

Metody Monte Carlo

Kurs dla kierunków Informatyka oraz Fizyka

Wydział Fizyki, Astronomii i Informatyki Stosowanej
Uniwersytet Jagielloński
ul. Reymonta 4, 30-059 Kraków

Rok akademicki 2010/2011

DR HAB. PROF. UJ WIESŁAW PŁACZEK

pokój 446, *e-mail:* Wieslaw.Placzek@uj.edu.pl

<http://th-www.if.uj.edu.pl/~placzek/dydaktyka/MMC/>

► **Forma zaliczenia:** zaliczenie ćwiczeń + egzamin ustny

1. R. Zieliński, „*Metody Monte Carlo*”, WNT 1970.
2. S. M. Jermakow, „*Metoda Monte Carlo i zagadnienia pokrewne*”, PWN 1976.
3. R. Zieliński, R. Wieczorkowski, „*Komputerowe generatory liczb losowych*”, WNT 1997.
4. R. Wit, „*Metody Monte Carlo – wykłady*”, Wydawnictwo Politechniki Częstochowskiej, 2004.
5. J. M. Hammersley, D. C. Handscomb, „*Monte Carlo Methods*”, London: Methuen & Co. Ltd., New York: J. Wiley & Sons Inc., 1964.
6. I. M. Sobol, „*The Monte Carlo Method*”, Mir Publishers, Moscow 1975.
7. F. James, „*Monte Carlo theory and practice*”, Rep. Prog. Phys. **43** (1981) 1145.
8. M. H. Kalos, P. A. Whitlock, „*Monte Carlo Methods*”, J. Wiley & Sons Inc., New York, 1986.
9. G. S. Fishman, „*Monte Carlo: Concepts, Algorithms and Applications*”, Springer, 1996.
10. R. Y. Rubinstein, D. P. Kroese, „*Simulation and the Monte Carlo Method*”, Second Edition, J. Wiley & Sons Inc., 2008.
11. R. Korn, E. Korn, G. Kroisandt, „*Monte Carlo methods and models in finance and insurance*”, CRC Press, Taylor & Francis Group, 2010.
12. S. Jadach, „*Practical Guide to Monte Carlo*”, arXiv: physics/9906056, <http://cern.ch/jadach/MCguide/>.
13. <http://random.mat.sbg.ac.at/>.

- **Definicje, historia, stosowalność metod Monte Carlo.**
- **Matematyczne podstawy metod Monte Carlo.**
- **Całkowanie metodami Monte Carlo oraz techniki redukcji wariancji.**
- **Adaptacyjne techniki całkowania Monte Carlo.**
- **Generatory liczb losowych o rozkładzie równomiernym i ich testowanie.**
- **Metody generowania liczb losowych o dowolnych rozkładach prawdopodobieństwa.**
- **Generatory liczb losowych dla podstawowych rozkładów prawdopodobieństwa.**
- **Zastosowanie metod błędzenia przypadkowego do rozwiązywania układów równań liniowych, równań różniczkowych, równań całkowych, problemów macierzowych, interpolacji funkcji wielu zmiennych itd.**
- **Zastosowanie metod Monte Carlo w zagadnieniach optymalizacji.**
- **Modelowanie (symulacje) Monte Carlo, m.in. zastosowanie w teorii gier.**

⇒ **Definicja ogólna:**

Technika Monte Carlo (MC) jest to dowolna technika używająca liczb losowych do rozwiązania problemu.

▶ **Liczba losowa** – na razie zakładamy, iż wiemy co to jest (choć jest to rzecz nietrywialna!)

⇒ **Definicja Haltona (1970)** – nieco węższa, ale bardziej pouczająca:

Metoda Monte Carlo jest to metoda reprezentująca rozwiązanie problemu w postaci parametru pewnej hipotetycznej populacji i używająca sekwencji liczb losowych do skonstruowania próby losowej danej populacji, z której to statystyczne oszacowania tego parametru mogą być otrzymane.

Niech F – wynik rozwiązania jakiegoś problemu (liczba rzeczywista, zbiór liczb, decyzja tak/nie itd.)

⇒ Oszacowanie Monte Carlo wyniku F :

$$\hat{F} = f(\{r_1, r_2, \dots, r_n\}; \dots),$$

gdzie $\{r_1, r_2, \dots, r_n\}$ – liczby losowe użyte w rachunku.

Problem nie musi być natury stochastycznej!

▷ Wprowadzenie losowości do danego, dobrze określonego problemu daje rozwiązania o specyficznych własnościach, które – jak się przekonamy później – są zaskakująco dobre!

- **G. Comte de Buffon (1777)** – prawdopodobnie najwcześniejsze udokumentowane użycie próbkowania losowego do obliczenia całki (przez rzucanie igły na poziomą płaszczyznę pokrytą równoległymi liniami prostymi).
- **Marquis Pierre-Simon de Laplace (1886)** – zastosowanie metody Buffona do wyznaczania wartości liczby π (**wielkość zupełnie nielosowa!**).
- **Lord Kelvin (1901)** – użycie próbkowania losowego (wyciąganie ponumerowanych kartek papieru z urny) do obliczenia pewnych całek w kinetycznej teorii gazów.
- **W. S. Gosset, ps. „Student”, (1908)** – podobne losowanie pomogło mu w odkryciu rozkładu współczynnika korelacji oraz potwierdzeniu rozkładu t -Studenta.
- **Enrico Fermi (lata 1930-te)** – eksperymenty losowania numerycznego dotyczące dyfuzji i transportu neutronów w reaktorach jądrowych (skonstruował specjalne mechaniczne urządzenie losujące, nazwane FERMIAK).
- **J. von Neumann, S. Ulam, N. Metropolis, R. Feynman i in. (lata 1940-te)** – pierwsze na dużą skalę rachunki oparte o użycie liczb losowych; dotyczyły rozpraszania i absorpcji neutronów w ramach projektu „Manhattan” (prace nad bombą jądrową w Los Alamos, USA). → Nazwa „Monte Carlo” została wymyślona jako kryptonim dla tego typu rachunków i odpowiednich metod matematycznych.

- **To czy metody Monte Carlo mogą być zastosowane do danego problemu nie zależy od stochastycznej natury układu, który jest badany, a jedynie od naszej zdolności do sformułowania problemu w taki sposób, aby liczby losowe mogły być użyte do jego rozwiązania.**
 - ▷ Rachunki takie jak te prowadzone w latach 1940-tych w Los Alamos nazywa się **bezpośrednią symulacją**, ponieważ „hipotetyczna populacja” (wg. definicji Haltona) odpowiada rzeczywistej badanej populacji. Jednakże wyniki numeryczne jakie tam uzyskano były całkowicie „**deterministyczne**” – w zasadzie możliwe od uzyskanie przy użyciu klasycznych technik obliczeniowych (całkowanie wielowymiarowe).
 - ▶ **Metody Monte Carlo mogą być zastosowane wszędzie, gdziekolwiek możliwe jest określenie równoważności pomiędzy żądanym rezultatem a oczekiwanym zachowaniem pewnego układu stochastycznego.**
 - ▷ Problem do rozwiązania może być natury **probabilistycznej** lub **statystycznej**, w którym to przypadku jego rozwiązanie MC będzie **bezpośrednią symulacją**, ale może też być natury **deterministycznej** lub **analitycznej**, w którym to przypadku odpowiednie sformułowanie Monte Carlo może wymagać pewnej **wyobraźni**, a czasem nawet wydawać się sztuczne i nienaturalne.
 - ⇒ **Stosowalność** wybranej metody będzie **zależać** od jej **matematycznych własności**, a nie od jej **powierzchnowego podobieństwa** do rozwiązywanego problemu.

Generujemy liczby losowe z rozkładu równomiernego na przedziale $(0, 1)$ do momentu, aż kolejna liczba będzie mniejsza od poprzedniej, tzn. otrzymujemy ciąg:

$$x_1 < \dots < x_{l-1} > x_l, \quad l \geq 2.$$

Zapamiętujemy wartość l . Powtarzamy tę czynność N razy, a następnie obliczamy średnią arytmetyczną wartości l . Otrzymana liczba jest oszacowaniem Monte Carlo stałej matematycznej e (podstawa logarytmu naturalnego, zwana też liczbą Eulera), tzn.

$$\hat{e} \equiv \frac{1}{N} \sum_{n=1}^N l_n \xrightarrow{N \rightarrow \infty} e.$$

⇒ Ćw. A1* (analityczne nadobowiązkowe): Udowodnić powyższe.

Przykładowe wyniki obliczeń Monte Carlo:

N	\hat{e}	$\hat{e} - e$
1	2.000000	-0.718282
100	2.750000	0.031718
10 000	2.722900	0.004618
1 000 000	2.717491	-0.000791
100 000 000	2.718239	-0.000042

⇒ Ćw. N1.1 (numeryczne obowiązkowe): Wykonać obliczenia Monte Carlo jak powyżej.

- **Wszystkie rachunki Monte Carlo są równoważne całkowaniu (w sensie formalnym)!**

▶ Niech F – wynik rachunku MC $\Rightarrow F$ – funkcja liczb losowych r_i .

▷ Założenie (dla prostoty): r_i – liczby losowe z rozkładu równomiernego na przedziale $(0, 1)$.

\Rightarrow Wynik MC:

$$F = F(r_1, r_2, \dots, r_n)$$

jest **estymatorem nieobciążonym** całki wielowymiarowej:

$$I = \int_0^1 \dots \int_0^1 F(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n$$

lub inaczej – **wartością oczekiwaną** wielkości F jest całka I :

$$E(F) = I.$$

Uwaga: Jeżeli rozwiązywany problem polega na obliczeniu całki funkcji f , to powyższe F nie powinno być identyfikowane z funkcją f , ale z oszacowaniem MC jej całki.

- ▶ Ta formalna **równoważność** da nam mocne **teoretyczne uzasadnienie** dla **technik Monte Carlo**, a także będzie prowadzić do wielu **rezultatów** o istotnym znaczeniu **praktycznym**.

Zmienne losowe i ich rozkłady

Zmienna losowa jest zmienną, która może przyjmować więcej niż jedną wartość (ogólnie przedział wartości z kontinuum) i dla której żadna z wartości, jakie przyjmuje nie może być przewidziana z góry.

▷ Chociaż **wartość zmiennej losowej jest nieprzewidywalna**, to jej **rozkład** może być znany!

Rozkład zmiennej losowej określa prawdopodobieństwo przyjęcia przez nią danej wartości (lub infinitezymalnego przedziału wartości).

▶ Dla **zmiennych ciągłych** definiujemy:

$$\rho(u) du = \mathcal{P} [u < u' < u + du] ,$$

$\rho(u)$ – **funkcja gęstości prawdopodobieństwa zmiennej u** (daje prawdopodobieństwo znalezienia zmiennej losowej u' wewnątrz przedziału du dla danej wartości u).

▶ **Scałkowana funkcja rozkładu – dystrybuanta:**

$$R(u) = \int_{-\infty}^u \rho(x) dx , \quad \rho(u) = \frac{dR(u)}{du} .$$

Uwaga: $R(u)$ jest **monotonicznie niemalejącą funkcją** i $0 \leq R(u) \leq 1$, a ρ jest zawsze tak unormowana, aby całka po całym zakresie zmiennej u była równa 1.

Wartość oczekiwana, wariancja i odchylenie standardowe

- Matematycznie **wartość oczekiwana** funkcji $f(u')$ zdefiniowana jest jako średnia funkcji:

$$E(f) = \int f(u) dR(u) = \int f(u) \rho(u) du .$$

gdzie $R(u)$ – dystrybuanta zmiennej losowej u' , a $\rho(u)$ – jej gęstość prawdopodobieństwa.

▷ Jeżeli $u' \in \mathcal{U}(a, b)$, tzn. u' jest równomiernie rozłożona między a i b , to:

$$dR = \frac{du}{b-a} \quad \Rightarrow \quad E(f) = \frac{1}{b-a} \int_a^b f(u) du .$$

- **Wariancja funkcji** $f(u')$ – średnia kwadratu odchylenia od wartości oczekiwanej:

$$V(f) = E \left([f - E(f)]^2 \right) = \int [f - E(f)]^2 dR .$$

► **Odchylenie standardowe:**

$$\sigma(f) = \sqrt{V(f)} .$$

→ W praktyce częściej używane niż wariancja, bo ma ten sam wymiar co jego argument (ważne np. dla wielkości fizycznych).

Własności wartości oczekiwanej i wariancji

- Niech x i y – zmienne losowe, a c – stała, wówczas:

$$E(cx + y) = cE(x) + E(y),$$

$$V(cx + y) = c^2 V(x) + V(y) + 2c \cdot Cov(x, y),$$

gdzie: $Cov(x, y) = E([x - E(x)][y - E(y)])$ – **kowariancja** między x i y .

⇒ **Ćw. A1 (analityczne obowiązkowe):** Udowodnić powyższe własności.

$$Cov(x, y) \begin{cases} = 0, & x \text{ i } y \text{ są nieskorelowane,} \\ > 0, & x \text{ i } y \text{ są dodatnio skorelowane,} \\ < 0, & x \text{ i } y \text{ są ujemnie skorelowane.} \end{cases}$$

▶ Jeżeli x i y są **niezależne**, to: $Cov(x, y) = 0$ i $V(x + y) = V(x) + V(y)$.

⇒ **Ćw. A2:** Pokazać powyższe.

▷ Jeżeli zmienne są **niezależne**, to są **nieskorelowane!** Natomiast jeżeli są **nieskorelowane**, to **nie zawsze** muszą być **niezależne!** ⇒ **Ćw. A3:** Znaleźć przykład na to drugie.

- Wnioski:** $V(f) = E(f^2) - [E(f)]^2,$ $Cov(x, y) = E(xy) - E(x)E(y).$

⇒ **Ćw. A4:** Udowodnić powyższe własności.

Prawo wielkich liczb (PWL)

Wyberzmy losowo n liczb u_i z jednorodnym prawdopodobieństwem na przedziale (a, b) i dla każdego u_i obliczmy wartość funkcji $f(u_i)$.

- **PWL:**

$$\frac{1}{n} \sum_{i=1}^n f(u_i) \xrightarrow{n \rightarrow \infty} E(f) = \frac{1}{b-a} \int_a^b f(u) du.$$

- ▶ W języku statystycznym: lewa strona jest **estymatorem zgodnym** całki po prawej stronie, ponieważ (pod „pewnymi warunkami”) jest zbieżna do dokładnej wartości całki kiedy $n \rightarrow \infty$.
- ▷ „Pewne warunki” dotyczą zachowania funkcji f , która musi być:
 - * całkowna,
 - * przynajmniej kawałkami ciągła (może mieć skończoną liczbę nieciągłości),
 - * wszędzie skończona (powszechnie stawiane wymaganie).

Prawo wielkich liczb może być interpretowane jako stwierdzenie, że oszacowanie Monte Carlo całki zbiega się do poprawnej wartości, jeżeli rozmiar próby losowej staje się bardzo duży!

Zbieżność

► Problem zbieżności jest bardziej skomplikowany niż dla zwykłego rachunku!

- **Zwykły rachunek:**

Mówimy, że ciąg $\{a_n\}$ jest zbieżny do wartości a , jeżeli dla dowolnie małej liczby δ można znaleźć taki element ciągu $\{a_n\}$, że wszystkie następne elementy tego ciągu są **na pewno** w odległości mniejszej niż δ od wartości a .

- **Statystyka:**

Mówimy, że $a(n)$ jest zbieżne do a , jeżeli dla dowolnego prawdopodobieństwa $\mathcal{P}[0 < \mathcal{P} < 1]$ i dowolnej dodatniej liczby δ można znaleźć takie k , że dla wszystkich $n > k$ **prawdopodobieństwo** tego, że $a(n)$ będzie w odległości $< \delta$ od a jest **większe** niż \mathcal{P} .

„na pewno” \longrightarrow **prawdopodobieństwo**

▷ **Uwaga:** Choćby nie wiadomo jak duże było n , to nie **ma pełnej gwarancji**, że $a(n)$ będzie w danym przedziale wokół a !

► **Zbieżność jest dana jedynie z pewnym prawdopodobieństwem!**

→ Dlatego rachunki MC wiążą się zawsze z pewnym ryzykiem – stąd też nazwa „**Monte Carlo**” (od kasyn Monte Carlo w Księstwie Monako).

Centralne twierdzenie graniczne (CTG)

Suma dużej liczby niezależnych zmiennych losowych ma zawsze rozkład normalny (tzn. rozkład Gaussa), niezależnie od tego jakie rozkłady mają indywidualne zmienne losowe, pod warunkiem, że mają skończone wartości oczekiwane i wariancje, oraz że n jest „dostatecznie duże”.

- Jak duże musi być n ?

- ▷ Zależy to od indywidualnych rozkładów, ale w praktyce **zbieżność** do rozkładu normalnego jest zaskakująco **szybka** (nawet jeśli poszczególne rozkłady są istotnie różne od normalnego, np. rozkłady równomierne).

- ▶ Rozkład Gaussa (normalny) $N(\mu, \sigma^2)$ o wartości oczekiwanej μ i wariancji σ^2 :

$$\rho(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x - \mu)^2}{2\sigma^2}\right].$$

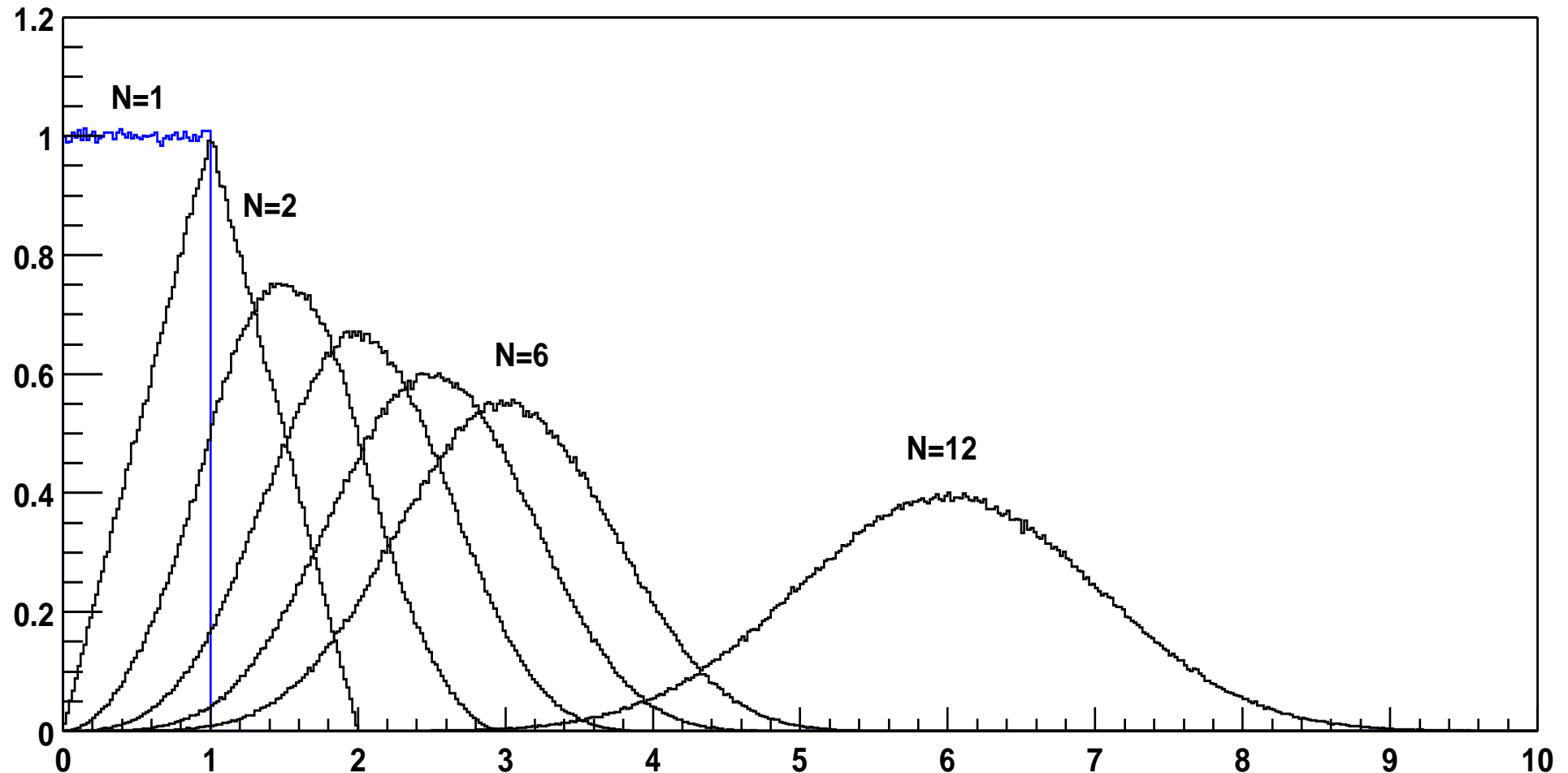
⇒ **Ćw. N1.2:** Narysować ten rozkład przy pomocy systemu ROOT dla kilku wartości (μ, σ) i pokazać:

$$\mathcal{P}\left(-1.64 \leq \frac{x - \mu}{\sigma} < 1.64\right) = 0.9,$$

$$\mathcal{P}\left(-2.58 \leq \frac{x - \mu}{\sigma} < 2.58\right) = 0.99,$$

$$\mathcal{P}\left(-3.29 \leq \frac{x - \mu}{\sigma} < 3.29\right) = 0.999.$$

Ilustracja CTG dla $x_i \in \mathcal{U}(0, 1)$, $i = 1, \dots, 12$:



Estymator Monte Carlo całki ma asymptotycznie rozkład Gaussa!

Gaussowski generator liczb losowych oparty o CTG

- Niech $x_i \in \mathcal{U}(0, 1)$, $i = 1, 2, \dots, n$, weźmy $R_n = \sum_{i=1}^n x_i$ wówczas:

$$\left. \begin{array}{l} E(x_i) = \frac{1}{2} \\ V(x_i) = \frac{1}{12} \end{array} \right\} \implies \left\{ \begin{array}{l} E(R_n) = \frac{n}{2} \\ V(R_n) = \frac{n}{12} \end{array} \right.$$

\implies **Ćw. A5:** Obliczyć powyższe wartości oczekiwane i wariancje.

- ▶ Z powyższego mamy:

$$\frac{R_n - n/2}{\sqrt{n/12}} \xrightarrow{n \rightarrow \infty} N(0, 1),$$

tzn. dostajemy **standardowy rozkład Gaussa**.

- ▷ **Wygodny wybór:** $n = 12$

$\implies (R_{12} - 6)$ – „praktyczny” **Gaussowski generator liczb losowych**.

- ▶ **UWAGA:** Ogony rozkładu Gaussa nie są dobrze reprodukowane przez tego typu generator (zawsze są skończone, podczas gdy w dokładnym rozkładzie są nieskończone!).

\implies **Ćw. N1.3:** Skonstruować Gaussowski generator liczb losowych oparty o CTG i porównać z dokładnym rozkładem Gaussa dla kilku wartości n – użyć systemu ROOT.

Résumé: matematyczne własności metody Monte Carlo

Niech $u_i \in \mathcal{U}(a, b)$, wówczas z **PWL** mamy:

$$\underbrace{\frac{1}{n} \sum_{i=1}^n f(u_i)}_{\text{estymator MC całki}} \xrightarrow{n \rightarrow \infty} \frac{1}{b-a} \int_a^b f(u) du .$$

Matematyczne własności estymatora Monte Carlo:

- (i) Jeżeli $V(f) < \infty$, to estymator MC jest **zgodny**, tzn. jest zbieżny do prawdziwej wartości całki dla dużych n .
- (ii) Estymator MC jest **nieobciążony** dla wszystkich n , tzn. wartość oczekiwana estymatora jest prawdziwą wartością całki. \Rightarrow **Ćw. A6: Pokazać.**
- (iii) Estymator MC ma asymptotycznie **rozkład normalny** (osiąga rozkład Gaussa dla dużych n).
- (iv) **Odchylenie standardowe** estymatora MC jest dane wzorem:

$$\sigma = \frac{1}{\sqrt{n}} \sqrt{V(f)} .$$

\Rightarrow **Ćw. A7: Pokazać powyższe.**

Praktyczne oszacowanie wariancji estymatora Monte Carlo

Niech

$$I \equiv \frac{1}{b-a} \int_a^b f(x) dx = E(f).$$

► **Estymator MC całki:**

$$\hat{I} = \frac{1}{n} \sum_{i=1}^n f(x_i), \quad x_i \in \mathcal{U}(a, b).$$

► **Wariancja estymatora MC całki:**

$$V(\hat{I}) = \frac{1}{n} V(f) = \frac{1}{n} \left\{ E(f^2) - [E(f)]^2 \right\} = \frac{1}{n} \left\{ \frac{1}{b-a} \int_a^b f^2(x) dx - I^2 \right\}.$$

▷ Do obliczenia $V(\hat{I})$ potrzebna jest znajomość całki I !

→ $V(\hat{I})$ – **wariancja teoretyczna estymatora \hat{I}** .

► W praktyce $V(\hat{I})$ oszacowuje się przez (\Rightarrow **Ćw. A8: Pokazać, że jest to estymator nieobciążony**):

$$\hat{V}(\hat{I}) = \frac{1}{n} \hat{V}(f), \quad \hat{V}(f) = \frac{1}{n-1} \sum_{i=1}^n \left[f(x_i) - \frac{1}{n} \sum_{i=1}^n f(x_i) \right]^2.$$

► **Estymator MC odchylenia standardowego:** $\hat{\sigma} = \sqrt{\hat{V}(\hat{I})}$.