MC metoda – opakování

① Formulovat úlohu jako střední hodnotu funkce náhodné proměnné x:

$$\langle f \rangle = \int_\Omega f(x)\, p(x)\, dx$$

spojitá náhodná prom. x

$x \in \Omega$ .. hodnoty

$p(x)$ na Ω ... distrib. funkce (pravděpodobnostní rozdělení)

míra $p(x)dx$

$$\langle f \rangle = \sum_i f(x_i)\, p_i$$

diskrétní: $x \in \{x_i\}$

pravděpodobnosti: $\{p_i\}$

② Nahradit středováním přes statistický soubor: $x \to \{\xi_1, \xi_2, \xi_3, \dots \xi_N\}$

$$\langle f \rangle \simeq \frac{1}{N} \sum_i f(\xi_i) \longrightarrow \langle f \rangle$$

$N \to \infty$

③ Pak platí: zákon velkých čísel $\Rightarrow$

Centrální limit věta: $\frac{1}{N} \sum_{i=1}^{N} f(\xi_i)$ je kolem $\langle f \rangle$ rozděleno Gaussovsky s chybou $\sigma = \frac{\sigma_f}{\sqrt{N}}$

Použití na integraci — integrace MC (Monte Carlo)

přímočaré: $\int_\Omega f(\vec{x})\, d\vec{x} = \langle f(x)\rangle \cdot V_\Omega$ kde $\vec{x}$ rozdělena rovnoměrně v obl. Ω

$$\simeq \frac{V_\Omega}{N}\sum_{i=1}^{N} f(\vec{x}_i)$$

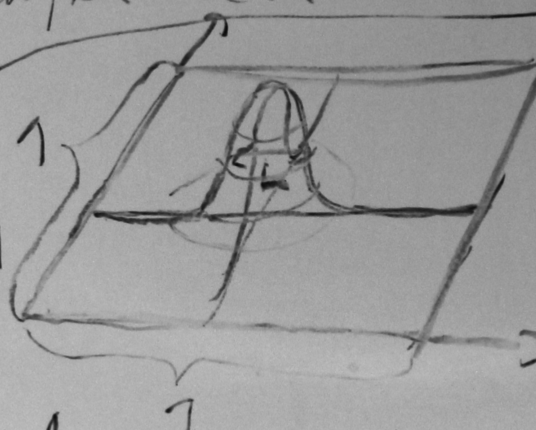
tj $p(x) = \frac{1}{V_\Omega}$ ... konst. ← objem Ω

... chyba $\sim \frac{\sigma_f}{\sqrt{N}}$

$$\sigma_f^2 = \langle (f-\langle f\rangle)^2\rangle = \int_\Omega \left[f(\vec{x}) - \langle f\rangle\right]^2 d\vec{x}$$

pozn.: složitost implementace ani chyba nezávisí na dimenzi d ... $\vec{x} \in \mathbb{R}^d$

! · σ_f může! PŘ: peak šířky $L<1$ v oblasti délky 1



$$V\langle f\rangle \simeq f\cdot L^d$$

$$\sigma_f^2 = \langle f^2\rangle - \langle f\rangle^2 \simeq f^2 L^d - f^2 L^{2d} \simeq f^2 L^d$$

tj relat. chyba $\simeq \frac{\sigma_f}{\sqrt{N}} \cdot \frac{1}{\langle f\rangle V} \simeq \frac{1}{\sqrt{N}} \frac{\sqrt{f^2 L^d}}{f L^d} = \frac{1}{\sqrt{N L^d}}$ ... diverguje pro $d \to \infty$

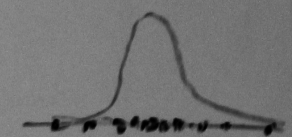
Optimalizace MC integrace ... nerovnoměrné rozdělení

$$I = \int_\Omega f(x)\,dx = \int_\Omega \frac{f(x)}{p(x)}\, p(x)\,dx$$ ... nová fce $\tilde f(x) = \frac{f(x)}{p(x)}$

... nová náhodná proměnná $\tilde X$:

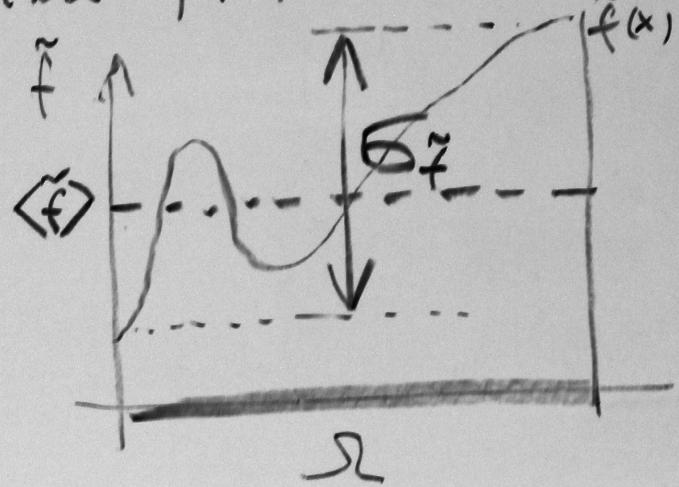
$x \in \Omega$

$$I = \langle \tilde f \rangle_{\tilde X} \simeq \frac{1}{N} \sum_{i=1}^{N} \tilde f(\xi_i)$$

pravd. rozdělení $p(x)$: 

Stat. soubor $X \to \{\xi_1, \xi_2, \xi_3 \dots \xi_N\}$

Volba $p(x)$ ... minimalizovat variaci $\sigma_{\tilde f}$ fce $\tilde f = \frac{f(x)}{p(x)}$



→ snaha, aby $\tilde f \sim$ konst.

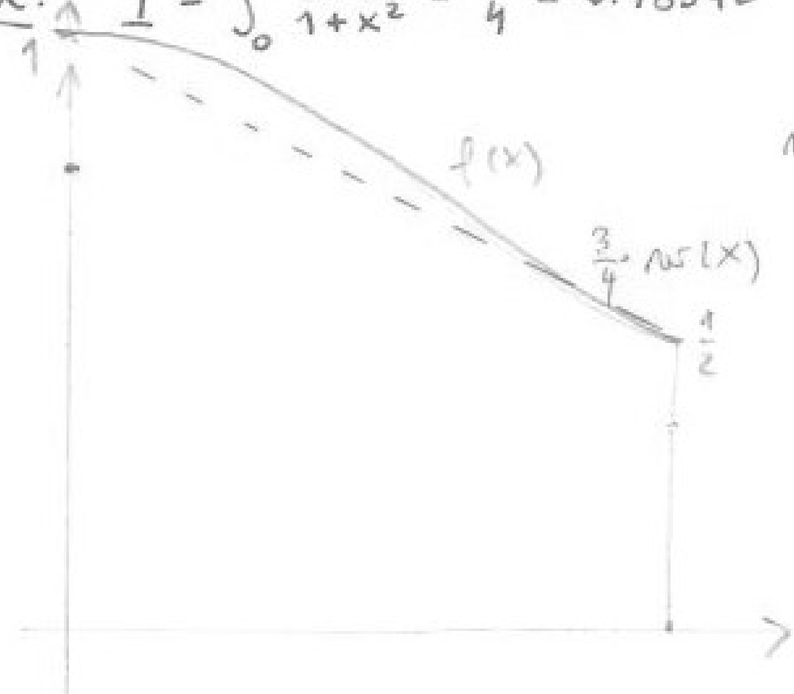
tj. $p(x) \sim f(x) \cdot c$

(a přitom jsme uměli generovat $\{\xi_i\}$ s rozdělením $p(x)$)

Table 8.1 Monte Carlo evaluation of the integral (8.4) using two different weight functions, $w(x)$. The exact value is 0.78540.

N	$w(x)=1$ I	$w(x)=1$ σ_I	$w(x)=\frac{1}{3}(4-2x)$ I	$w(x)=\frac{1}{3}(4-2x)$ σ_I
10	0.81491	0.04638	0.79982	0.00418
20	0.73535	0.03392	0.79071	0.00392
50	0.79606	0.02259	0.78472	0.00258
100	0.79513	0.01632	0.78838	0.00194
200	0.78677	0.01108	0.78529	0.00140
500	0.78242	0.00719	0.78428	0.00091
1000	0.78809	0.00508	0.78524	0.00064
2000	0.78790	0.00363	0.78648	0.00045
5000	0.78963	0.00227	0.78530	0.00028

Příklad: $I = \int_0^1 \frac{dx}{1+x^2} = \frac{\pi}{4} \doteq 0.78540$

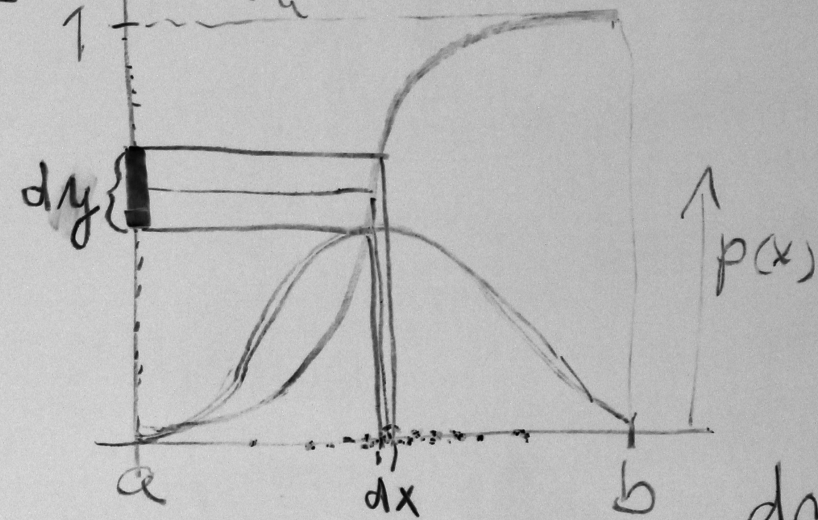


$w(x) = \frac{1}{3}(4-2x)$

Generování náhodné proměnné s předem danou distribuční funkcí

① Přímá metoda v 1D

trik: $y=\int_a^x p(x)\,dx$



náhodná prom $X \in \langle a,b\rangle \ldots p(x)$

$\updownarrow$ rovnoměrné rozd. $y\in\langle 0,1\rangle$

$\hookrightarrow \xi_i^{(y)} \longleftrightarrow \xi_i^{(x)} = X\left[\xi_i^{(y)}\right]$

$y[x] \ldots$ inverze $X[y]$

počet bodů:

$$dn = N\,dy = N\,y'\,dx = N\,p(x)\,dx$$

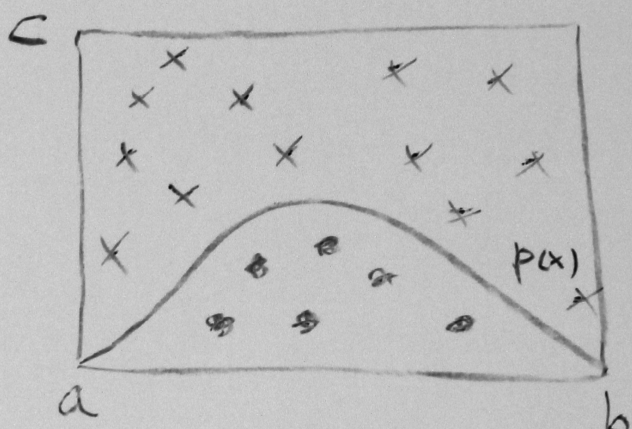
Návod: Ⓐ najít $y(x)\equiv\int_a^x p(t)\,dt$ a $y^{-1}\equiv X[y]$

Ⓑ generovat rovnoměrné $y\in\langle 0,1\rangle \longrightarrow \left\{\xi_i^{(y)}\right\}_{i=1}^{N}$

$\longrightarrow \xi_i^{(x)} = X\left[\xi_i^{(y)}\right] \ldots$ dá soubor $\left\{\xi_i^{(x)}\right\}_{i=1}^{N}$

pozn: pokud nelze analyticky $\int^{x} p(x)$ nebo y^{-1} je třeba numericky

② von Neumann rejection (odhazování)



a) generujeme náhodně (rovnoměrně)

$x_i \in \langle a,b \rangle$; $y_i \in \langle a, c \rangle$

$c > p(x) \ \forall x$

b) hodnoty $y_i > p(x_i)$ zahodíme

hodnoty $y_i \leq p(x_i)$ přijmeme

do souboru $\xi_j = x_i$

pozn: dá se snadno zobecnit do více dim.

generujeme $\vec{x}_i \in \Omega$ a $y_i \in \langle 0, c \rangle$

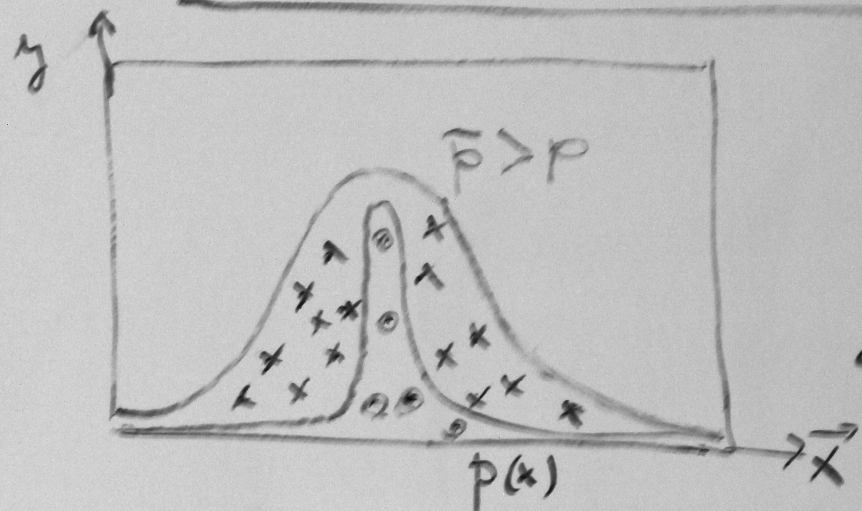
a podle

$y_i > p(\vec{x}_i)$ zahodíme

$y_i \leq p(\vec{x}_i)$ přijmeme

pozn: pokud nelze analyticky $\int^x p(x)$ nelze y^{-1}
je třeba numericky

(2) von Neumann rejection pro úzká rozdělení



... zahazujeme moc bodů
→ trvá dlouho

vylepšení: srovnávací rozdělení
... x ... $\bar{p}(x)$

a) generujeme rovnoměrně v oblasti pod $\bar{p}$ a b) přijmáme pod p

technické detaily postupu:

a) generujeme $\vec{x}_i$ náhodnou prom s distr. fcí $\bar{p}(x)$
generujeme $t \in \langle 0,1 \rangle$ rovnoměrně: $y_i = t\,\bar{p}(x_i)$

b) ? $\boxed{y_i > p(x_i)}$ zahodíme
$\boxed{y_i \leq p(x_i)}$ přijmeme $\boxed{\vec{\xi}_j = \vec{x}_i}$

③ Přímé metody pro Gaussovo (normální) rozdl.

$p_0(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$ … $x \in \mathbb{R}$ ….. rozdělení $\langle x \rangle = 0$, $\sigma = 1$

stačí → generuje $\{\xi_i\}$

potom $\{\sigma(\xi_i + \mu)\} \equiv \{\bar{\xi}_i\}$ $\langle\rangle = \mu$, σ

přeškálovaní na

metoda ① vyžaduje spec. fci $Erf(x)$ a její inverzi.

… jednodušší … generovat najednou dvě x_1, x_2 … $p(\vec{x}) \sim e^{-\frac{\vec{x}^2}{2}} = p_0(x_1) \cdot p_0(x_2)$

platí: $e^{-\frac{1}{2}(x_1^2 + x_2^2)} dx_1 dx_2 = e^{-\frac{1}{2}r^2} \cdot r\, dr\, d\theta = e^{-t} dt\, d\theta$ kde $t = \frac{1}{2}r^2$

tj ekvivaletně generujeme $\boxed{\theta \in \langle 0, 2\pi) \text{ rovnoměrně}}$

použitím ① … $\int_0^t -e^{-t} = 1 - e^{-t}$ ← $t \in (0, \infty)$ s distr $p(t) = e^{-t}$

→ $y(t) = 1 - e^{-t}$ ….. $t = y^{-1} = -\ln(1 - y)$ … $\boxed{\text{rovnoměrně } y \in (0, 1\rangle}$

pak $x_1 = \sqrt{-2\ln(1-y)}\cos\theta$ a $x_2 = \sqrt{-2\ln(1-y)}\sin\theta$ … Gauss.

③ Přímé metody pro Gaussovo (normální) rozdělení

a) metoda: generujeme rovnoměrně $\theta \in \langle 0, 2\pi)$
$\eta \in \langle 0, 1)$

pak $x_1 = \sqrt{-2\ln(1-\eta)}\cos\theta$
$x_2 = \sqrt{-2\ln(1-\eta)}\sin\theta$ jsou nezávislé Gauss náhodné prom $\langle x \rangle = 0$ $\sigma = 1$

b) jednoduchá přibližná metoda založená na Centrální limitní větě

... náhodná prom $\eta \in \langle 0, 1)$ rovnoměrné: $\langle \eta \rangle = \frac{1}{2}$
$\sigma_\eta^2 = \frac{1}{12}$

návod: generujeme $\eta_1, \eta_2, \dots, \eta_{12}$

pak $x = \left[\sum_{i=1}^{12} \eta_i - 6 \right]$ je blízká normálnímu rozdělení s $\langle x \rangle = 0$ a $\sigma = 1$

VYZKOUŠEJTE
... chyba distr. pro $x \in (-2, 2)$ je menší než 2 %.

4) Metropolisův algoritmus

Metropolis, Rosenbluth & Rose...
Teller & Teller 1953

(obecné metody založené na náhodné procházce,
Markovovský proces – řetězec)

generujeme posl. $\vec{\xi}_1 \to \vec{\xi}_2 \to \vec{\xi}_3 \to \dots \to \vec{\xi}_i \to \dots \to \vec{\xi}_N$

náhodným procesem —— Markov proces

s takovými pravidly, aby pro $N \to \infty$ byla hustota bodů $\{\vec{\xi}_i\} \in \Omega$ předem daná distr. fce $p(\vec{x})$

pravidla konstrukce překvapivě jednoduchá a flexibilní

4 Metropolisův algoritmus

Metropolis, Rosenbluth & Rose... Teller & Teller 1953

Algoritmus: (1) vybrat $\vec{\xi}_1$

OPAKOVAT $N \times$

(2) z bodu $\vec{\xi}_n$ pokusně vykročíme do $\vec{\xi}_t$

(např. rovnoměrně náhodně v δ-okolí $\vec{\xi}_n$)

(3) def: $r \equiv p(\vec{\xi}_t)/p(\vec{\xi}_n)$

$\boxed{r > 1}$ přijmeme $\vec{\xi}_{n+1} = \vec{\xi}_t$

$\boxed{r < 1}$ generujeme $\eta \in \langle 0,1 \rangle$ rovnoměrně

přijetí s pravd. r

$\boxed{\eta < r}$ přijmeme

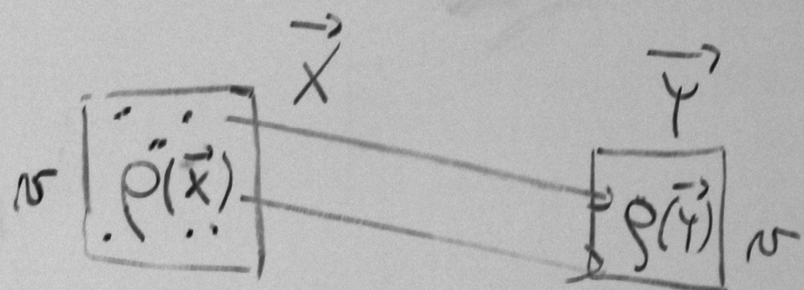
$\boxed{\eta > r}$ zahodíme

Tvrzení: $\{\vec{\xi}_n\}$ pro $N \to \infty$ jsou rozděleny s hust. pravd. $p(\vec{x})$

Náznak důkazu (intuitivní ... detailní rovnováha)

po dlouhé době ... distribuce $\vec{\xi}_n$ kolem $\vec{x}$ ... $\rho(\vec{x})$

$$N(x) = \nu\,\rho(x)$$



$$\Delta N(Y) = N(X)\,P(X\to Y) - N(Y)\,P(Y\to X) = 0$$

detailní rovnováha

.. – ustálený stav ρ_e

tj.
$$\frac{P(Y\to X)}{P(X\to Y)} = \frac{N(X)}{N(Y)} = \frac{\rho_e(X)}{\rho_e(Y)} = \frac{p(X)}{p(Y)}$$

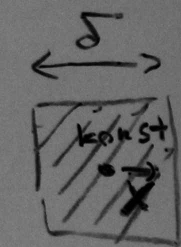
potřebujeme ukázat $\rho_e(\vec{x}) \sim p(\vec{x})$ tj.

Náznak důkazu

v našem algoritmu:

$$P(X\to Y) = \underbrace{T(X\to Y)}_{\text{pravd. udělat krok do } Y}\, A(X\to Y)$$

pravd. udělat krok do Y ... např



jediný požadavek ... symetrické! ∴ $T(X\to Y) = T(Y\to X)$

pravděpodobnost přijetí:

$p(x) > p(y)$: $A(x\to Y) = \dfrac{p(Y)}{p(X)}$, $A(Y\to X) = 1$

$p(Y) > p(x)$: $A(Y\to X) = \dfrac{p(X)}{p(Y)}$, $A(X\to Y) = 1$

v obou př.:

$$\frac{P(Y\to X)}{P(X\to Y)} = \frac{A(Y\to X)}{A(X\to Y)} = \frac{p(X)}{p(Y)}$$

$\boxed{\text{CBD QED}}$

Komentáře a praktické otázky

- **volba délky kroku δ**

  - malý ... skoro vše přijmeme; hodně kroků, abychom se někam dostali, Následující ξ_n, ξ_{n+1} hodně korelované (závislé)

  - velký .... hodně zahozených kroků (zpomaluje), a pomalu vzorkujeme důležité obl., ale rychle propátráme velké oblasti (i odlehlé)

optimum ... cca $\frac{1}{2}$ kroků přijata

- **volba poč. podmínky:** vhodné relaxovat (pár kroků na poč. zahodit)

- statistická závislost násl. $\vec{\xi}_n, \vec{\xi}_{n+1}$

ODBOČKA: statistická závislost — korelace

dvě náhodné veličiny $a \dots p_a(a) \dots \{\alpha_i\}_{i=1}^N$

$b \dots p_b(b) \quad \{\beta_i\}_{i=1}^N$

Def: korelační koeficient

$$C_{ab} \equiv \frac{\langle ab\rangle - \langle a\rangle\langle b\rangle}{\sigma_a \sigma_b} = \frac{\langle ab\rangle - \mu^2}{\underbrace{\langle a^2\rangle - \mu^2}_{\sigma^2}}$$

pro totožné veličiny $\underbrace{\langle a\rangle = \langle b\rangle}_{=\mu}$; $\sigma_a = \sigma_b$

výpočet statist. souborem: $C_{ab} \xleftarrow[N\to\infty]{} \dfrac{\sum_i (\alpha_i - \mu_a)(\beta_i - \mu_b)}{\sqrt{\sum_i (\alpha_i - \mu_a)^2}\sqrt{\sum_i (\beta_i - \mu_b)^2}}$

pozn: $C_{ab} = \dfrac{\langle (a-\mu_a)(b-\mu_b)\rangle}{\sigma_a \sigma_b}$

… důsl … $0 < C_{ab} \leq 1$ … Schwartz … $\dfrac{\langle x|y\rangle}{\|x\|\,\|y\|} = \cos\theta$

význam korelač. koef C_{ab} – statistická (ne)závislost?

... (a,b) chápu jako 1 vektor náhod. prom. ... $p(a,b)$

... tj $\langle ab\rangle = \iint ab\, p(a,b)\, da\, db$

$$\langle a\rangle = \iint a \underbrace{\left[p(a,b)\, db\right]}_{p_a(a)} da = \int a\, p_a(a)\, da$$

statisticky nezávislé: $p(a,b) = p_a(a)\, p_b(b)$

... stat. vlast. a nezávisí na b ... $\langle f(a)\rangle$ ... jen z $p_a(a)$ $\forall f$

a $\langle ab\rangle = \iint ab\, p_a(a)\, p_b(b)\, da\, db = \langle a\rangle\langle b\rangle$

$\Rightarrow$ $\boxed{C_{ab} = 0}$

opačný extrém: 100 % korelace

když naměřím $a \longrightarrow \alpha_i$ pak též b: $\beta_i = \alpha_i$

tj. $p(a,b) = p_a(a)\,\delta(a-b)$ ... $p_a(x) = p_b(x) = p(x)$

$$\langle ab \rangle = \iint ab\, \delta(a-b)\, p(a)\, da\, db = \int a^2 p(a)\, da = \langle a^2 \rangle$$

tj. $C_{ab} = \dfrac{\sigma_\alpha^2}{\sigma_\alpha \sigma_\alpha} = 1$

Význam pro posl. náhodných prom:

autokorelační fce $C_{\xi_0 \xi_k} \equiv C(k) \equiv \dfrac{\langle \xi_0 \xi_k \rangle - \langle \xi_0 \rangle^2}{\langle \xi_0^2 \rangle - \langle \xi_0 \rangle^2}$

prakticky spočteme za předp. transl. invariance:

$$C(k) = \frac{\langle \xi_i \xi_{i+k} \rangle - \langle \xi_i \rangle^2}{\langle \xi_i^2 \rangle - \langle \xi_i \rangle^2}$$

+ přes soubor:

$$\langle \xi_i \xi_{i+k} \rangle = \frac{1}{N-k} \sum_{i=1}^{N-k} \xi_i \xi_{i+k}$$