

Komputerowa analiza danych doświadczalnych

Wykład 4
18.03.2016

dr inż. Łukasz Graczykowski
lgraczyk@if.pw.edu.pl

Semestr letni 2015/2016



Zamiana zmiennych

Transformacje liniowe

Propagacja niepewności

Metody Monte Carlo



Zamiana zmiennych

Zamiana zmiennych

- Na poprzednich wykładach zayważyliśmy, że **dowolna funkcja zmiennej losowej X jest także zmienną losową:**

$$Y = H(X)$$

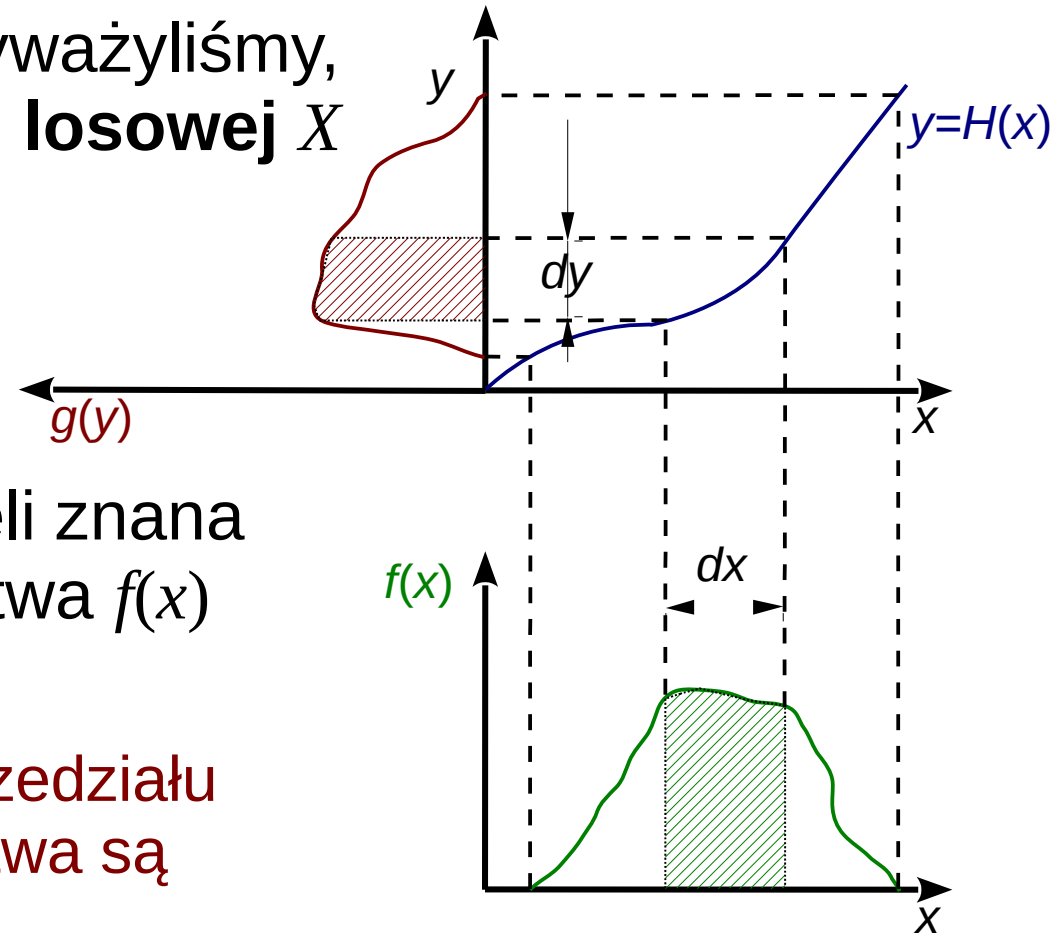
- Pytanie:** jaka jest gęstość prawdopodobieństwa $g(y)$, jeżeli znana jest gęstość prawdopodobieństwa $f(x)$ i oczywiście funkcja $Y=H(X)$?

- dla **infinitesimalnie małego przedziału zmienności** prawdopodobieństwa są

równe: $f(x)dx = g(y)dy$

- z tego wynika: $dy = \left| \frac{dy}{dx} \right| dx$ $dx = \left| \frac{dx}{dy} \right| dy$

- zatem rozkłady są powiązane: $g(y) = \left| \frac{dx}{dy} \right| f(x)$

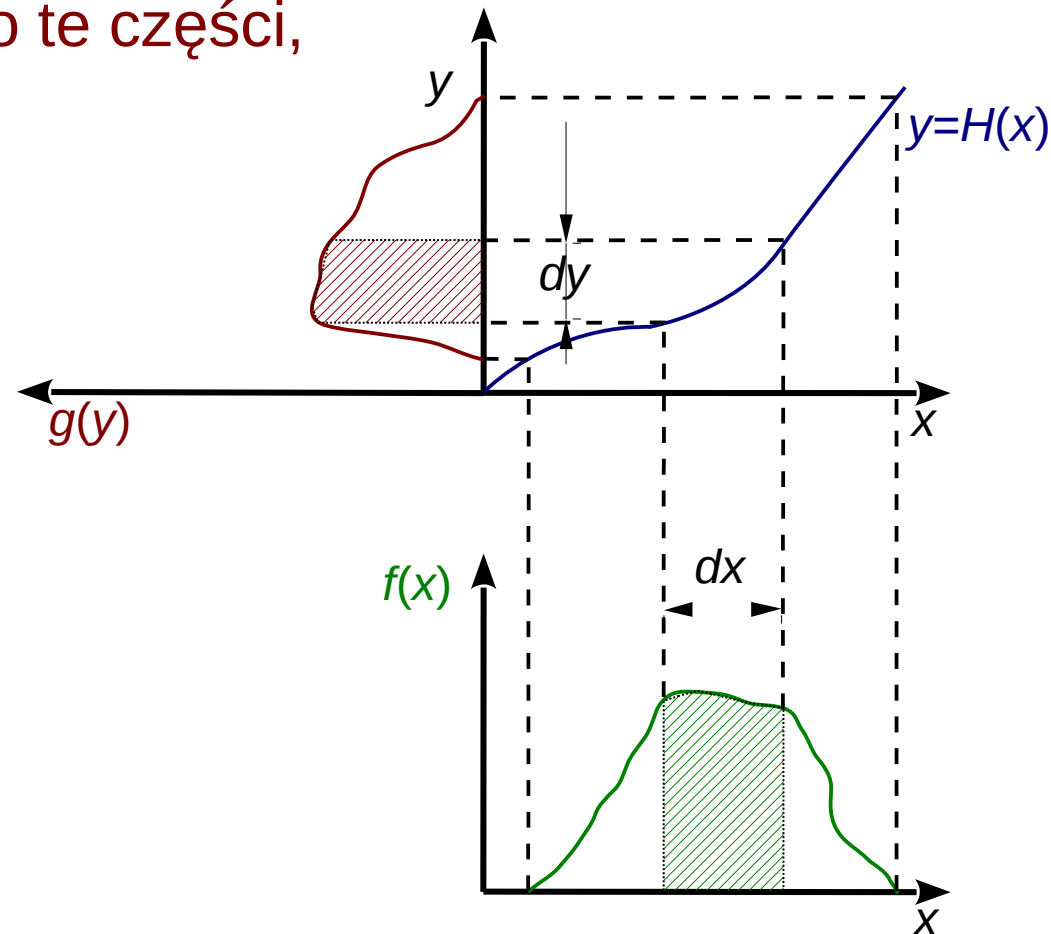


Zamiana zmiennych

- Warunki:

- funkcja $y=H(x)$ musi być wzajemnie jednoznaczna
- funkcje wieloznaczne (np. $y=\sqrt{(x)}$) rozpatrujemy oddzielnie – tylko te części, które są dodatnie $y=+\sqrt{(x)}$
- rozkłady są unormowane:

$$\int_{-\infty}^{\infty} g(y) dy = \int_{-\infty}^{\infty} f(x) dx = 1$$



Zamiana zmiennych - przykład

- Pytanie:** mamy zmienną losową X opisaną rozkładem jednorodnym $f(x)=1$ na przedziale od 0 do 1. Jaka będzie postać funkcji $Y=H(X)$, aby otrzymać zadany (znany) rozkład $g(y)$?

- **metoda odwracania dystrybuanty**

$$f(x) dx = g(y) dy$$

$$\text{gdy } f(x) \equiv 1 \Rightarrow dx = g(y) dy = dG(y)$$

$$g(y) = G'(y)$$

↑
dystrybuanta

$$\int dx = \int dG(y)$$

$$x = G(y)$$

musi istnieć

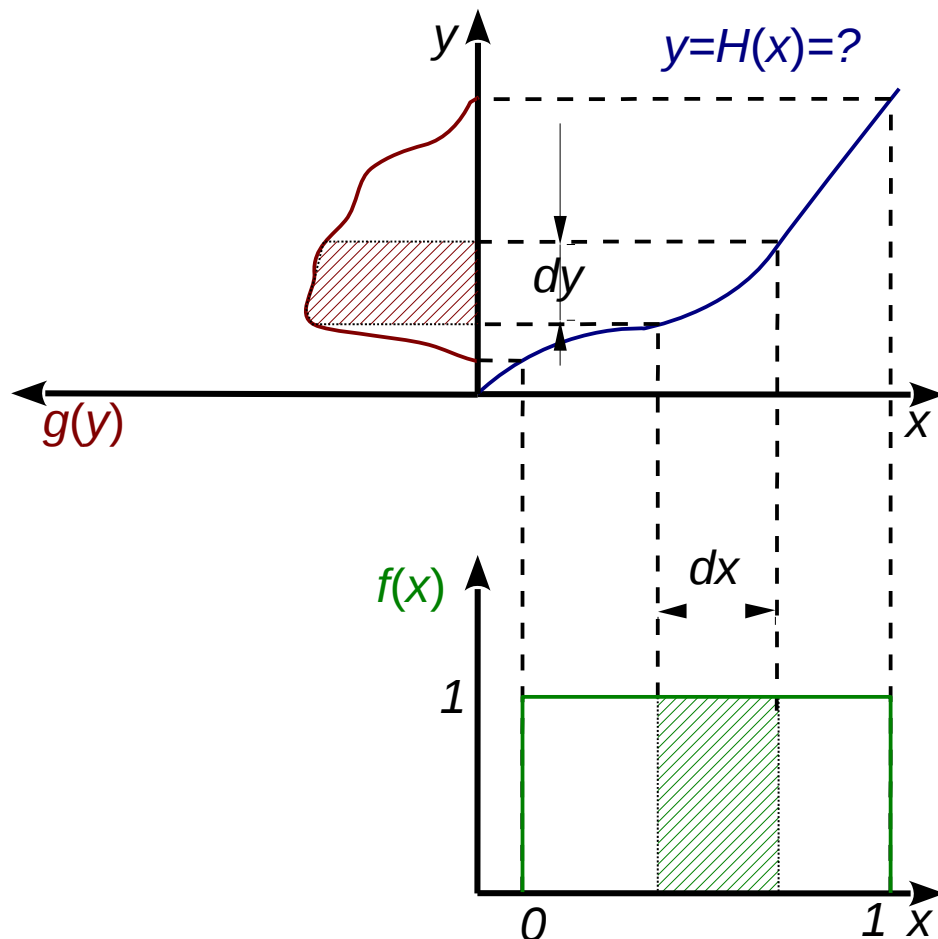
$$y = G^{-1}(x) \equiv H(x)$$

$$x_{min} = G(y_{min}), x_{max} = G(y_{max})$$

$$y_{min} = G^{-1}(0), y_{max} = G^{-1}(1)$$

Czyli: liczymy dystrybuantę $x=G(y)$
a następnie funkcję odwrotną $y=G^{-1}(x)$

Zmienna losowa X po transformacji
 $y=G^{-1}(x)$ ma rozkład $g(y)$



Zamiana zmiennych - przypadek wielowym.

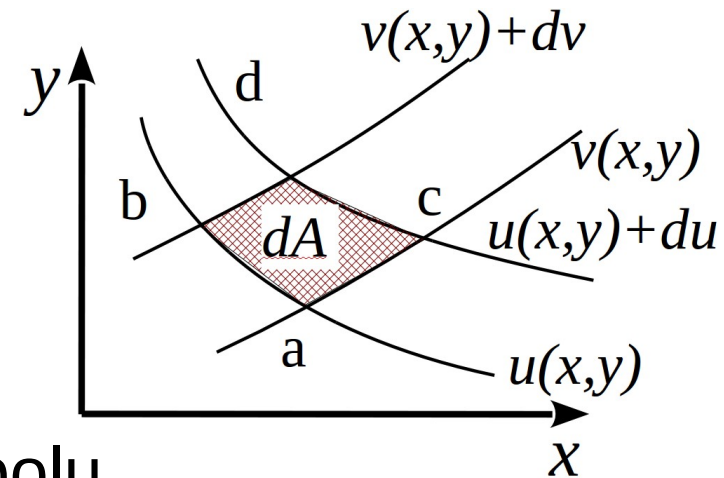
- Mamy dwie zmienne losowe X i Y , dokonujemy zamiany zmiennych na U i V :

$$(X, Y) \rightarrow (U, V) \quad U = U(X, Y) \quad V = V(X, Y)$$

- Szukamy funkcji J (jakobian):

$$g(u, v) = f(x, y) \left| J \left(\frac{x, y}{u, v} \right) \right|$$

- Rysunek przedstawia płaszczyznę (x, y) , z układem krzywych dla $u = \text{const}$ i $v = \text{const}$



Mamy więc mały element powierzchni o polu $dA = dx dy$ w zmiennych $x, y \rightarrow$ zatem jego pole możemy policzyć jako pole równoległoboku o wierzchołkach a, b, c, d :

- współrzędne pierwszych 3 wierzchołków:

$$\begin{aligned} x_a &= x(u, v) & y_a &= y(u, v) \\ x_b &= x(u, v + dv) & y_b &= y(u, v + dv) \\ x_c &= x(u + du, v) & y_c &= y(u + du, v) \end{aligned}$$

Zamiana zmiennych - przypadek wielowym.

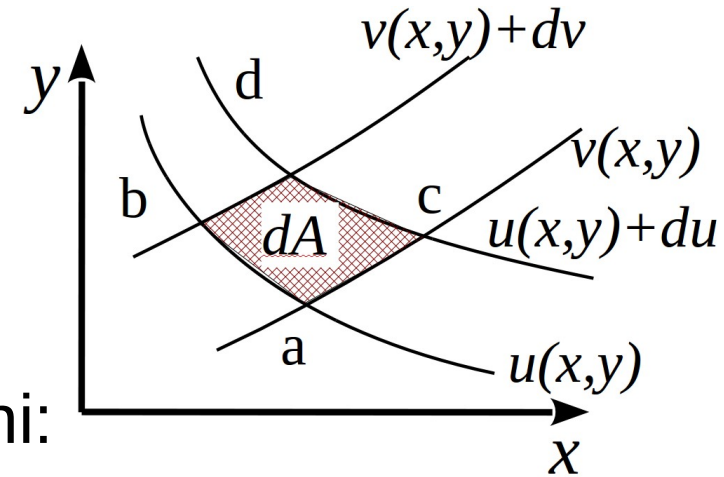
- Rozwijamy w szereg Taylora:

$$x_b = x(u, v) + \frac{\partial x}{\partial v} dv \quad y_b = y(u, v) + \frac{\partial y}{\partial v} dv$$

$$x_c = x(u, v) + \frac{\partial x}{\partial u} du \quad y_c = y(u, v) + \frac{\partial y}{\partial u} du$$

- Z dokładnością do znaku pole powierzchni:

$$dA = \begin{vmatrix} 1 & x_a & y_a \\ 1 & x_b & y_b \\ 1 & x_c & y_c \end{vmatrix} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{vmatrix} du dv \equiv J \left(\frac{x, y}{u, v} \right) du dv$$



jakobian przejścia
(transformacji)

- W ogólnym przypadku, dla n zmiennych:

$$Y_1 = Y_1(\mathbf{X})$$

$$Y_2 = Y_2(\mathbf{X})$$

⋮

$$Y_n = Y_n(\mathbf{X})$$

$$g(\mathbf{y}) = \left| J \left(\frac{\mathbf{x}}{\mathbf{y}} \right) \right| f(\mathbf{x})$$

$$J \left(\frac{\mathbf{x}}{\mathbf{y}} \right) = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_2}{\partial y_1} & \dots & \frac{\partial x_n}{\partial y_1} \\ \frac{\partial x_1}{\partial y_2} & \frac{\partial x_2}{\partial y_2} & \dots & \frac{\partial x_n}{\partial y_2} \\ \dots & \dots & \dots & \dots \\ \frac{\partial x_1}{\partial y_n} & \frac{\partial x_2}{\partial y_n} & \dots & \frac{\partial x_n}{\partial y_n} \end{vmatrix}$$

Zamiana zmiennych - przypadek wielowym.

$$f(x, y) = 1/(2 \cdot a^2), |x| + |y| < a$$

- Dokonujemy zamiany zmiennych:

$$\begin{aligned} u(x, y) &= x + y & x(u, v) &= \frac{1}{2}(u + v) \\ v(x, y) &= x - y & y(u, v) &= \frac{1}{2}(u - v) \end{aligned} \Leftrightarrow$$

- Obliczamy jacobian:

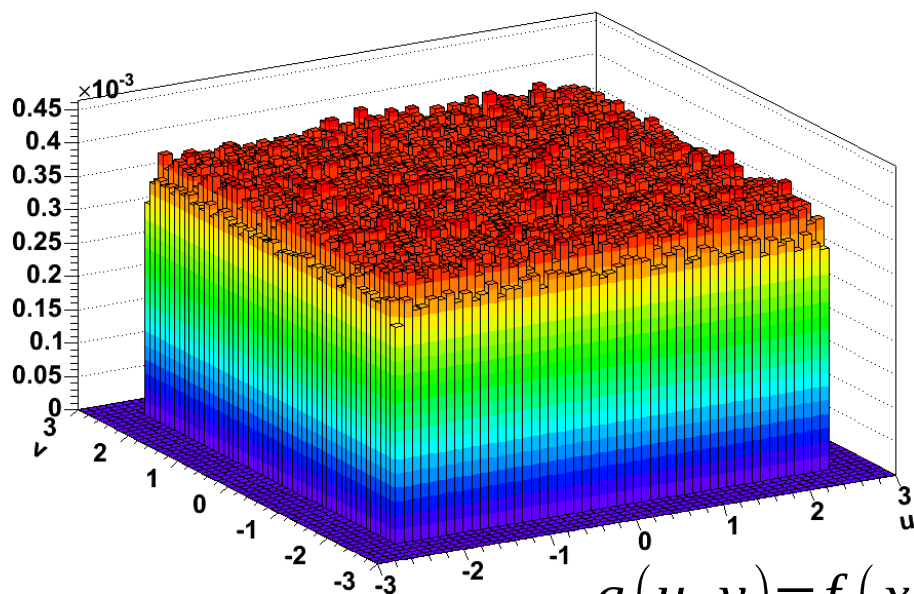
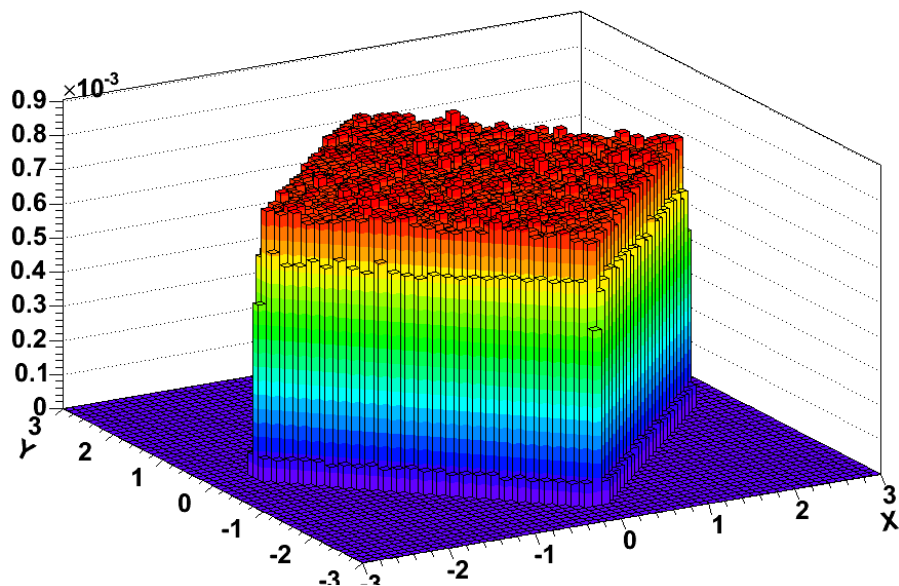
$$\frac{\partial x}{\partial u} = \frac{1}{2} \qquad \frac{\partial y}{\partial u} = \frac{1}{2}$$

$$\frac{\partial x}{\partial v} = \frac{1}{2} \qquad \frac{\partial y}{\partial v} = -\frac{1}{2}$$

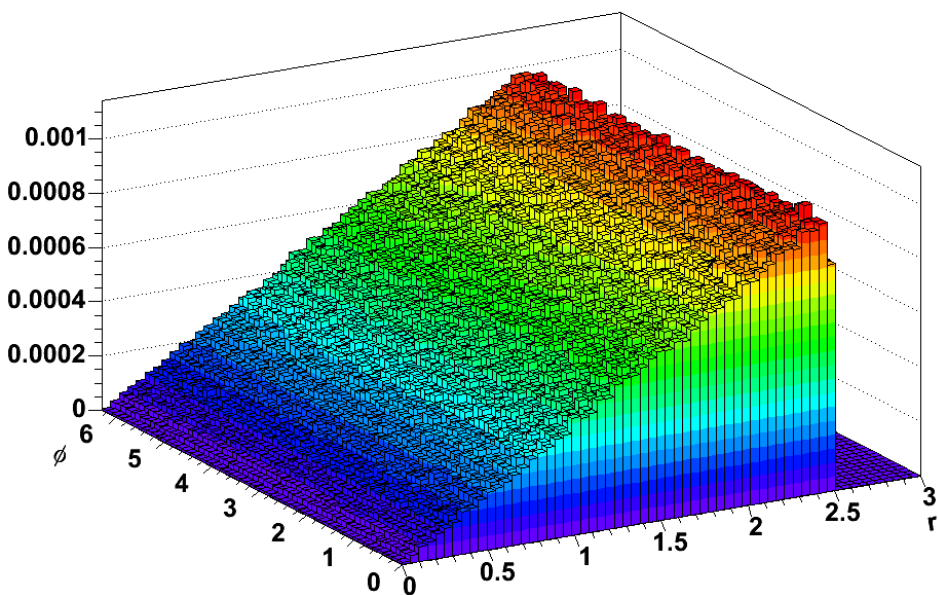
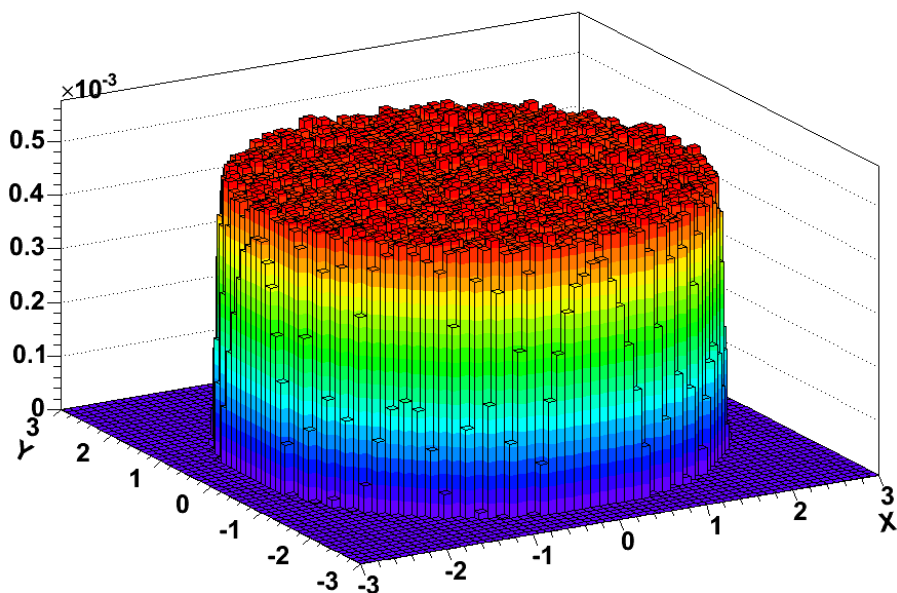
$$J \begin{pmatrix} x, y \\ u, v \end{pmatrix} = \frac{1}{2} \cdot \frac{1}{2} - \frac{1}{2} \cdot -\frac{1}{2} = \frac{1}{2}$$

- Otrzymujemy $g(u, v)$:

$$g(u, v) = f(x, y) \cdot J \begin{pmatrix} x, y \\ u, v \end{pmatrix} = \frac{1}{2} f(x, y) = 1/(4 \cdot a^2), |u| < a; |v| < a$$



Zamiana zmiennych - przypadek wielowym.



$$f(x, y) = \frac{1}{\pi R^2}, \sqrt{x^2 + y^2} < R$$

- Dokonujemy zamiany zmiennych (do wsp. bieg.):

$$r(x, y) = \sqrt{x^2 + y^2} \quad x(r, \varphi) = r \cdot \cos \varphi$$
$$\varphi(x, y) = \operatorname{arctg} \frac{x}{y} \quad y(r, \varphi) = r \cdot \sin \varphi$$

- Obliczamy jacobian:

$$\frac{\partial x}{\partial r} = \cos \varphi \quad \frac{\partial y}{\partial r} = \sin \varphi$$

$$\frac{\partial x}{\partial \varphi} = -r \cdot \sin \varphi \quad \frac{\partial y}{\partial \varphi} = r \cdot \cos \varphi$$

$$J \left(\begin{array}{c} x, y \\ r, \varphi \end{array} \right) = r \cdot \cos^2 \varphi + r \cdot \sin^2 \varphi = r$$

- Otrzymujemy $g(r, \varphi)$:

$$g(r, \varphi) = f(x, y) \left| J \left(\begin{array}{c} x, y \\ r, \varphi \end{array} \right) \right| = \frac{r}{\pi R^2}$$



Transformacje liniowe

Propagacja niepewności

Transformacje liniowe

- Najczęściej, ze względu na prototę, posługujemy się transformacjami liniowymi (inne transformacje najczęściej aproksymujemy liniowymi, rozwijając na szereg Taylora)

– **funkcje** $Y = (Y_1, Y_2, \dots, Y_r)$ **są liniowymi funkcjami zmiennych**

$$\begin{aligned} Y_1 &= a_1 + t_{11} X_1 + t_{12} X_2 + \dots + t_{1n} X_n \\ Y_2 &= a_2 + t_{21} X_1 + t_{22} X_2 + \dots + t_{2n} X_n \\ &\vdots \\ Y_r &= a_r + t_{r1} X_1 + t_{r2} X_2 + \dots + t_{rn} X_n \end{aligned} \quad X = (X_1, X_2, \dots, X_n)$$

- W zapisie macierzowym: $Y = T X + a$
- Wartość oczekiwana Y : $E(Y) = \hat{y} = T \hat{x} + a$
- Macierz kowariancji Y :
$$\begin{aligned} C_Y &= E((Y - \hat{y})(Y - \hat{y})^T) \\ &= E((T X + a - T \hat{x} - a)(T X + a - T \hat{x} - a)^T) \\ &= E(T(X - \hat{x})(X - \hat{x})^T T^T) \\ &= T E((X - \hat{x})(X - \hat{x})^T) T^T \\ C_Y &= T C_X T^T \end{aligned}$$

Propagacja niepewności

- Załóżmy, że znamy pewne wartości oczekiwane (wyniki pomiarów) \hat{x}_i oraz ich niepewności $\sigma(x_i)$ i kowariancje $cov(x_i, x_j)$. Szukamy niepewności funkcji: $Y(\mathbf{X})$
- Jeśli niepewności są małe, to możemy dokonać rozwinięcia na szereg Taylora wokół wartości oczekiwanych:

$$Y_i = Y_i(\hat{\mathbf{x}}) + \left(\frac{\partial y_i}{\partial x_1} \right)_{x=\hat{\mathbf{x}}} (X_1 - \hat{x}_1) + \dots + \left(\frac{\partial y_i}{\partial x_n} \right)_{x=\hat{\mathbf{x}}} (X_n - \hat{x}_n) + \text{wyrazy wyższego rzędu}$$

- w notacji macierzowej: $\mathbf{Y} = \mathbf{Y}(\hat{\mathbf{x}}) + \mathbf{T}(\mathbf{X} - \hat{\mathbf{x}}) + \text{wyrazy wyższego rzędu}$

- gdzie:

$$\mathbf{T} = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \dots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \dots & \frac{\partial y_2}{\partial x_n} \\ \vdots & & & \\ \frac{\partial y_n}{\partial x_1} & \frac{\partial y_n}{\partial x_2} & \dots & \frac{\partial y_n}{\partial x_n} \end{pmatrix}_{\mathbf{X}=\hat{\mathbf{x}}}$$

Propagacja niepewności

- Niepewności Y to elementy diagonalne macierzy kowariancji:

$$C_Y = T C_X T^T$$

- jak widać, zależą one nie tylko od elementów diagonalnych macierzy C_X , ale również od jej elementów pozadiagonalnych
- tylko i wyłącznie jeżeli wszystkie zmienne X są niezależne, tj. $c_{ij}=0$, dla $i \neq j$ możemy zapisać:

$$\sigma^2(Y_i) = \sum_{j=1}^n \left(\frac{\partial y_i}{\partial x_j} \right)^2 \sigma^2(X_j)$$

- co daje nam **prawo propagacji błędów** znane z Wykładu 1:

$$\sigma(y_i) = \sqrt{\sum_{j=1}^n \left(\frac{\partial y_i}{\partial x_j} \right)^2 \sigma^2(x_j)}$$

Znaczenie macierzy kowariancji - przykład

- Mierzmy w układzie kartezjańskim współrzędne punktu (x,y) . Pomiar x i y są niezależne.
- Z jakiejś przyczyny (np. inne urządzenie pomiarowe) pomiar daje trzykrotnie większą niepewność współrzędnej y niż x .

- Macierz kowariancji: $C_{xy} = \begin{pmatrix} 1 & 0 \\ 0 & 9 \end{pmatrix}$

- Dokonujemy transformacji na współrzędne biegunowe:

$$r(x, y) = \sqrt{x^2 + y^2} \quad \varphi(x, y) = \operatorname{arctg} \frac{x}{y} \quad x(r, \varphi) = r \cdot \cos \varphi \quad y(r, \varphi) = r \cdot \sin \varphi$$

- Policzmy macierz transformacji (dla prostoty – w punkcie $(1,1)$)

$$T = \begin{pmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial \varphi}{\partial x} & \frac{\partial \varphi}{\partial y} \end{pmatrix} = \begin{pmatrix} \frac{x}{r} & \frac{y}{r} \\ -\frac{y}{r^2} & \frac{x}{r^2} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \quad T^T = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{2} \\ \frac{1}{\sqrt{2}} & \frac{1}{2} \end{pmatrix}$$

Znaczenie macierzy kowariancji - przykład

- Macierz kowariancji dla zmiennych biegunowych:

$$C_{r\phi} = T C_{xy} T^T = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 9 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{2} \\ \frac{1}{\sqrt{2}} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 5 & \frac{4}{\sqrt{2}} \\ \frac{4}{\sqrt{2}} & \frac{5}{2} \end{pmatrix}$$

- Widzimy, że **nie są one niezależne**

- Licząc w drugą stronę: $C_{r\phi} = \begin{pmatrix} 5 & \frac{4}{\sqrt{2}} \\ \frac{4}{\sqrt{2}} & \frac{5}{2} \end{pmatrix}$

- Pozadiagonalne elementy macierzy kowariancji są BARDZO WAŻNE**

$$T' = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \phi} \end{pmatrix} = \begin{pmatrix} \cos \phi & -r \sin \phi \\ -\sin \phi & r \cos \phi \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -1 \\ \frac{1}{\sqrt{2}} & 1 \end{pmatrix}$$

$$C_{xy} = T' C_{r\phi} T'^T = \begin{pmatrix} \frac{1}{\sqrt{2}} & -1 \\ -\frac{1}{\sqrt{2}} & 1 \end{pmatrix} \begin{pmatrix} 5 & \frac{4}{\sqrt{2}} \\ \frac{4}{\sqrt{2}} & \frac{5}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 9 \end{pmatrix}$$

ŻLE! gdyby uwzględniać tylko diagonalne niepewności są zupełnie inne →

$$C_{xy} = T' C_{r\phi} T'^T = \begin{pmatrix} \frac{1}{\sqrt{2}} & -1 \\ -\frac{1}{\sqrt{2}} & 1 \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & \frac{5}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix}$$



Generacja liczb (pseudo)losowych za pomocą komputera

Liczby (pseudo)losowe

- Do tej pory zajmowaliśmy się jedynie opisem zmiennych losowych (ich właściwości) – nie zajmowaliśmy się tym, jak je otrzymać
- Bardzo często potrzebujemy jednak posłużyć się ciągiem (losowych) wartości jakiejś zmiennej losowej, która opisana jest danym rozkładem prawdopodobieństwa
- Metoda otrzymywania takich liczb może pochodzić z badania zjawiska fizycznego (np. rozpad promieniotwórczy, szumy fal elektromagnetycznych)
- Liczby takie możemy wygenerować również w komputerze – trzeba jednak pamiętać, że taki ciąg będzie ciągiem pseudolosowym, gdyż komputer cechuje się zachowaniem deterministycznym (można więc te liczby “przewidzieć”)
- Metody analizy danych z wykorzystaniem liczb (pseudo)losowych nazywamy **metodami Monte Carlo**



Zapis liczb w komputerze

- Najmniejsza jednostka informacji – **bit** (0 lub 1), system dwójkowy)

- Przykład:

$$213 = 1 \cdot 128 + 1 \cdot 64 + 0 \cdot 32 + 1 \cdot 16 + 0 \cdot 8 + 1 \cdot 4 + 0 \cdot 2 + 1 \cdot 1 \Rightarrow 011010101$$

- Ogólnie – $k-1$ bitów \rightarrow wartość bezwzględna kodowanej liczby, 1 bit \rightarrow znak. Wartość bezwzględna:

$$a = a^{(k-2)} 2^{k-2} + a^{(k-3)} 2^{k-3} + \dots + a^{(1)} 2^1 + a^{(0)} 2^0$$

- W komputerze zapisujemy liczby za pomocą **bajtów** składających się z 8 bitów (najmniejsza adresowalna jednostka pamięci)
- Współcześnie komputery pracują na (16), 32 i 64-bitowych liczbach
- Liczby bez znaku przyjmują wartości od 0 do 2^k-1 , zaś liczby ze znakiem od -2^{k-1} do $2^{k-1}-1$

Zapis liczb w komputerze

- Liczbę zmiennoprzecinkową zapisujemy jako:

$$x = (-1)^s \cdot m \cdot b^c$$

- Gdzie: m – mantysa (część ułamkowa liczby), b – podstawa, c – wykładnik (albo “cecha”, część całkowita liczby), s – bit znaku
- Komputery zapisują liczbę z $b=2$, notacja naukowa to $b=10$
- Do zapisu **liczby zmiennoprzecinkowej** wystarczy zatem **znajomość dwóch liczb całkowitych**
- Przykłady:

$$00010110_{(FP2)}$$

$$s = 0$$

$$c = 0010_{(U2)} = 2$$

$$m = 1,110_{(2)} = 1\frac{3}{4}$$

$$L = (-1)^0 \cdot 1\frac{3}{4} \cdot 2^2 = 7_{(10)}$$

$$0,875_{(10)} = 0,111_{(2)}$$

$$L = (-1)^0 \cdot 1,11_{(2)} \cdot 2^{(-1)}$$

$$c = -1 = 1111_{(U2)}$$

$$m = 1,110_{(2)}$$

$$0,875_{(10)} = 01111110_{(FP2)}$$

Zapis liczb w komputerze

- Zakres zmienności liczb zmiennoprzecinkowych:

$$2^{C_{min}} < |x| < 2^{C_{max}}$$

- Dwie liczby zmiennoprzecinkowe są różne, jeżeli ich mantysy różnią się o minimalną wartość α :

$$x_1 = m \cdot 2^e, x_2 = (m + \alpha) \cdot 2^e, \Delta x = x_2 - x_1 = \alpha \cdot 2^e, \frac{\Delta x}{x} = \frac{\alpha}{m} = \frac{2^0}{2^n} = 2^{-n}$$

- Przykład – typy danych w języku C++:

TYPE	Bit width		Common range
Char	8	(1byte)	-128 to 127
unsigned char	8	(1byte)	0 to 255
signed char	8	(1byte)	-128 to 127
int	32	(4bytes)	-2,147,483,648 to 2,147,483,647
unsigned int	32	(4bytes)	0 to 4,294,967,295
signed int	32	(4bytes)	-2,147,483,648 to 2,147,483,647
short int	16	(2bytes)	-32,768 to 32,767
unsigned short int	16	(2bytes)	0 to 65,535
signed short int	16	(2bytes)	-32,768 to 32,767
long int	32	(4bytes)	Same as int
unsigned long int	32	(4bytes)	Same as unsigned int
signed long int	32	(4bytes)	Same as signed int
float	32	(4bytes)	3.4E-38 to 3.4E+38
double	64		1.7E-308 to 1.7E+308
long double	80		3.4E-4932 to 1.1E+4932
bool	N/A		true or false

Generatory liniowe kongruentne

- Komputer, urządzenie deterministyczne, może generować tylko liczby pseudolosowe
 - kolejna generowana liczba jest funkcją liczb wcześniej wygenerowanych

- Generator liniowy kongruentny (*LCG – Linear Congruential Generator*):

$$x_{j+1} = (a \cdot x_j + c) \bmod m$$

- LCG generuje okresowy ciąg liczb (po jakimś czasie powtarza się)
 - m – maksymalna długość okresu generatora (szczegóły – Brandt)
- Multiplikatywny generator liniowy kongruentny (*MLCG – multiplicative linear congruential generator*), $c=0$:

$$x_{j+1} = (a \cdot x_j) \bmod m$$

- szybsze od LCG, ale nigdy nie dają wartości 0 i mają krótkie okresy

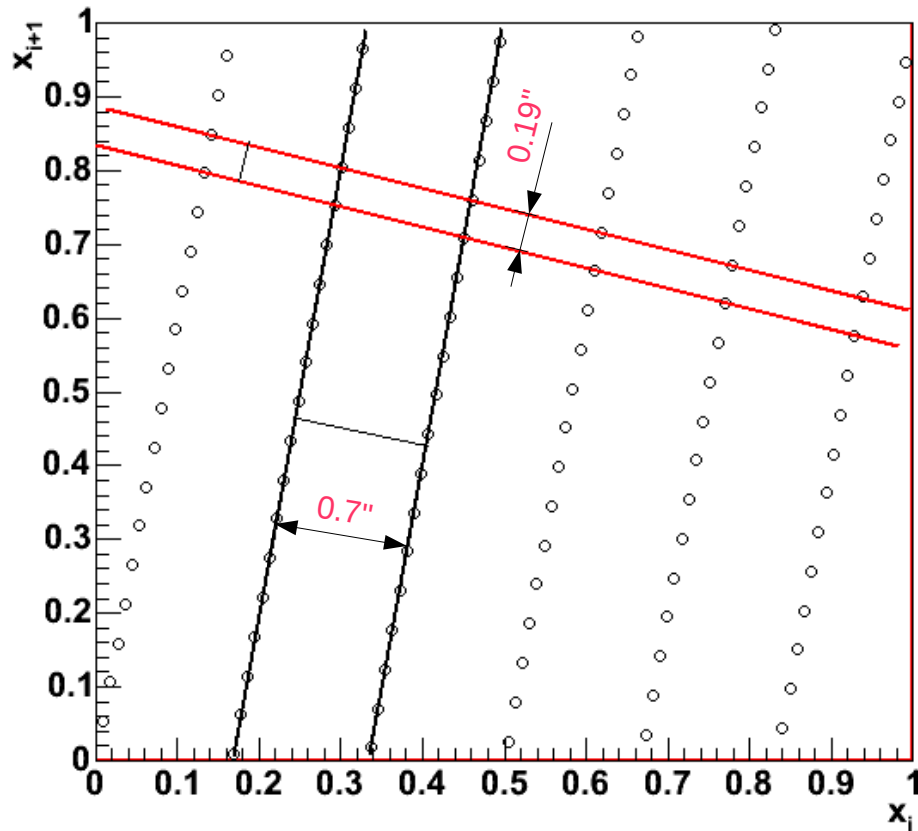
Jakość generatorów losowych

- Najdłuższy okres nie jest jedyną porządaną cechą generatora
- Ważniejsze jest, aby liczby następujące po sobie występowały w sposób jak najbardziej “przypadkowy” - w tym celu wykonujemy tzw. **test widmowy**:
 - wykonujemy dwuwymiarowy wykres (sieć) par: (x_i, x_{i+1})
 - obsadzamy (maksymalnie) m z m^2 węzłów
 - szukamy prostych łączących obsadzone węzły sieci, a następnie wybieramy największą z odległości d_t między nimi
 - jeśli odległości między sąsiednimi prostymi są podobne → jednostajny rozkład węzłów sieci (**tego oczekujemy!**)
 - oczekujemy, że: $d_t \approx m^{-1/2}$
 - jeśli parametry a i m są źle dobrane, to: $d_t \gg m^{-1/2}$

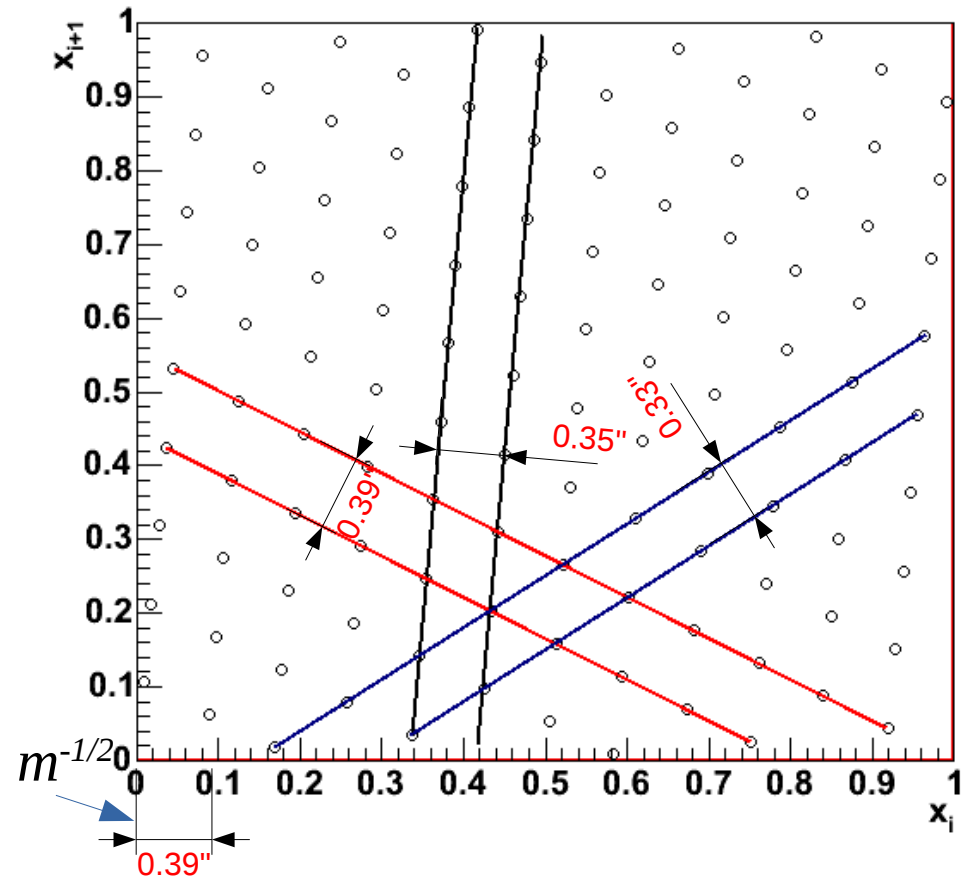
Jakość generatorów losowych

- Najdłuższy okres nie jest jedyną porządaną cechą generatora
- Ważniejsze jest, aby liczby następujące po sobie występowały w sposób jak najbardziej "przypadkowy" - w tym celu wykonujemy tzw. **test widmowy**:
 - **prawy** – poprawny (dobre parametry), **lewy** - niepoprawny

$m = 113, a = 6$



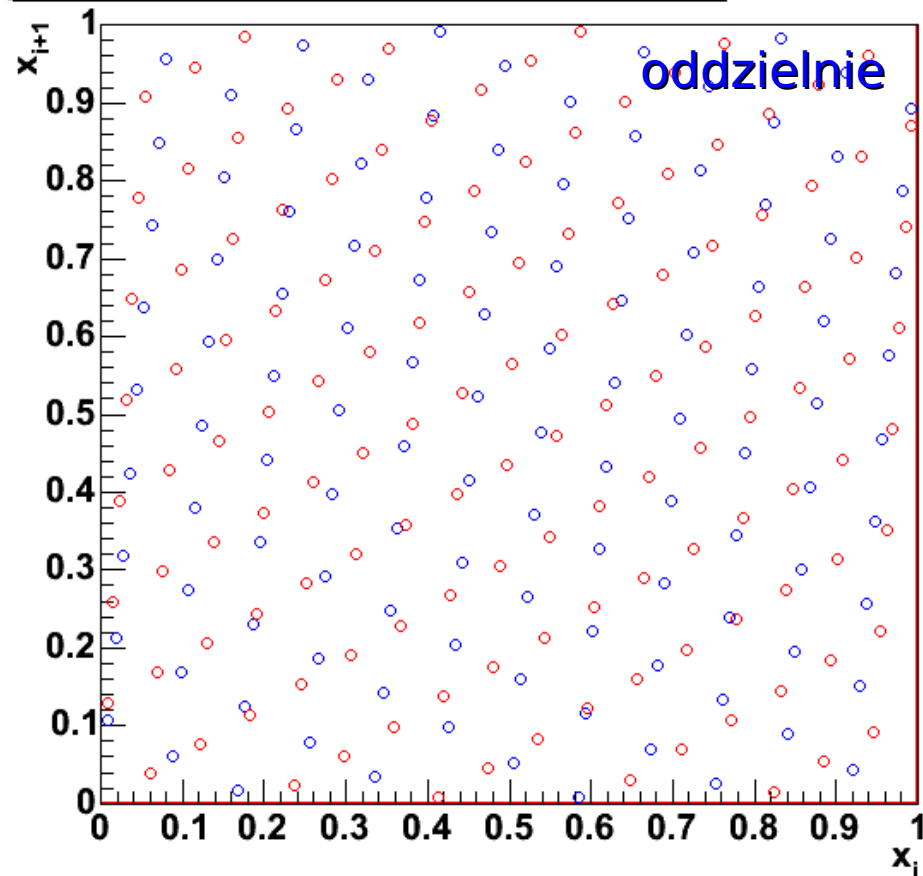
$m = 113, a = 12$



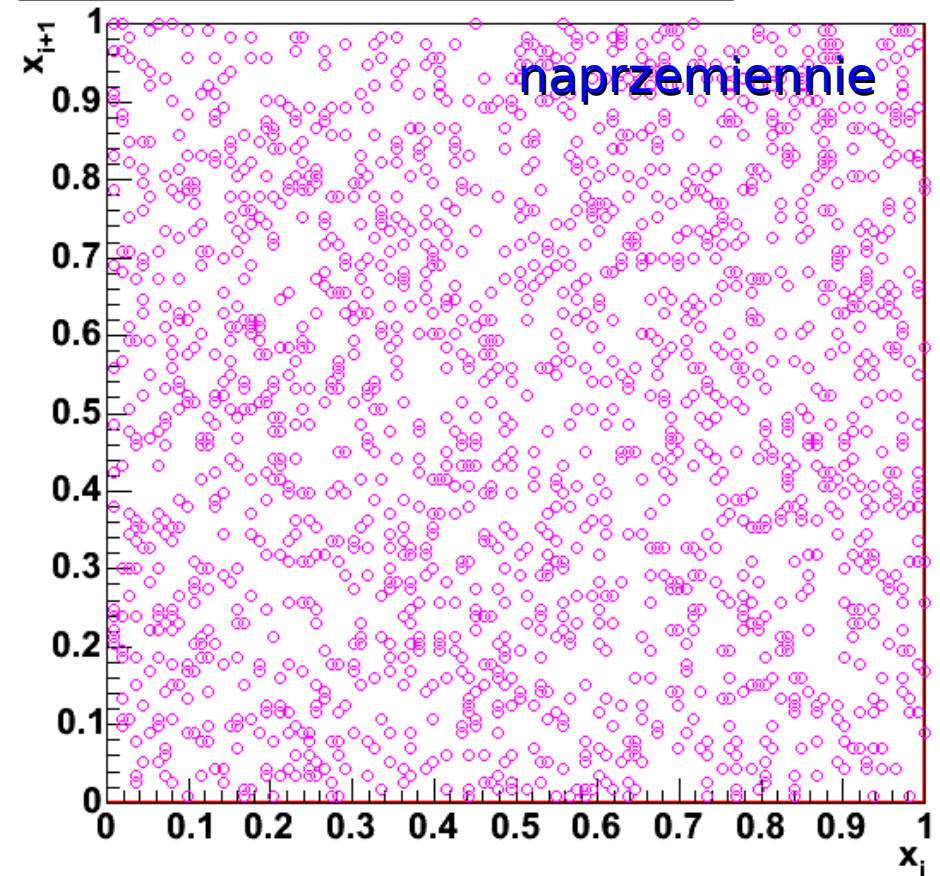
Jakość generatorów losowych

- Najlepsze wyniki (długie okresy) można otrzymać łącząc ze sobą kilka (np. l) generatorów liniowych (o różnych parametrach)
- Ich maksymalny okres wynosi wtedy: $p = 2^{-(l-1)} \cdot \prod_{j=1}^l (m_j - 1)$

$m_1 = 113, a_1 = 12, m_2 = 131, a_2 = 17$



$m_1 = 113, a_1 = 12, m_2 = 131, a_2 = 17$

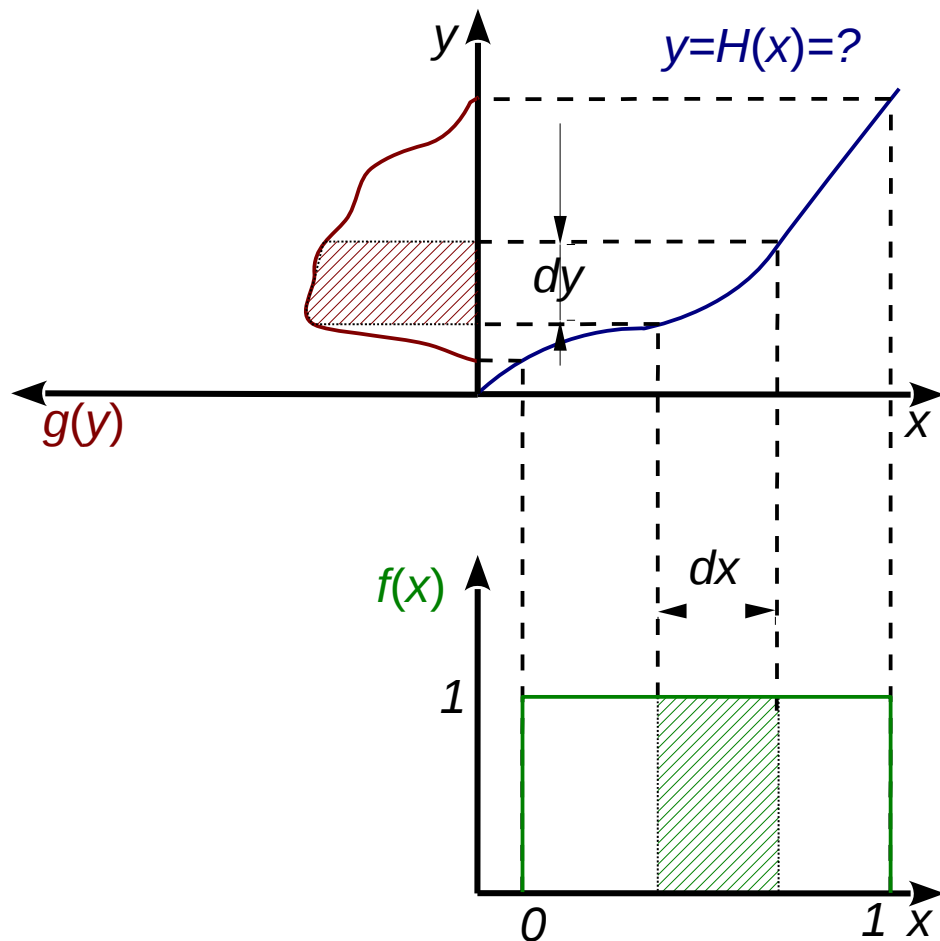




Generacja liczb (pseudo)losowych metodą transformy rozkładu jednorodnego

Transformacja rozkładu jednorodnego

- To już wiemy :) **Metoda z odwrotnością dystrybuanty**
- Transformację rozkładu jednostajnego możemy wykorzystać do generowania liczb losowych o skomplikowanych gęstościach prawdopodobieństwa



$$f(x) dx = g(y) dy$$

$$\text{gdy } f(x) \equiv 1 \Rightarrow dx = g(y) dy = dG(y)$$

$$g(y) = G'(y)$$

↑
dystrybuanta

$$\int dx = \int dG(y)$$

$$x = G(y)$$

↖ musi istnieć

$$y = G^{-1}(x) \equiv H(x)$$

$$x_{min} = G(y_{min}), x_{max} = G(y_{max})$$

$$y_{min} = G^{-1}(0), y_{max} = G^{-1}(1)$$

Czyli: liczymy dystrybuantę $G(y)$
a następnie funkcję odwrotną $G^{-1}(y)$

Zmienna losowa X po transformacji
 $y = G^{-1}(x)$ ma rozkład $g(y)$

Przykład - rozkład Breita-Wignera

- Rozkład Breita-Wignera ma następującą postać:

$$g(y) = \frac{2}{\pi\Gamma} \cdot \frac{\Gamma^2}{4(y-a)^2 + \Gamma^2}$$

- Liczmy dystrybuantę:

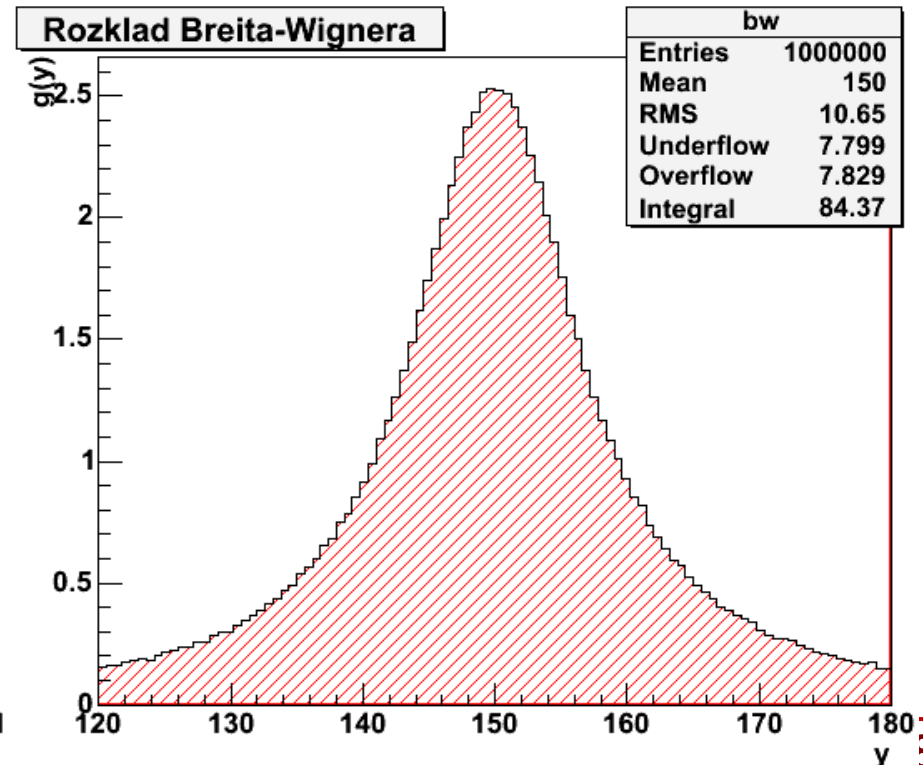
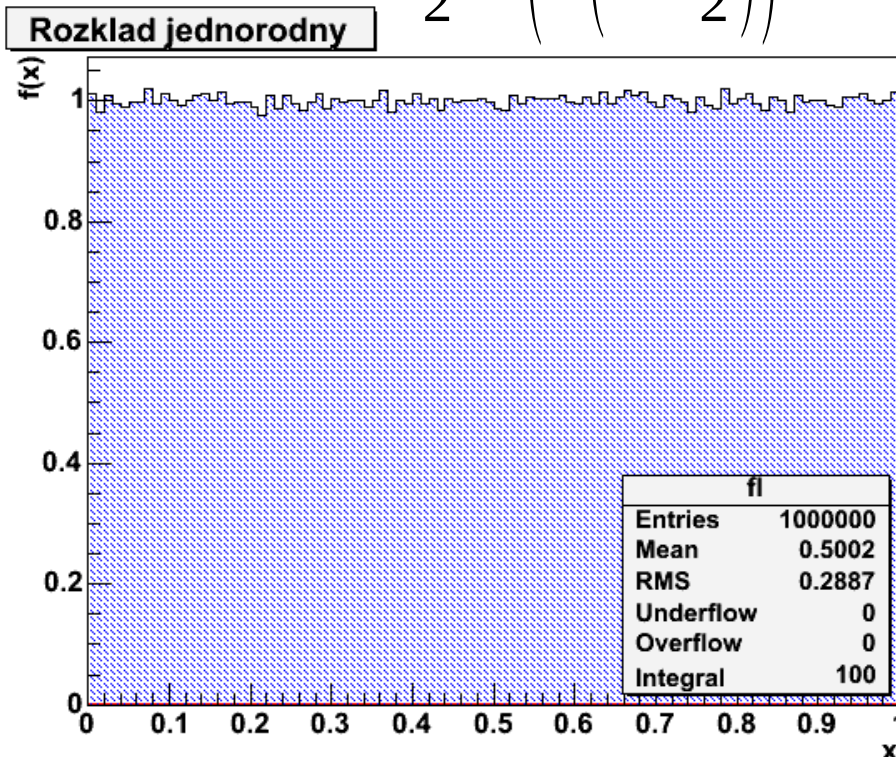
$$x = G(y) = \frac{2}{\pi\Gamma} \cdot \int_{-\infty}^y \frac{\Gamma^2}{4(z-a)^2 + \Gamma^2} dz = \frac{1}{\pi} \operatorname{arctg} \left(\frac{2(y-a)}{\Gamma} \right) + \frac{1}{2}$$

- Odwrotna dystrybuanta:

$$y(x) = G^{-1}(x) = a + \frac{\Gamma}{2} \operatorname{tg} \left(\pi \left(x - \frac{1}{2} \right) \right)$$

$$y(0) = -\infty$$

$$y(1) = \infty$$





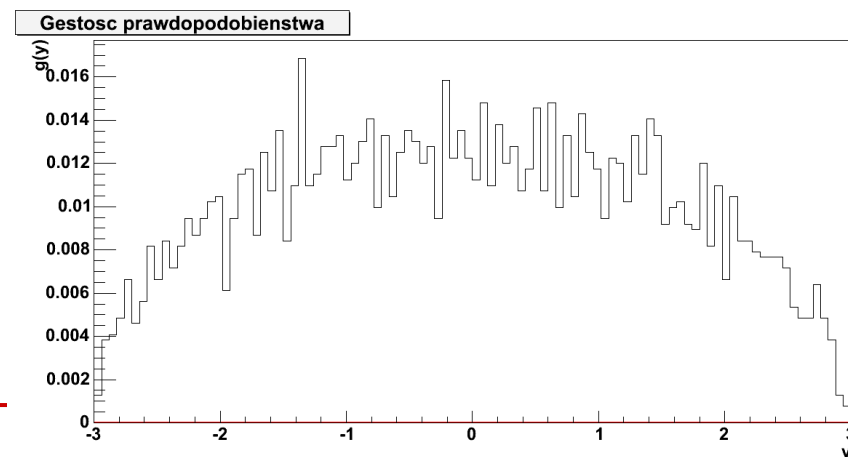
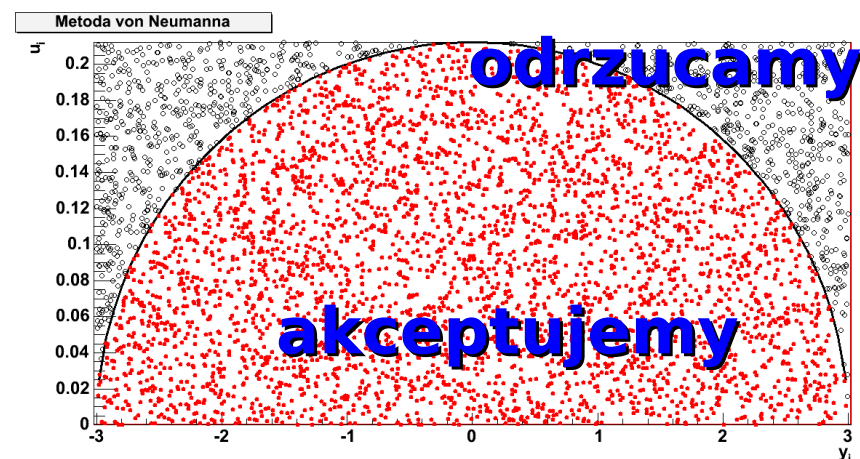
Metoda akceptacji-odrzuceń von Neumanna

Metoda (akceptacji) von Neumanna

- Metoda generacji liczb metodą odwrotnej dystrybuanty ma swoje ograniczenia – dystrybuanta musi być znana (i być funkcją wzajemnie jednoznaczną), czyli musi istnieć funkcja odwrotna
- Metoda (akceptacji-odrzuceń) von Neumanna wymaga znajomości jedynie gęstości prawdopodobieństwa i pozwala na otrzymanie liczb z praktycznie dowolnego rozkładu



- Jak to działa?
 - generujemy parę liczb z rozkładu jednorodnego: (y_i, u_i)
 - sprawdzamy, czy $u_i < g(y_i)$
 - jeśli warunek jest spełniony, akceptujemy liczbę y_i ,
jeśli nie - odrzucamy



Metoda von Neumanna - definicje

- Geometryczny opis metody von Neumanna:
 - chcemy wygenerować liczby losowe opisane gęstością $g(y)$ w przedziale: $a \leq y \leq b$
 - rozważamy krzywą: $u = g(y)$ oraz funkcję stałą: $u = d, d \geq g_{max}$
 - losujemy z rozkładu jednorodnego liczby (y_i, u_i) , które spełniają warunki: $a \leq y_i \leq b, 0 \leq u_i \leq d$
 - jak łatwo zauważyć, nasze punkty układają się w prostokącie na płaszczyźnie (y, u)
 - odrzucaamy wszystkie punkty spełniające nierówność: $u_i \geq g(y_i)$
 - pozostają jedynie punkty położone pod krzywą: $u = g(y)$
 - zaakceptowane wartości y_i podlegają rozkładowi $g(y)$
- Wada metody: w zależności od kształtu funkcji $g(y)$, znaczna część liczb y_i jest odrzucana $\int_a^b g(y) dy$
- Wydajność metody: $E = \frac{a}{(b-a)d}$

Metoda von Neumanna - przypadek wielowym.

- Metodę von Neumanna można uogólnić na wiele wymiarów:
 - mamy funkcję gęstości wielu zmiennych: $g(y_1, y_2, \dots, y_n)$
 - generujemy zbiór liczb z rozkładu jednorodnego: $(y_1^i, y_2^i, \dots, y_n^i, u^i)$
 - dla każdego i sprawdzamy warunek akceptacji: $u^i < g_{\max}(y_1^i, y_2^i, \dots, y_n^i)$
 - akceptujemy lub odrzucamy cały zestaw wygenerowanych liczb dla danego i
- Kilka uwag do metody von Neumanna:
 - za jej generować możemy liczby z dowolnego, nawet bardzo skomplikowanego rozkładu
 - rozkład (czy w ogólności funkcja) $g(y)$ nie musi być nawet unormowany
 - metodę tę można stosować do **obliczania całek oznaczonych**:

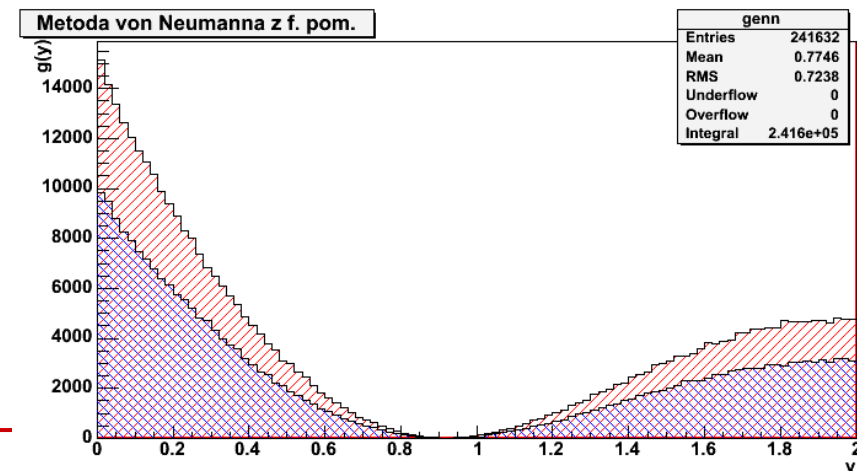
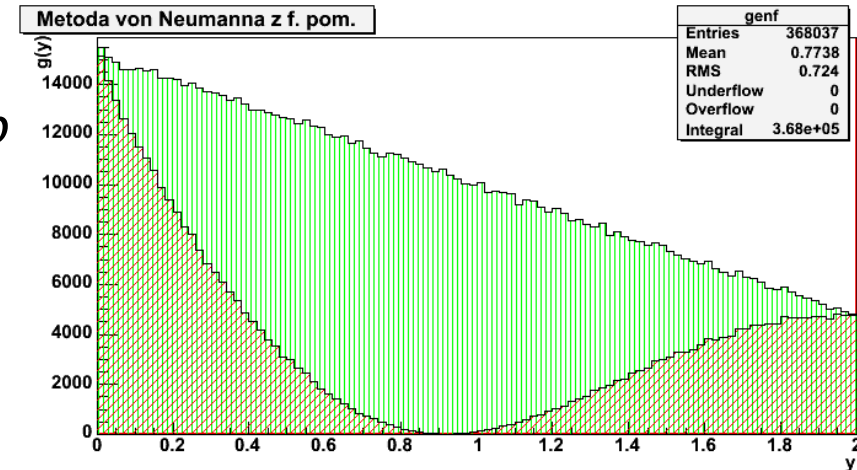
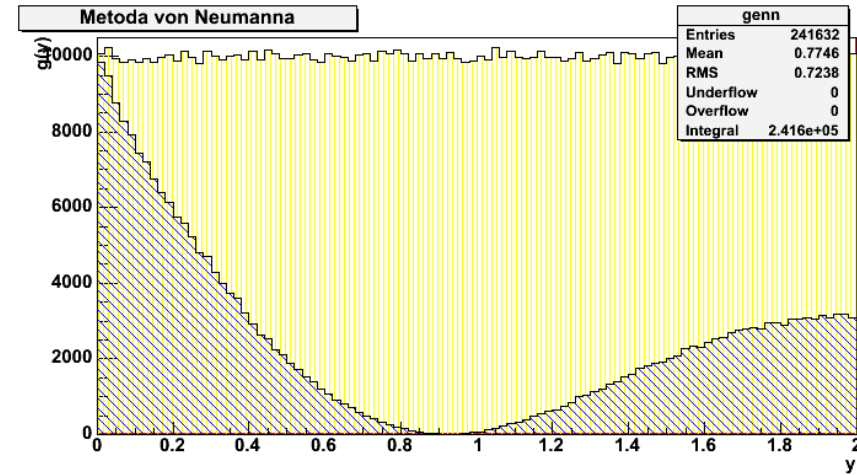
$$E = \frac{\int_a^b g(y) dy}{(b-a)d} \approx \frac{N_{\text{zaakceptowane}}}{N_{\text{wszystkie}}}$$

$$I = \int_a^b g(y) dy \approx \frac{N_{\text{zaakceptowane}}}{N_{\text{wszystkie}}} (b-a)d$$

Metoda von Neumanna z funkcją pomocniczą

- Wydajność metody von Neumanna można poprawić, jeśli odpowiednio zawężymy obszar losowania:
 - wprowadzamy funkcję pomocniczą $s(y)$, z której “łatwo” wygenerować zmienne losowe (np. metodą odwrotnej dystrybuanty), i która spełnia warunek: $g(y) \leq c \cdot s(y)$, $a < y < b$
 - generujemy liczbę losową y_i z rozkładu $s(y)$ na przedziale $a < y_i < b$ oraz liczbę u_i z rozkładu jednorodnego na przedziale $0 < u_i < 1$
 - odrzucaamy liczbę y_i , jeżeli: $u_i \geq \frac{g(y_i)}{c \cdot s(y_i)}$
 - wydajność metody:

$$E = \frac{\int_a^b g(y) dy}{c \int_a^b s(y) dy}$$





KONIEC