаберимо правоугаоник над $[a,b]$ чија је висина h и површина $A = h(b-a)$. Функција $f(x)$ је ограничена границама правоугаоника. У такозваној Монте Карло методи погодак-промашај одређени интеграл је оцењен на квадрату који ограничава функцију $f(x)$ на интервалу $[a,b]$. Квадрат се креће од a ка b и од 0 до f_{max} где је $f_{max} > f(x)$ на целом интервалу. Постоји n тачака које су одабране на случајан начин на датом квадрату. Оцена интеграла може бити добијена рачунањем укупног броја тачака n_S које се налазе испод криве $f(x)$. Следи да је оцена F_n у методи погодак - промашај дата са

$$F_n = A\frac{n_S}{n}$$

где је $A = (b-a)f_{max}$ област на којој се тачке бирају. Сваких n тачака је добијено генерисањем униформно распоређених бројева S_1 и S_2 тако да је

$$x = a + S_1\dot{(}b-a) \qquad y = S_2\dot{f}_{max}$$

.

Оцена интеграла постаје прецизнија како број покушаја $n \to \infty$ па ће евентуално конвергирати тачном одговору тј. решењу. Квалитет одговора зависи од низа случајних бројева који се приликом израчунавања користе. Могу се добити независне оцене понављајући израчунавања користећи различите низове случајних бројева. Упоређујући вредности добија се идеја прецизног израчунавања.

2. Оцена средње вредности

Други Монте Карло интеграциони метод је заснован на теореми о средњој вредности где је интеграл одређен просечном вредности интегранда $f(x)$ на интервалу $a \leq x \leq b$. Да би одредили ову просечну вредност, бирамо тачке x_i на случајан начин као и вредности функције у тим тачкама и рачунамо одговарајуће $f(x_i)$. Код једноструких интеграла оцена F_n интеграла у методи оцена средње вредности је дата са

$$F_n = (b-a)(f) = (b-a)\tfrac{1}{n}\sum_{i=1}^{n} f(x_i)$$

Тачке x_i су случајне вредности распоређене униформно на интервалу $a \leq x_i \leq b$ и n је број покушаја.

3. Монте Карло анализа грешке

Методе нумеричке интеграције и Монте Карло метод дају приближна решења чија тачност зависи од броја покушаја. До сада се користило знање о тачној вредности која би се добила на различитим интервалима да би се утврдила грешка Монте Карло методе која тежи нули као приближна вредност $n^{-1/2}$ за велико n, где је n број покушаја. Треба видети како оценити грешку када је тачна вредност интеграла непозната. Зависност грешке је независна од природе интеграла, и много важније, не зависи од димензије интеграла.
Посматрајмо Монте Карло оцењивање интеграла функције $f(x) = 4\sqrt{1-x^2}$ на интервалу $[0,1]$. Добијени резултат је $F_n = 3.1489$ за случајан низ од $n = 10^4$ покушаја користећи метод средње вредности. Упоређујући F_n са тачним резултатом $F = \pi \approx 3.1416$, налазимо да је грешка придружена $n = 10^4$ покушајима апроксимативно 0.0073.

Ако је интегранд константна функција онда би грешка била нула, па би F_n било једнако F за било које n. Зашто то важи? Понашање функција указују да је могућа величина грешке **дисперзија** σ^2, дефинисана са

$$\sigma^2 = f^2 - (f)^2 \tag{29}$$

где је

$$f = \frac{1}{n}\sum_{i=1}^{n} f(x_i) \qquad f^2 = \frac{1}{n}\sum_{i=1}^{n} f(x_i)^2$$

Из дефиниције **стандардне девијације** σ, видимо да ако је f независно од x, вредност за σ је нула. За пример који је претходно поменут и за низ случајних бројева добија се $F_n = 3.1489$, па је $\sigma = 0.8850$. Пошто је вредност σ два реда величина већа од грешке, закључићемо да σ не може бити директна величина грешке. Уместо тога, σ је мера колико функција $f(x)$ варира на посматраном интервалу.

Један од начина како наћи одговарајућу меру грешке је пратити вредности грешке како опада ако n расте. У табели 1.1 видимо да како се n креће од 10^2 до 10^4, грешка опада по броју покушаја n тј. као $\sim 1/n^{\frac{1}{2}}$. Међутим, такође видимо да је σ_n константна и много већа од посматране грешке.

n	F_n	грешка	σ_n
10^2	3.0692	0.0724	0.8550
10^3	3.1704	0.0288	0.8790
10^4	3.1489	0.0073	0.8850

Табела 1.1: Примери Монте Карло мерења за оцену средње вредности функције $f(x) = 4\sqrt{1-x^2}$ на интервалу $[0,1]$. Права грешка је дата разликом $|F_n - \pi|$. Стандардна девијација σ_n је нађена из једнакости (26).

Сваким низом од n покушаја добија се средња вредност мерења коју означавамо са M_α. У општем случају, мерења нису једнака јер свако мерење се односи на различити низ случајних бројева. Табела 1.2 приказује резултате од десет мерења за $n = 10^4$ покушаја. Видимо да грешка варира од мерења до мерења. Квалитативно, величина разлика међу мерењима је слична правим грешкама и ове разлике су вредност грешке придружене сваком појединачном мерењу.

низ α	M_α	права грешка
1	3.1489	0.0073
2	3.1326	0.0090
3	3.1404	0.0012
4	3.1460	0.0044
5	3.1526	0.0110
6	3.1311	0.0105
7	3.1311	0.0105
8	3.1358	0.0058
9	3.1344	0.0072
10	3.1405	0.0011

Табела 1.2: Пример Монте Карло мерења средње вредности функције $f(x) = 4\sqrt{1-x^2}$ на интервалу $[0,1]$. Урађен је тотал за свако од 10 мерења за $n = 10^4$ покушаја. Приказани су средња вредност M_α и праве грешке $|M_\alpha - \pi|$ за свако мерење.

Да би се добила квантитативна мера грешке, одредићемо разлике ових мерења користећи стандардну девијацију средње вредности σ_m која се дефинише као

$$
{\sigma_m}^2 = M^2 - (M)^2 \tag{30}
$$

где је

$$
M = \frac{1}{m} \sum_{\alpha=1}^{m} M_\alpha
$$

и

$$
(M^2) = \frac{1}{m} \sum_{\alpha=1}^{m} {M_\alpha}^2
$$

Мада σ_m даје вероватноћу грешке, наш метод добијања σ_m додавањем нових мерења је непрактичан, пошто би требало комбиновати додатна мерења да би се добила боља оцена. Следи релација

$$
\sigma_m = \frac{\sigma}{\sqrt{n-1}} \approx \frac{\sigma}{\sqrt{n}} \tag{31}
$$

Један начин да се провери релација (31) је да се иницијална мерења од n покушаја поделе на s подскупова. Ова процедура не захтева додатна мерења. Забележимо да је средња вредност за $f(x_i)$ у k-том подскупу означена као S_k. Као пример, поделићемо 10^4 покушаја првог мерења на $s = 10$ подскупова и $n/s = 10^3$ покушаја за сваки подскуп. Средња вредност функције $f(x)$ за сваки подскуп k није једнака. Да бисмо измерили грешку користићемо стандардну девијацију средње вредности за сваки подскуп. Означимо ову величину са σ_s где је

$$
{\sigma_s}^2 = (S^2) - (S)^2
$$

где се просечна вредност односи на подскупове. Закључујемо да можемо протумачити n покушаја или као појединачно мерење или као скуп мерења од n/s покушаја. Вероватноћа грешке је дата стандардном девијацијом за n покушаја подељеном са квадратним кореном броја покушаја. Друго тумачење подразумева да је вероватноћа грешке дата стандардном девијацијом за s мерења која се односе на подинтервале подељене са квадратним кореном броја мерења.

Можемо направити грешку што мању са повећањем броја покушаја или повећањем ефикасности индивидуалних покушаја смањујући стандардну девијацију.

4. Важност бирања узорка

У претходним секцијама је показано да је грешка Монте Карло оцене пропорционална σ и инверзно пропорционална квадратном корену броја покушаја. Постоје два начина да се грешка смањи у Монте Карло оцењивању - повећавати број покушаја или смањити дисперзију. Други избор је пожељнији јер не захтева много компијутерског времена. Уводимо технике **важности бирања узорка** које смањују σ и побољшавају ефикасност сваког покушаја.
У контексту нумеричке интеграције, уводимо позитивну функцију $p(x)$ тако да је

$$\int_a^b p(x)dx = 1$$

и ако интеграл запишемо као

$$F = \int_a^b [\frac{f(x)}{p(x)}]p(x)dx$$

Интеграл се може оценити бирањем узорка у складу са расподелом вероватноће $p(x)$ па добијамо суму

$$F_n = \frac{1}{n}\sum_{i=1}^{n}\frac{f(x_i)}{p(x_i)}$$

Сума F_n постаје оцена средње вредности за функцију густине вероватноће $p(x) = 1/(b-a)$. Бирамо облик за $p(x)$ која минимизује дисперзију интегранда $f(x)/p(x)$. Пошто не можемо да оценимо σ аналитички у општем случају, одређујемо σ апостериори и бирамо облик $p(x)$ који имитира $f(x)$ што је више могуће, где је $f(x)$ веће. Избор замене $p(x)$ би успорио варирање интегранда $f(x)/p(x)$ и дисперзија ће бити смањена. Као пример, посматрајмо интеграл

$$F = \int_0^1 e^{-x^2}dx$$

Оцењивање интеграла F са $p(x) = 1$ за $0 \leq x \leq 1$ је приказано у првој колони табеле 1.4. Разуман избор тежине функције је $p(x) = Ae^{-x}$, где је A одабрана тако да је $p(x)$ слично $f(x)$. Вредности приказане у другој колони табеле 1.4 су тражене вредности. Видимо да је време израчунавања по покушају за неуниформни случај веће па мање вредности σ чине корисним неуниформну расподелу вероватноће још ефикаснијом.

	$p(x) = 1$	$p(x) = Ae^{-x}$
n покушаја	4×10^5	8×10^3
F_n	0.7471	0.7469
σ	0.2010	0.0550
$\sigma/\sqrt{n}$	3×10^{-4}	6×10^{-4}
Укупно време CPU	35	1.35
CPU време по покушају	10^{-4}	2×10^{-4}

Табела 1.4: Поређење Монте Карло оцена интеграла коришћењем униформне густине вероватноће $p(x) = 1$ и неуниформне густине вероватноће $p(x) = Ae^{-x}$. Нормализација константе A је одасбрана тако да је $p(x)$ нормализовано на јединичном интервалу. Вредност интеграла са четри децимална места је 0.7468. Оцењивање F_n дисперзије σ и вероватноћа грешке $\sigma/n^{1/2}$ су приказани. Време CPU је приказано једино за упоређивање.

Алтернативни приступ користи познату вредност функције $f(x)$ на датом интервалу за избор узорка чешће где је вредност функције $f(x)$ већа. Идеја је да се функцијом $f(x)$ одреди вероватноћа за бирани узорак, па једино узимамо у разматрање интегранде $f(x)$ који су ненегативни. Да би се израчунала груба оцена релативне вредности функције $f(x)$, прво рачунамо њену просечну вредност узимајући вредност за k једнако размакнутим тачкама s_i и израчунавања

$$S = \sum_{i=1}^{k} f(s_i)$$

Сума подељена по k даје просечну вредност функције f на интервалу. Приближну вредност даје интеграл F па следи $F \approx Sh$, где је $h = (b-a)/k$. Ова апроксимација интеграла је једнака правоугаоној апроксимацији. Следи генерисање n случајних узорака. Вероватноћа бирања подинтервала i је дата са

$$p_i = \frac{f(s_i)}{S}$$

Збир над свим подинтервалима вероватноћа p_i је једнак јединици. Да би се одабрао подинтервал са жељеном вероватноћом, генеришемо случајан број r униформно на интервалу $[a, b]$ и одређујемо подинтрвал i који задовољава неједнакост (30). Подинтервал је одабран са жељеном вероватноћо па генеришемо случајан број x_i на подинтервалу $[s_i, s_i + h]$ и израчунавамо однос $\frac{f(x_i)}{p(x_i)}$.

Оцена интеграла је дата следећим условима. Вероватноће p_i у претходној релацији су вероватноће одабира подинтервала i, не вероватноћа одабира вредности x између x и $x+\Delta x$. Вероватноћа $p(x)\Delta x$ је p_i пута вероватноћа бирања посебне вредности за x на подинтервалу i:

$$p(x_i)\Delta x = p_i \times \frac{\Delta x}{h}$$

Имамо да је

$$F_n = \frac{1}{n}\sum_{i=1}^{n}\frac{f(x_i)}{p(x_i)} = \frac{h}{n}\sum_{i=1}^{n}\frac{f(x_i)}{p_i}$$

5. Класична Монте Карло интеграција

Пре примењивања технике симулације на практичним проблемима, подсетимо се својства које оправдавају њихово коришћење. Општи проблем је око оцењивања интеграла

$$E_f[h(X)] = \int_{\chi} h(x)f(x)dx$$

где χ означава скуп где случајна промењлива X узима вредности, која је обично једнако подржана расподелом f. Принцип методе Монте Карло за апроксимирање претходне једнакости је генерисање узорка $(X_1, ..., X_n)$ за расподелу f и апроксимативан емпиријски просек је

$$\overline{h}_n = \frac{1}{n}\sum_{j=1}^{n} h(x_j)$$

израчунавајући средњу вредност $mean(h(x))$ у R, пошто $\overline{h}_n$ конвергира скоро сигурно ка $E_f[h(X)]$ по строгом закону великих бројева. Још прецизније, када $h^2(X)$ има коначно очекивање за f, брзина конвергенције $\overline{h}_n$ може бити оцењено као $O(\sqrt{n})$ и асимптоцка апроксимација дисперзије је

$$var(\overline{h}_n) = \frac{1}{n}\int_{\chi}(h(x) - E_f[h(X)])^2 f(x)dx$$

Дисперзија може бити оцењене на узорку $(X_1, ..., X_n)$ са

$$v_n = \frac{1}{n^2} \sum_{j=1}^{n} [h(x_j - \overline{h}_n)]^2$$

Још битније, због централне граничне теореме, за велико n важи да је

$$\frac{\overline{h}_n - E_f[h(X)]}{\sqrt{v_n}}$$

апроксимативно распоређено као $N(0,1)$ расподела и ово води конвергенцији и граници поверења апроксимације $E_f[h(X)]$.
Методологија Монте Карла илустрована на примеру може бити успешно имплементирана у широком спектру предмета где расподеле у моделу могу бити симулиране. На пример, можемо користити Монте Карло суме за рачунање кумулативне функције расподеле.

Пример 4: За дату нормалну расподелу $N(0,1)$ узорак величине n, $(x_1, ..., x_n)$ је апроксимиран са

$$\Phi(t) = \int_{\propto}^{t} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy$$

па је по методи Монте Карло

$$\hat{\phi}(t) = \frac{1}{n} \sum_{i=1}^{n} I_{x_i \leq t}$$

са тачном дисперзијом $\phi(t)[1 - \phi(t)]/n$. Може бити да је уведена методологија Монте Карла довољна за апроксимирање интеграла. Међутим, Монте Карло метод обезбеђује добру апроксимацију у најрегуларнијим случајевима. Постоји још ефикаснија алтернатива која не избегава директну симулацију функције f али такође може бити искоришћена више пута за неколико интеграла. Поновна употреба може бити за фамилију фреквенције h или за фамилију густине f. Проблеми репа расподеле могу бити обрађени ефикасније него симулирање функције f јер симулација догађаја са веома малом вероватноћом захтева велики број симулација над f да би се постигла прецизност.

6. Монте Карло израчунавање интеграла

Већ поменути метод средње вредности, који се користи за оцењивање једноструких интеграла, је метод Монте Карло интеграције.

$$F_n = (b-a)(f) = (b-a)\frac{1}{n}\sum_{i=1}^{n} f(x_i)$$

Случајни бројеви x_i су распоређени униформно $X : U[a, b]$ на интервалу $a \leq x \leq b$ и n би био број покушаја.

1. $\int_0^\pi \sin(x)dx = -\cos(x)|_0^\pi = -\cos\pi - (-\cos 0) = 1 + 1 = 2$

   $(f) = \frac{1}{n}\sum_{i=1}^{n}\sin(x_i)$

   Промењлива X има униформну расподелу $X : U[a, b]$. Функција густине расподеле је $g(x) = \frac{1}{b-a} = \frac{1}{\pi}$.

   Математичко очекивање: $E(\sin x) = \int_0^\pi \sin(x)g(x)dx = \int_0^\pi \sin(x)\frac{1}{b-a}dx = \frac{1}{\pi}\int_0^\pi \sin(x)dx$

   $I = \int_0^\pi \sin(x)dx \rightarrow I = \pi E(\sin x) = \underline{F_n}$

   Да би се реализовала случајна промењлива X са униформном расподелом $U[0, 1]$ користи се R-функција $runif(x < -runif(n, a, b)\#n = 10000, a = 0, b = \pi\ )$. Као њен резултат добија се неки низ цифара.

   Код у R-у:

   $> x < -round(runif(10000, 0, \pi), 5)$ /**бирамо униформно неки низ цифара на* $[0, \pi]$
   $> m < -mean(sin(x))$ /**рачунамо средњу вредност*
   $> m$
   [1] 0.6389945
   $> pi * m$ /**рачунамо интеграл*
   [1] 2.00746

   Упоређујући вредност F_n са тачном вредности резултата интеегрaла налази се да је грешка која се односи на $n = 10^4$ покушаја апроксимативно **0.007**46.

2. $\int_0^2 e^x dx = 6.389056$

$> x < -round(runif(10000, 0, 2), 5)$
$> m < -mean(exp(x))$
$> m$
$[1]3.205835$
$> 2 * m$
$[1]6.411671$

Упоређујући вредност F_n са тачном вредности резултата интеграла налази се да је грешка која се односи на $n = 10^4$ покушаја апроксимативно **0.02**262.

3. $\int_1^4 \sqrt{x} dx = 4.66667$

$> a < -function(x)\{sqrt(x)\}$
$> x < -round(runif(10000, 1, 4), 5)$
$> m < -mean(a(x))$
$> m$
[1] 1.552631
$> 3 * m$
[1] 4.657892

Упоређујући вредност F_n са тачном вредности резултата интеграла налази се да је грешка која се односи на $n = 10^4$ покушаја апроксимативно **0.0087**8.

Налази се да је најбоље оцењен интеграл $\int_0^2 e^x dx$ овом методом.

Метод Монте Карло је један од начина како се одређени интеграли могу израчунати. Приликом коришћења овог метода може се повећати број покушајај како би оцена била ефикаснија.

7. Расподела вероватноће која није униформна

У претходним секцијама смо видели да се униформна расподела случајних бројева може користити за оцењивање одређених интеграла. Пожељно је наћи интегранд $f(x)$ у околини x где понашање брзо варира. Због значаја бирања узорка, методе захтевају расподелу вероватноће која није униформна, па прво разматрамо неколико метода за генерисање случајних бројева који нису распоређени униформно пре разматрања значења бираног узорка. Означићемо са r низ униформно распоређених бројева на интервалу $[0,1]$ тако да је $0 \leq r \leq 1$.
Претпоставимо да се два дискретна догађаја дешавају са вероватноћама p_1 и p_2 тако да је $p_1 + p_2 = 1$. Како можемо одабрати два догађаја са тачним вероватноћама користећи униформну вероватноћу расподеле? Очигледно је да бирамо догађај 1 ако је $r < p_1$; иначе, бирамо догађај 2. Ако постоје три догађаја са вероватноћама p_1,p_2 и p_3, онда ако је $r < p_1$ бирамо догађај 1; ако је $r < p_1+p_2$, бирамо догађај 2; иначе бирамо догађај 3. Случајно одабрана тачка r на линије сегмента ће се наћи у i-том сегменту са вероватноћом једнаком p_i.
Посматрајмо n дискретних догађаја. Како одредити сваки догађај i, да би одабрали вредност за дато r? За $n = 2$ и $n = 3$ вредност i која задовољава услов је

$$\sum_{j=0}^{i-1} p_j \leq r \leq \sum_{j=0}^{i} p_j \tag{32}$$

где смо дефинисали $p_0 \equiv 0$.
Посматрајмо непрекидну расподелу вероватноће која није униформна. Један начин да се генерише таква расподела је да се нађу границе у једнакости (32) као и вероватноће p_i са $p(x)dx$. Густина вероватноће $p(x)$ је дефинисана са $p(x)dx$ тако да је догађај x између x и $x + dx$. Густина вероватноће $p(x)$ је нормализована тако да је

$$\int_{-\propto}^{+\propto} p(x)dx = 1 \tag{33}$$

У једнакости (32) суме постају већ поменути интеграл и неједнакости постају једнакости. Можемо записати да је

$$P(x) = \int_{-\propto}^{+\propto} p(x')dx' = r' \tag{34}$$

Видимо да униформно рапоређен број r одговара **кумулативној вероватноћи расподеле функције** $P(x)$, чија је вроватноћа одабира вредности мања или једнака x. Функцију $P(x)$ не би требало поистовећивати са густином расподеле $p(x)$ или са вероватноћом $p(x)dx$. У многим примерима смислени опсег вредности за x је позитиван. У том случају, важи да је $p(x) = 0$ за $x < 0$.
Релација (33) нас води ка методу **инверзне трансформације** за генерисање случајних бројева распоређених у складу са функцијом $p(x)$. Овај метод се односи на генерисање случајног броја r и решавања једнакости (34) за одговарајућу вредност x. Као пример за ову методу, користићемо рерлацију (34) за генерисање низа случајних бројева у складу са униформном расподелом вероватноће на интервалу $a \leq x \leq b$. Жељена густина расподеле $p(x)$ је

$$p(x) = \begin{cases} 1/(b-a) & a \leq x \leq b \\ 0 & \text{иначе} \end{cases}$$

Кумулативна вероватноћа расподеле функције $P(x)$ за $a \leq x \leq b$ може бити добијена са

$$P(x) = \tfrac{x-a}{b-a}$$

Ако заменимо $P(x)$ у (30) и решимо по x, наћићемо жељену релацију

$$x = a + (b-a)r$$

Примењујемо метод инверзне трансформације густине расподеле функције

$$p(x) = \begin{cases} (1-\lambda)e^{-x/\lambda} & \text{ако је } 0 \leq x \leq\propto \\ 0 & x < 0 \end{cases} \tag{35}$$

Даљим извођењем следи

$$r = P(x) = 1 - e^{-x/\lambda}$$

Решење r за дато x даје $x = -\lambda \ln(1-r)$. Пошто је $1-r$ распоређено на исти начин као r, записаћемо

$$x = -\lambda \ln x$$

Променљива x нађена у претходној једнакости је распоређена у складу са густином вероватноће $p(x)$ датом у релацији (35).

Многа израчунавања природног логаритма су спора па инверзна трансформација можда није најефикаснији метод за употребу.
Из претходног примера, видимо да два услова морају бити задовољена у циљу примене метода инверзне трансформације. Посебно, густина функције вероватноће $p(x)$ мора омогућити аналитичко интеграљење у једнакости (35) и да се инвертује релација $P(x) = r$ за x. Гаусова густина вероватноће

$$p(x) = \frac{1}{(2\pi\sigma^2)^{1/2}} e^{-x^2/2\sigma^2} \tag{36}$$

је пример густине вероватноће чија кумулативна расподела $P(x)$ не може бити добијена аналитично. Међутим, можемо генерисати дводимензионалну вероватноћу $p(x,y)dxdy$ дату као

$$p(x,y)dxdy = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} dxdy \tag{37}$$

Прво, правимо измену међу променљивим уводећи поларне координате:

$$r = (x^2+y^2)^{1/2} \qquad \theta = \tan^{-1}\frac{y}{x}$$

Нека је $\rho = r^2/2$ па следи да је дводимензионална вероватноћа

$$p(\rho,\theta)d\rho d\theta = \frac{1}{2\pi} e^{-\rho} d\rho d\theta$$

где имамо да је $\sigma = 1$. Ако генеришемо ρ у складу са експоненцијалном расподелом и θ униформно на интервалу $0 \leq \theta < 2\pi$, онда ће променљиве $(Box - Muller\ method)$

$$x = (2\rho)^{1/2}\cos\theta \quad \text{и} \quad y = (2\rho)^{1/2}\sin\theta$$

бити генерисане у складу са једнакости (37) са средњом вредношћу 0 и $\sigma = 1$. Метод који се користи за генерисање Гаусове расподеле је познат као $Box - Muller$ метод.

5. Упоређивање детерминистичког и стохастичког приступа

1. Неки примери

Пример 5: Израчунати следеће интеграле користећи се: **нумеричком интеграцијом** (на пет децимала), **класичним поступком** израчунавања интеграла и **методом Монте Карло** па анализом грешке упоредити интеграле (грешка: 0.5×10^{-5}).

1.$\int_0^{\pi} \sin(x)dx$

2. $\int_0^2 e^x dx$

3. $\int_1^4 \sqrt{x}dx$

Израчунавање интеграла **Трапезном формулом** и **Симпсоновом формулом**:

1. $\int_0^{\pi} sin(x)dx = -cos(x)|_0^{\pi} = -cos\pi - (-cos0) = 1 + 1 =$**2**

$[a, b] = [0, \pi]$ $\quad n = 10$ $\quad h = \frac{b-a}{n} = \frac{\pi}{10} = \frac{3.141593}{10} = 0.3141593 \approx 0.3$ $\quad x_i = x_0 + ih$

i	x_i	y_i
1	0.3	0.29552
2	0.6	0.56464
3	0.9	0.78333
4	1.2	0.93204
5	1.5	0.99750
6	1.8	0.97385
7	2.1	0.86321
8	2.4	0.67546
9	2.7	0.42738

Трапезна формула: $I_T = \frac{h}{2}(y_0 + y_n + 2\sum_{i=1}^{n-1} y_i)$ $\qquad y_0 = 0$ и $y_n = 0.14112$

$I_T = \frac{0.3}{2}(0 + 0.14112 + 2 \times 6.51293) = \frac{0.3}{2}(0.14112 + 13.02586) = 0.15 \times 13.16698 = \underline{1.975047}$

$E_T = |I - I_T| = |2 - 1.97505| = \mathbf{0.02}495$

Процена грешке Трапезне формуле је дала оцену на 2 децимале: 2.5×10^{-2}

Симпсонова формула: $I_S = \frac{h}{3}(y_0 + y_n + 4\sum_{i=1}^{n-1} y_i + 2\sum_{i=2}^{n-2} y_i)$ $y_0 = 0$ и $y_n = 0.14112$

$I_S = \frac{0.3}{3}(0.14112 + 4 \times 3.36694 + 2 \times 3.14599) = 0.1(0.14112 + 13.46776 + 6.29198) =$
$0.1 \times 19.90086 = \underline{1.99009}$

$E_S = |I - I_S| = |2 - 1.99009| = \mathbf{0.009}91 \approx 0.01101$

Процена грешке Симпсонове формуле је дала оцену на 2 децимале: 1.1×10^{-2}

2. $\int_0^2 e^x dx = e^x|_0^2 = e^2 - e^0 = 7.389056 - 1 = \mathbf{6.38906}$

$[a, b] = [0, 2] \qquad n = 10 \qquad h = \frac{b-a}{n} = \frac{2}{10} = 0.2 \qquad x_i = x_0 + ih$

i	x_i	y_i
1	0.2	1.22140
2	0.4	1.49182
3	0.6	1.82212
4	0.8	2.22554
5	1	2.71828
6	1.2	3.32012
7	1.4	4.05520
8	1.6	4.95303
9	1.8	6.04965

Трапезна формула: $I_T = \frac{h}{2}(y_0 + y_n + 2\sum_{i=1}^{n-1} y_i)$ $y_0 = 1$ и $y_n = 7.38906$

$I_T = \frac{0.2}{2}(1 + 7.38906 + 2 \times 27.85716) = 0.1(8.38906 + 55.71432) = 0.1 \times 64.10338 = \underline{6.41034}$

$E_T = |I - I_T| = |6.38906 - 6.41034| = \mathbf{0.02}128$

Процена грешке Трапезне формуле је дала оцену на 2 децимале: 2.13×10^{-2}

Симпсонова формула: $I_S = \frac{h}{3}(y_0 + y_n + 4\sum_{i=1}^{n-1} y_i + 2\sum_{i=2}^{n-2} y_i)$ $y_0 = 1$ и $y_n = 7.38906$

$I_S = \frac{0.2}{3}(1 + 7.38906 + 4 \times 15.86665 + 2 \times 11.99051) = 0.0667(8.38906 + 63.4666 + 23.98102) =$
$0.0667 \times 95.83668 = \underline{6.39231}$

$E_S = |I - I_S| = |6.38906 - 6.39231| = \mathbf{0.003}25$

Процена грешке Симпсонове формуле је дала оцену на 3 децимале: 3.25×10^{-3}

3. $\int_1^4 \sqrt{x}dx = \int_1^4 x^{\frac{1}{2}}dx = \frac{x^{\frac{1}{2}+1}}{\frac{1}{2}+1}\Big|_1^4 = \frac{2}{3}x^{\frac{3}{2}}\Big|_1^4 = \frac{16}{3} - \frac{2}{3} = \frac{14}{3} = \mathbf{4.66667}$

$[a, b] = [1, 4]$ $n = 10$ $h = \frac{b-a}{n} = \frac{3}{10} = 0.3$ $x_i = x_0 + ih$

i	x_i	y_i
1	1.3	1.14018
2	1.6	1.26491
3	1.9	1.37840
4	2.2	1.48324
5	2.5	1.58114
6	2.8	1.67332
7	3.1	1.76068
8	3.4	1.84391
9	3.7	1.92354

Трапезна формула: $I_T = \frac{h}{2}(y_0 + y_n + 2\sum_{i=1}^{n-1} y_i)$ $\quad y_0 = 1$ и $y_n = 2$

$I_T = \frac{0.3}{2}(1 + 2 + 2 \times 14.04932) = \frac{0.3}{2} \times (3 + 28.09864) = 0.15 \times 31.09864 = \underline{4.66480}$

$E_T = |I - I_{10}| = |4.66667 - 4.66480| =$**0.001**87

Процена грешке Трапезне формуле је дала оцену на 3 децимале: 1.87×10^{-3}

Симпсонова формула: $I_S = \frac{h}{3}(y_0 + y_n + 4\sum_{i=1}^{n-1} y_i + 2\sum_{i=2}^{n-2} y_i)$ $y_0 = 1$ и $y_n = 2$

$I_S = \frac{0.3}{3}(1+2+4\times7.78394+2\times6.26538) = 0.1(3+31.13576+12.53076) = 0.1\times46.6665 = \underline{4.66665}$

$E_S = |I - I_S| = |4.66667 - 4.66665| =$**0.00002**

Процена грешке Симпсонове формуле је дала оцену на свих 5 децимале: 0.2×10^{-5}

Методом интеграције, користећи се Трапезном и Симпсоновом формулом, добија се најбоље оцена за интеграл $\int_1^4 \sqrt{x}dx$ док је методом Монте Карло најбоље оцењени интеграл $\int_0^\pi sinxdx$.

интеграл	**Трапезна ф.**	**Симсонова ф.**	**МК метод**
$\int_0^\pi sinx = 2$	1.975047	1.99009	2.00746
$\int_0^2 e^x dx = 6.389056$	6.41034	6.39231	6.411671
$\int_1^4 \sqrt{x}dx = 4.66667$	4.66480	4.66665	4.657892

2. Нумеричка интеграција вишеструких интеграла

Ако почнемо од тога да се грешка смањује као $n^{-\alpha}$ за $d = 1$, онда грешка опада као $n^{-\alpha/d}$ у d димензија. У супротном, грешка која се односи на све методе Монте Карло интеграције опада као $n^{-\frac{1}{2}}$ независно од димензије интеграла. Пошто је време израчунавања грешке пропорционално са n код класичне методе Монте Карло, закључујемо да су за ниже димензије, класичне нумеричке методе учесталије као што је Симпсоново правило. Међутим, грешка у конвенционалним нумеричким методама расте са димензијом и методе Монте Карло су суштински за вишедимензионалне интеграле.
Да би се илустровао општи метод оцењивања вишедимензионалног интеграла,посматрајмо дводимензионални интеграл

$$F = \int_R f(x, y)dxdy$$

где R означава област интеграције. Формирати правоугаоник који затвара регион R и дели га на квадрате дужине h. Претпоставимо да се правоугаоник протеже од x_a ка x_b у правцу x осе и од y_a до y_b у смеру y-осе. Укупан број квадрата је $n_x n_y$, где је $n_x = (x_b - x_a)/h$ и $n_y = (y_b - y_a)/h$. Ако користимо средњу тачку апроксимације, интеграл F је оцењен као

$$F \approx \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} f(x_i, y_j) H(x_i, y_j) h^2$$

где $x_i = x_\alpha + (i - \frac{1}{2})h$, $y_j = y_\alpha + (j - \frac{1}{2})h$ и функција $H(x, y)$ једнака јединици ако $(x, y) \in R$ или нула иначе.
Једноставна метода Монте Карло за оцењивање дводимензионалног интеграла користи исту правоугаону област као и изнад али n тачака (x_i, y_i) је одабрано случајно на правоугаонику. Оцена интеграла је онда

$$F_n = \frac{A}{n} \sum_{i=1}^{n} f(x_i, y_i) H(x_i, y_i)$$

где је A област правоугаоника. Забележимо да ако је $f(x, y) = 1$ свуда, онда је једнакост изнад еквивалентна методи погодак или промашај за израчунавање области R.

3. Вишеструки интеграли

Монте Карло метод добија на већем значају када посматрамо интеграцију у већим димензијама. На пример,

$$I = \int_0^1 \int_0^1 \int_0^1 f(x, y, z) dx dy dz \approx \frac{1}{N} \sum_{i=1}^{N} f(x_i, y_i, z_i)$$

где је (x_i, y_i, z_i) случајан низ тачака за $x, y, z \in [0, 1]$. Укупно, потребно нам је $3N$ случајних бројева да би се генерисало N случајних тачака. Интеграл над произвољном области која се интеграли је

$$\begin{aligned} I &= \int_{a_1}^{a_2} \int_{b_1}^{b_2} \int_{c_1}^{c_2} f(x, y, z) dx dy dz \\ &= \int_{a_1}^{a_2} [\int_{b_1}^{b_2} (\int_{c_1}^{c_2} f(x, y, z) dx)] dz \\ &\approx (a_2 - a_1)(b_2 - b_1)(c_2 - c_1) \frac{1}{N} \sum_{i=1}^{N} f(x_i, y_i, z_i) \end{aligned}$$

У општем случају, метод средње вредности је дат са

$$I = \int_A f \approx$$

$\approx$ (површина A) $\times$ (просечна вредност функције f над n случајних тачака на A)

Грешка код стандардне нумеричке интеграције опада као $n^{-a}, a \in N$ за $d = 1$ (једнодимензионални интеграл); на пример код трапезног правила је n^{-2} . У d димензија, грешка опада као $n^{-a/d}$. Овде n означава укупан број подинтервала; број подинтервала по димензији је $n^{1/d}$.
Грешка Монте Карло интеграције увек опада као $N^{-1/2}$.
Трошак израчунавања је релативан у односу на укупан број подинтервала n или укупан број тачака N. Из овога можемо закључити да је Монте Карло интеграција најефикаснија за $d > 2a$.

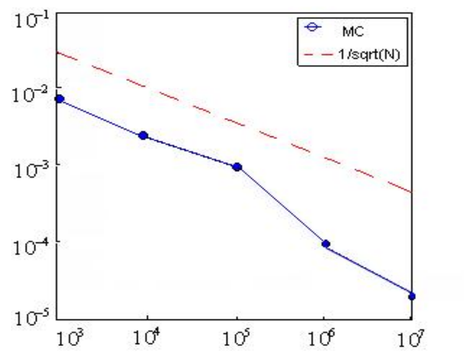
Пример 6: Двоструки интеграл

Посматрајмо интеграл функције $f(x,y) = xye^{-x^2y}$ на квадрату $x, y \in [0,1]$. Користећи Монте Карло метод средње вредности (Monte Carlo sample mean - MC), вредност овог интеграла је оцењена генерисањем N тачака случајно на датом квадрату и рачунањем просечних вредности функције $f(x,y)$.
Тада је

$$I_{MC} = \frac{1}{N}\sum_{i=1}^{N} f(x_i, y_i)$$

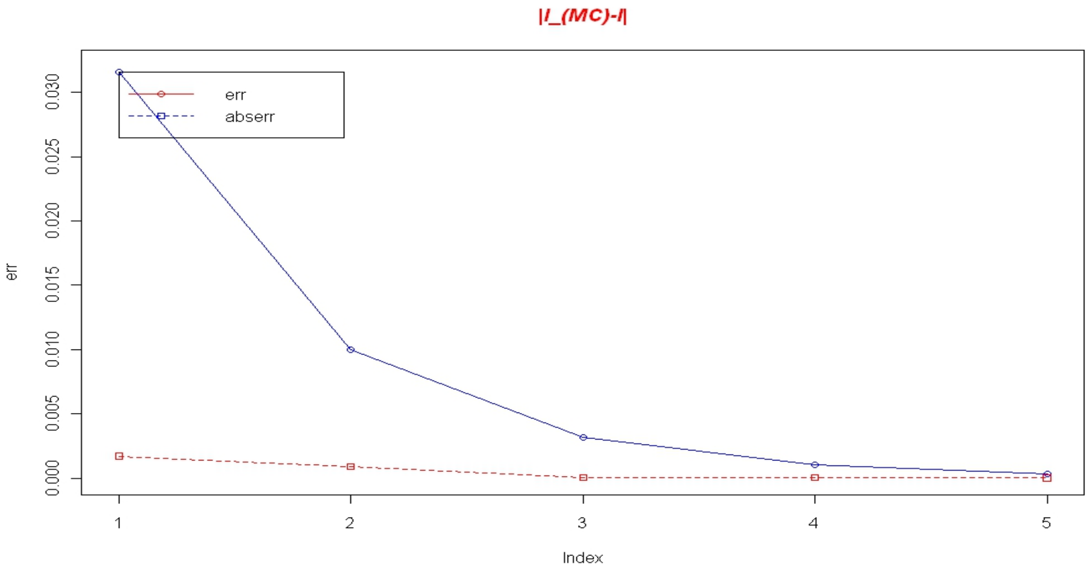
Овде је фактор области $A = 1$.
Тачна вредност интеграла је

$$I = \int_0^1 \int_0^1 xye^{-x^2y}dxdy = \frac{1}{2e} \approx 0.18393972$$



Слика 1.5: Апсолутна грешка Монте Карло оцене.

Грешка се смањује као $1/\sqrt{N}$ али постоји непоузданост која се односи на сваку оцену (линија која није права). Ово је последица тога што користимо случајне бројеве за израчунавање оцена. Ако користимо низ од N случајних бројева и поновимо рачунање са другим низом од N случајних бројева, добићемо два различита одговора. Расподела одговора ће постати ужа како повећавамо број N и низ ће конвергирати тачном одговору.
Грешка код Монте Карло израчунавања је различита него код других метода интеграције. На пример, посматрајмо трапезно правило.
У овом случају, грешка је последица линеарне апроксимације која се јавља на сваком подинтервалу. Делећи подинтервале што мање, склоп је бољи и грешка опада.



Слика 1.6: Апсолутна грешка Монте Карло оцене.

Код у R-у:

$> i < -0.18393972$
$> i\_mc < -c(0.1822299, 0.1830448, 0.1839015, 0.1839026, 0.1839296)$
$> abserr < -c(0.00170982, 0.00089492, 3.822e-05, 3.712e-05, 1.012e-05)$
$> err < -c(0.031622777, 0.01, 0.003162278, 0.001, 0.000316228)$
$> g\_range < -range(0, err, abserr)$
$> plot(err, type = "o", col = "blue", ylim = c(0, 0.032))$
$> lines(abserr, type = "o", pch = 22, lty = 2, col = "red")$
$> title(main = "|I\_(MC) - I|", col.main = "red", font.main = 4)$
$> legend(1, g\_range[2], c("err", "abserr"), col = c("red", "blue"), pch = 21 : 22, lty = 1 : 2)$

Видимо на графику да је грешка приликом првог израчунавања интеграла највећа и да се повећавањем броја покушаја она смањује. За $n = 10^7$ добија се најмања апсолутна грешка Монте Карло оцене.

Пример 7. Израчунати интеграл $I = \int_{0.2}^{0.5} x^3(1-x)^{2.5}dx$ који има расподелу $X : U[0.2, 0.5]$.

$$I = \int_{0.2}^{0.5} x^3(1-x)^{2.5}dx = 0.00428$$

$> x < -runif(100, 0.2, 0.5)$
$> g < -function(x)\{$
$+x^3 * (1-x)^{2.5}\}$
$> mean(g(x))$
$[1]0.01432740$
$> mean(g(x)) * 0.3$
$[1]0.004298220$

Пример 8: Израчунати интеграл $I = \int_1^7 x^{1.5}e^{-2x}dx$ који има расподелу $X : U[1, 7]$.

$$I = \int_1^7 x^{1.5}e^{-2x}dx = 0.12685$$

$> x < -runif(100, 1, 7)$
$> g < -function(x)\{$
$+x^{1.5} * exp(-2 * x)\}$
$> mean(g(x))$
$[1]0.02114137$
$> mean(g(x)) * 6$
$[1]0.1268482$

Закључак

Детерминистички и стохастички приступ представљају начине израчунавања одређених интеграла поред елементарног начина. Кроз рад се може приметити да је нумеричка интеграција један начин израчунавања одређених интегралал који се види кроз примену Трапезне и Симпсонове формуле. Метод Монте Карло је други начин и при том израчунавању је бржи. Овај метод користи једноставне алгоритме за израчунавање одређених интеграла како једнодимензионалних тако и вишедимензионалних. Проучавајући ову тему може се закључити да су оба приступа корисна. Стохастички приступ је ефикаснији и најчешће се примењује код вишеструких интеграла. Примери одређених интеграла код методе нумеричке интеграције као и примери одређених интеграла код методе Монте Карло су упоређивани у овом раду да би се видела ефикасност елементарног, детерминистичког и стохастичког приступа израчунавања. Суштина рада је у томе да ако се одређени интеграли не могу израчунати елементарно онда се могу израчунати једним од ова два приступа.

Литература

[1] В. Јевремовић, Ј. Малишић, Статистичке методе у метеорологији и инжењерству, Савезни хидрометеоролошки завод, Београд, 2002.

[2] Весна Јевремовић, Јован Малишић, Драган Ђорић, Емилија Николић - Ђорић, Атлас расподела, Грађевински факултет, Београд, 2007.

[3] З. Каделбург, Д. Аднађевић, Анализа 1 - *Одређени интеграли*, Математички факултет, 2008.

[4] Зоран Станић, Нумеричка интеграција - *poincare.matf.bg.ac.rs zstanic/*

[5] Р. Оморјан, Нумеричка интеграција, 2006. - *www.tf.uns.ac.rs*

[6] $Krauth, W.(1998) Introduction To Monte Carlo Algorithms, CNRS - Laboratoire de Physique Statistique$

[7] $Gould, H and Tobochnik, J. and Christian, W.(2001) Numerical Integration And Monte Carlo Methods$

[8] $www.Ice.hut.fi/teaching/S-114.1100/lect\_9.pdf$