



Komputerowa analiza danych doświadczalnych

Wykład 6
6.04.2018

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

Semestr letni 2017/2018



Własności rozkładu normalnego

Centralne twierdzenie graniczne

Funkcja charakterystyczna

Rozkłady wielowymiarowe - elipsa kowariancji

Rozkład normalny standardowy

- Gęstość prawdopodobieństwa:

$$f(x) \equiv \phi_0(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

- rozkład o średniej 0 i wariancji 1

- Dystrybuanta nie ma postaci analitycznej (korzystamy z tabel)

- Rozkład jest unormowany:

$$\int_{-\infty}^{\infty} e^{-x^2/2} dx = \sqrt{2\pi}$$

- Jeśli wprowadzimy zmienną:

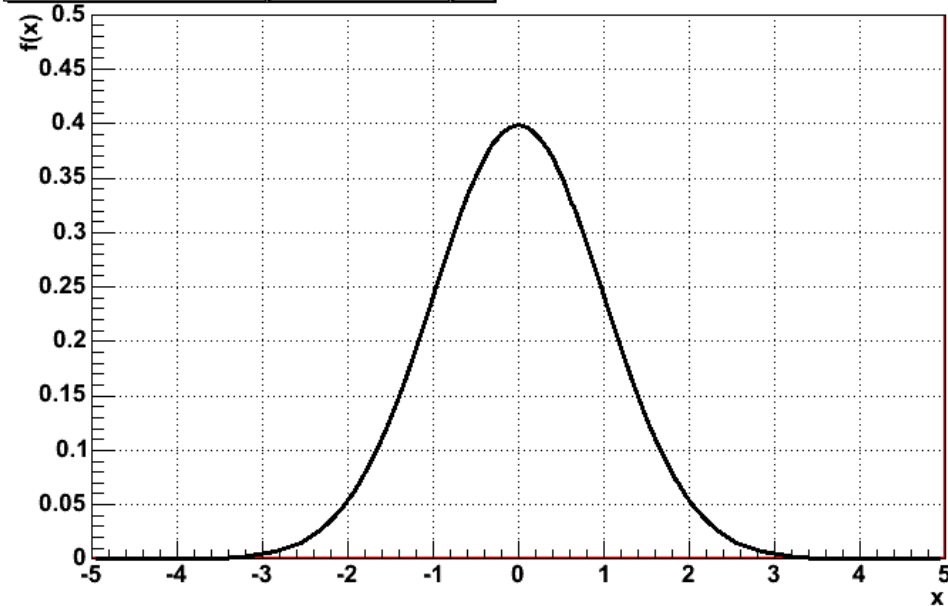
$$Y = (X - a)/b$$

- Otrzymamy rozkład Gaussa:

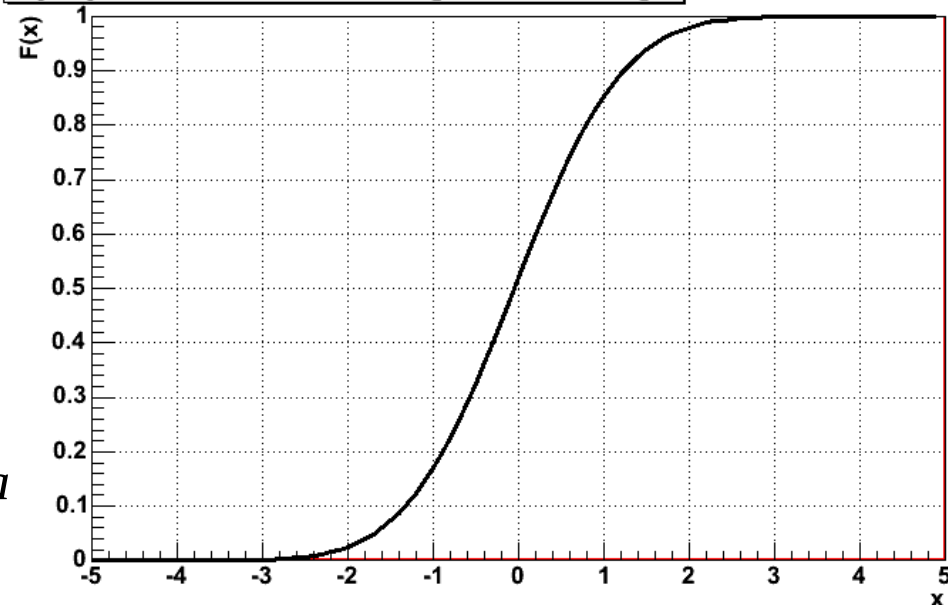
$$f(y) \equiv \phi(y) = \frac{1}{\sqrt{2\pi}b} e^{-(y-a)^2/2b^2}$$

- średnia (przesunięcie): $E(Y) = \hat{y} = a$
- odch. std. (szerokość): $\sigma(Y) = b$

Rozkład normalny standardowy

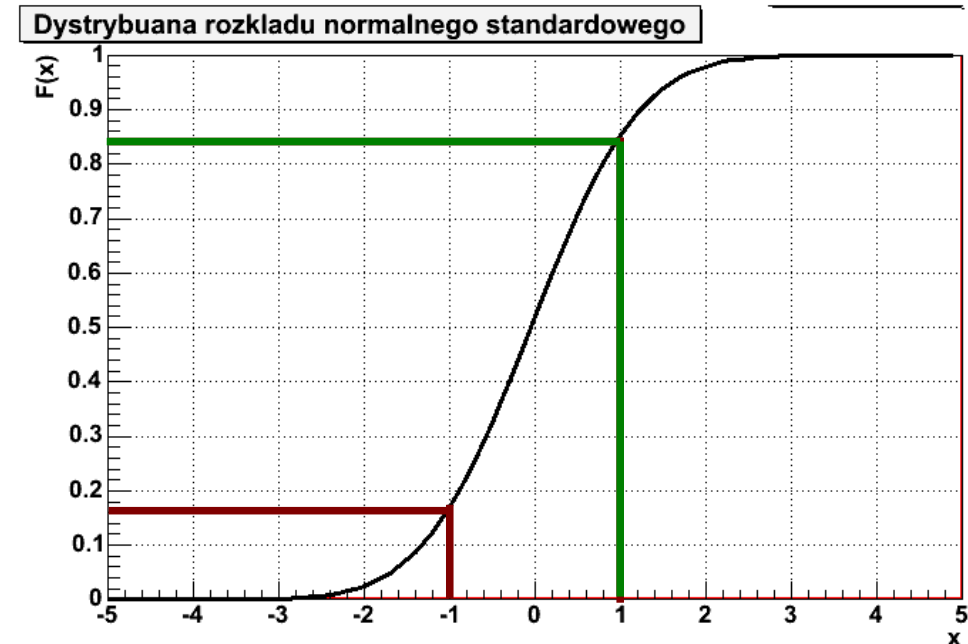
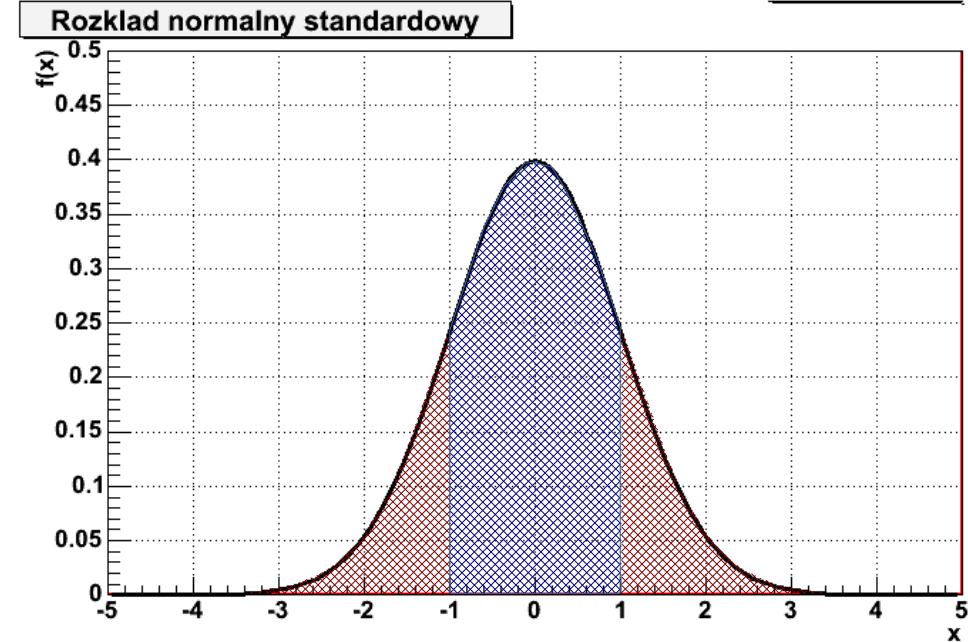


Dystrybuanta rozkładu normalnego standardowego



Rozkład normalny standardowy - własności

- Punkt przegięcia rozkładu:
 - **standardowego** $x=\pm 1$
 - **Gaussa** $x=a\pm b$
- Załóżmy, że znamy dystrybuantę:
$$F_0(x) \equiv \Phi_0(x) = P(X \leq x)$$
- Ze względu na asymetrię gęstości:
$$P(|X| > x) = 2 \Phi_0(-|x|) = 2(1 - \Phi_0(|x|))$$
- Analogicznie, wewnątrz przedziału $2x$:
$$P(|X| \leq x) = 2 \Phi_0(|x|) - 1$$
- Dystrybuantę r. norm. można uogólnić na r. Gaussa:
$$\Phi(y) = \Phi_0\left(\frac{x-a}{b}\right)$$



Rozkład normalny standardowy - własności

- Wtedy szczególnie interesujące jest obliczenie występowania zmiennej los. dla wielokrotności odchylenia standardowego:

$$P(|Y - a| \leq n\sigma) = 2\Phi_0\left(\frac{nb}{b}\right) - 1 = 2\Phi_0(n) - 1$$

- Otrzymamy wtedy:

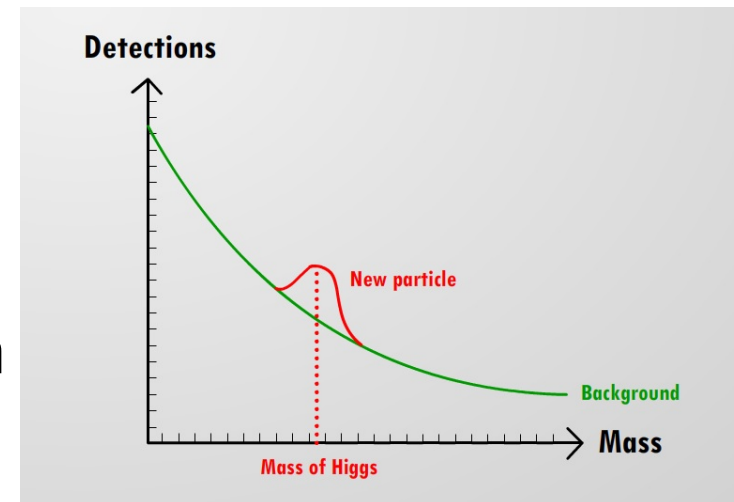
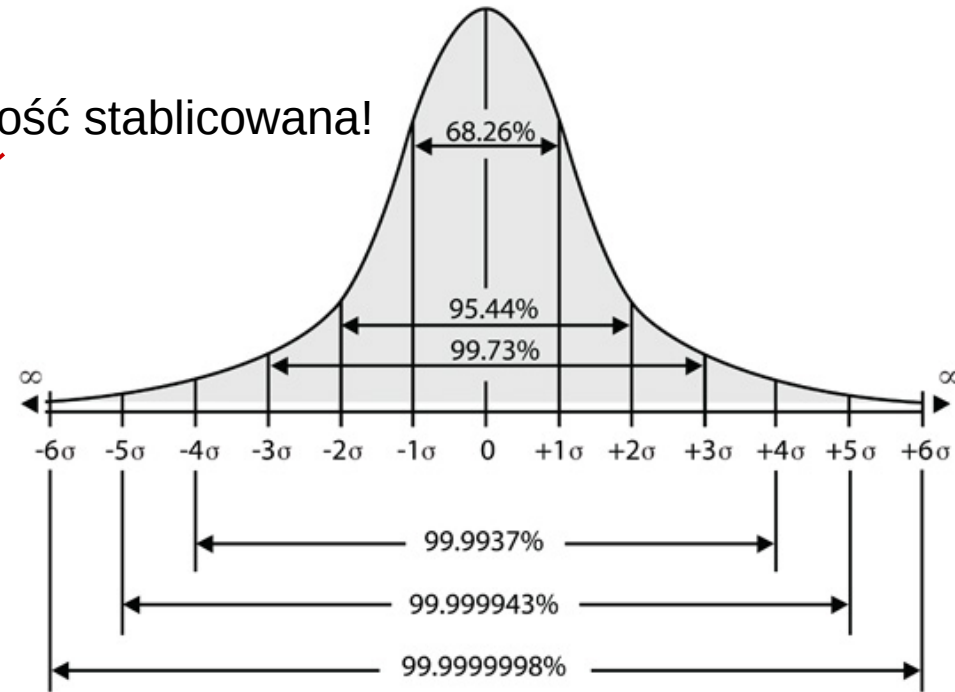
$$P(|Y - a| \leq \sigma) = 68,3\% \quad P(|Y - a| > \sigma) = 31,7\%$$

$$P(|Y - a| \leq 2\sigma) = 95,4\% \quad P(|Y - a| > 2\sigma) = 4,6\%$$

$$P(|Y - a| \leq 3\sigma) = 99,8\% \quad P(|Y - a| > 3\sigma) = 0,2\%$$

- Z Wykładu 1 pamiętamy, że **współczynnik rozszerzenia** niepewność typu A zwykle jest między 2 a 3 – tu widać dlaczego
- W nauce przez odchylenie standardowe określamy różnice w obserwowanym sygnale eksperymentalnym w stosunku do sytuacji, gdy efektu fizycznego nie ma

Wartość stabilizowana!





Centralne twierdzenie graniczne

Centralne twierdzenie graniczne

- Dlaczego rozkład normalny jest tak ważny w rachunku prawdopodobieństwa i statystyce?
- Mówi o tym **centralne twierdzenie graniczne** (*ang. central limit theorem*) – jedno z najważniejszych twierdzeń rachunku prawdopodobieństwa:

- jeżeli zmienne losowe X_i są zmiennymi niezależnymi o jednakowych wartościach średnich a i odchyleniach standardowych b , to **rozkład normalny** ma zmienna:

$$X = \lim_{n \rightarrow \infty} \sum_{i=1}^n X_i \quad E(X) = na, \quad \sigma^2(X) = nb^2$$

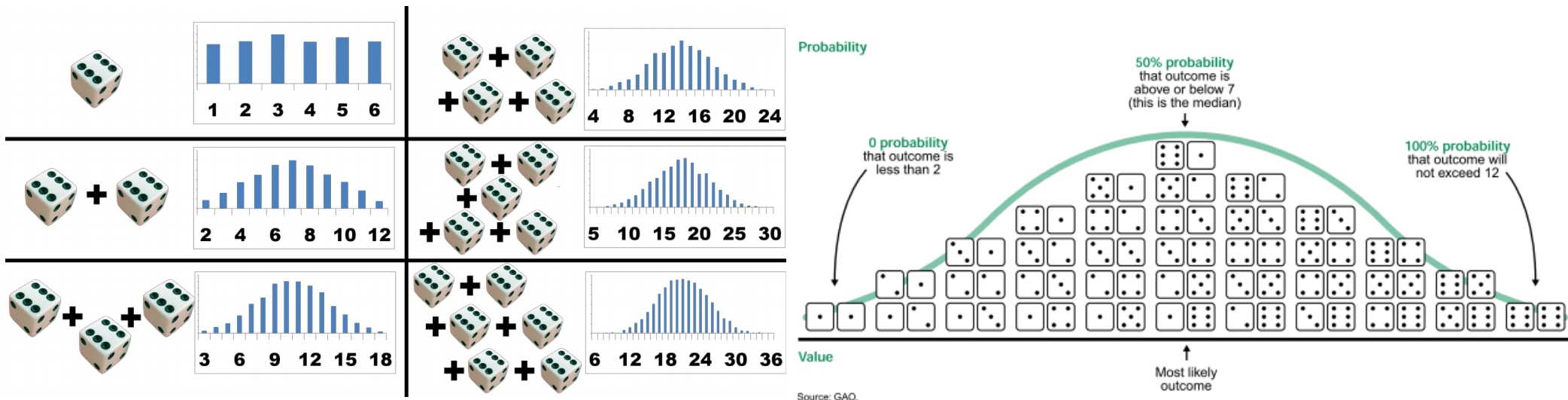
- ponadto, zmienna $\xi = \frac{1}{n} X = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n X_i$ ma rozkład normalny z:

$$E(\xi) = a, \quad \sigma^2(\xi) = b^2/n$$

- Innymi słowy – mając n niezależnych zmiennych o jednakowym (ale **dowolnym!**) rozkładzie, to ich suma dla dużych n zbiega do rozkładu normalnego

Centralne twierdzenie graniczne - przykład 1

- Wyobraźmy sobie eksperyment polegający na rzucie kostką (kostkami) i obserwowaniu całkowitej liczby oczek:
 - kolejne rzuty kostką (kostkami) są niezależne
 - jeśli rzucamy kostką jednokrotnie (albo 1 kostką), to prawdopodobieństwo uzyskania danej wartości jest jednakowe
 - jeśli rzucamy kostką dwukrotnie (albo 2 kostkami), to prawdopodobieństwo uzyskania sumy oczek nie jest już jednakowe
 - jeśli rzucimy kostką n -krotnie (n -kostkami) → rozkład normalny



Centralne twierdzenie graniczne - przykład 2

- Wykonajmy doświadczenie rzutu monetą – zmienne losowe X_i przybierają wartości 1 (orzeł) i 0 (reszka) z prawdopodobieństwem p i $q=1-p$:

- **rozkład dwumianow** (n rzutów, k razy wypada wartość 1):

$$P(k) = P(X=k) = W_k^n = \binom{n}{k} p^k q^{n-k}$$

$$X = \sum_{i=1}^n X_i$$

$$E(X_i) = p, \quad \sigma^2(X_i) = pq$$

$$E(X) = np, \quad \sigma^2(X) = npq$$

- wprowadźmy teraz zmienną unormowaną U :

$$U = \sum_{i=1}^n \frac{X_i - p}{\sqrt{np(1-p)}} = \frac{1}{\sqrt{np(1-p)}} \left(\sum_{i=1}^n X_i - np \right) = \frac{1}{\sqrt{np(1-p)}} (X - np)$$

- wtedy rozkład prawdopodobieństwa przybiera postać:

$$P(k) = P(X=k) = W_k^n = P\left(U = \frac{k - np}{\sqrt{np(1-p)}} \right)$$

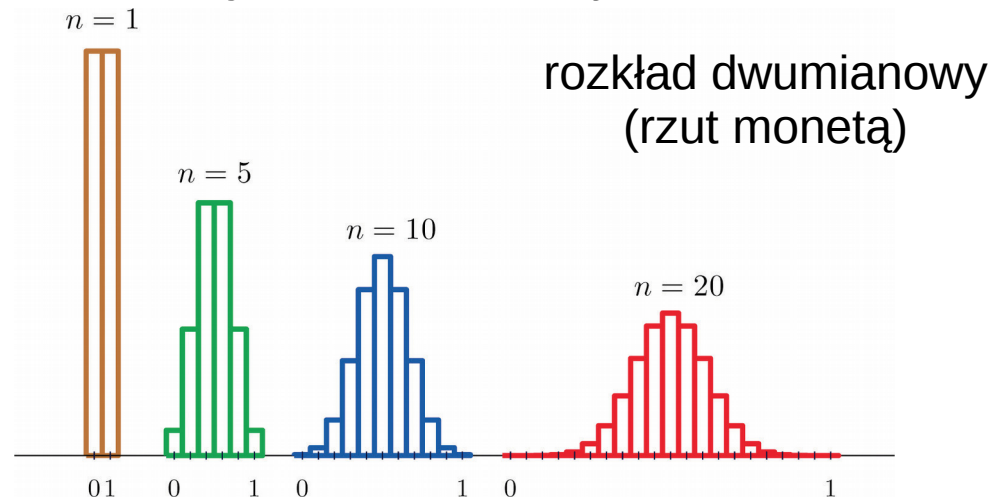
- zmienna losowa $U^{(n)}$ ma **rozkład normalny** dla $n \rightarrow \infty$:

$$P(a < U < b) = P\left(a < \frac{k - np}{\sqrt{np(1-p)}} < b \right) \xrightarrow{n \rightarrow \infty} \Phi_0(b) - \Phi_0(a)$$

**Twierdzenie
de Moivre'a-Laplace'a**

Centralne twierdzenie graniczne - przykład 2

- Przykład 1: wielokrotny rzut monetą



- Przykład 2:

Na linii produkcyjnej procesorów prawdopodobieństwo defektu (zmienna losowa X) jest $p=0.02$. System online monitorowania jakości sprawdza $n=1000$ procesorów dziennie i zlicza liczbę defektów. Wiemy, że:

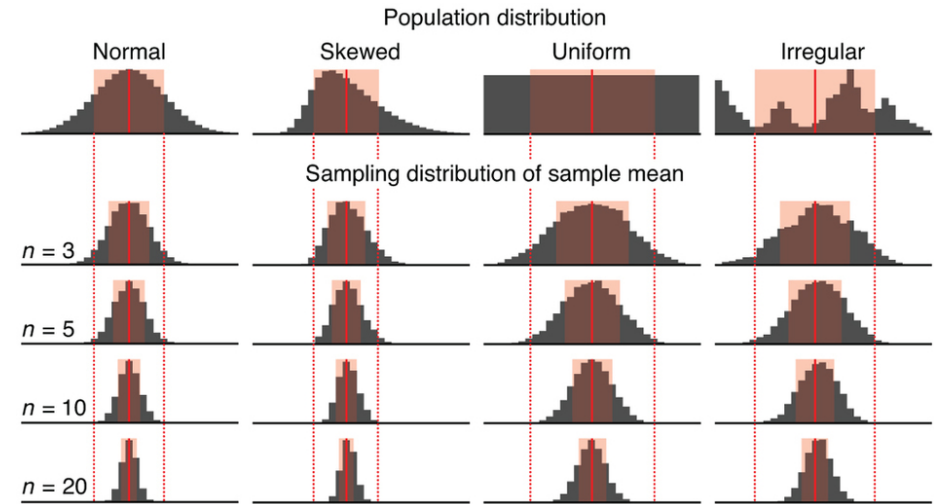
$$E(X) = np = 1000 \cdot 0.02 = 20$$

$$\sigma(X) = \sqrt{np(1-p)} = \sqrt{1000 \cdot 0.02 \cdot 0.98} = 4.427$$

- Z uwagi na tw. de Moivre'a - Laplace'a zmienna losowa X podlega w przybliżeniu rozkładowi Gaussa o wartości oczekiwanej 20 oraz odchyleniu standardowemu 4.427. Przykładowo, pozwala to nam sprawdzić, czy dana partia nie "odbiega" zbyt daleko od 20 (mieści się w dopuszczalnym zakresie 'sigma')

Centralne twierdzenie graniczne - przykład 2

- Ogólniej: CTG mówi, że rozkład średniej z pewnej próby losowej z dowolnego rozkładu będzie dążył do rozkładu normalnego:
 - mamy dowolny rozkład losowy pewnej zmiennej losowej o wart. ocz. μ i odch. std. σ
 - pobieramy próbę losową (wybieramy n elementów z rozkładu)
 - liczymy średnią \bar{X} z próby losowej
 - rozkład normalny ma:



$$P\left(a < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < b\right) \xrightarrow{n \rightarrow \infty} \Phi_0(b) - \Phi_0(a)$$

**Centralne Twierdzenie
Graniczne**

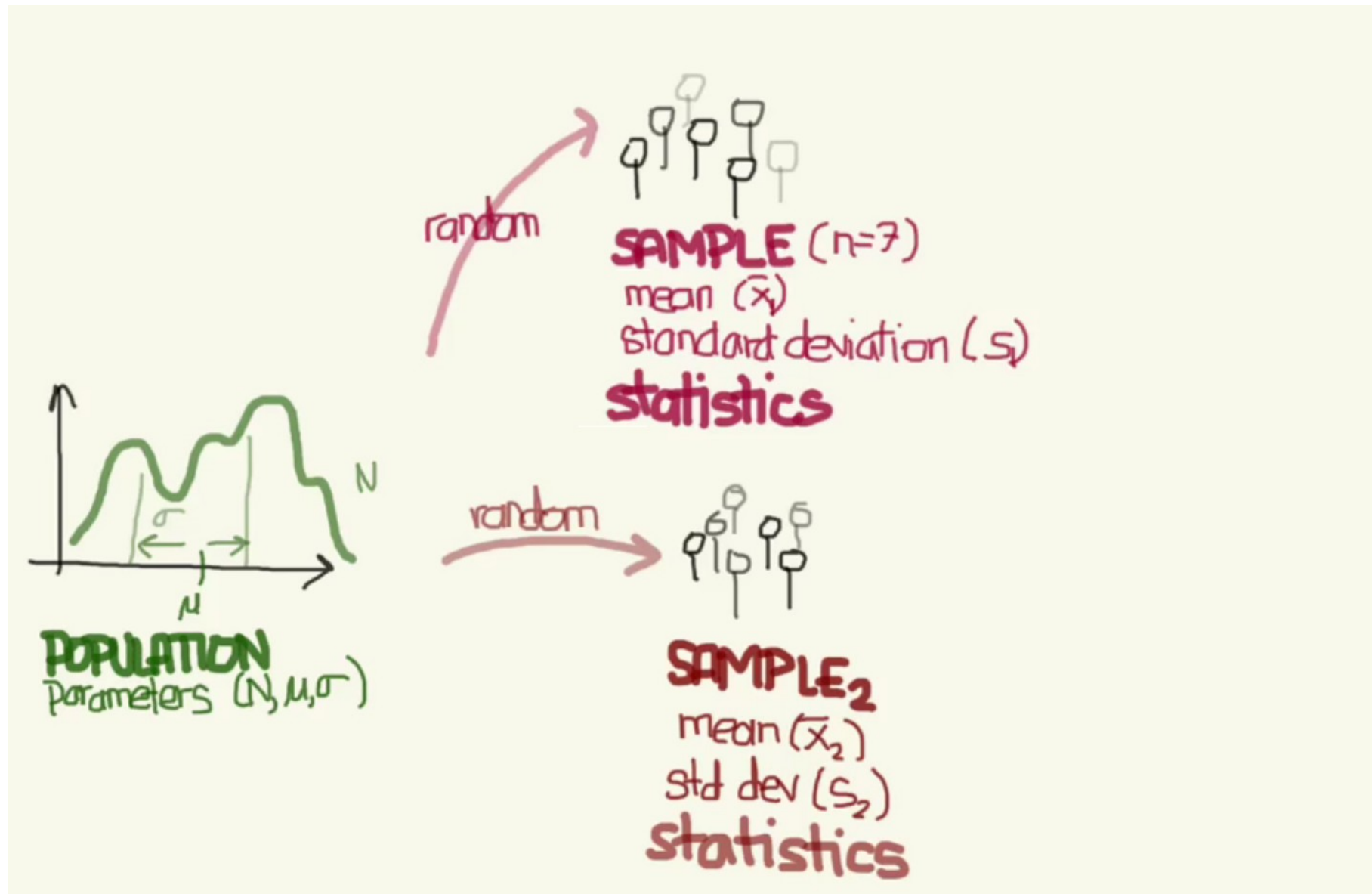
Centralne twierdzenie graniczne - przykład 2

- Przykład:

- wyobraźmy sobie, że szacujemy wzrost w **populacji** ośmioletnich dzieci w Polsce. Rozkład populacji ma parametry: μ, σ
- wybieramy losowo 100 8-latków i liczymy średnią wartość z próby losowej – \bar{X}_1
- nasz kolega wykonuje analogiczne doświadczenie → dostaje inny wynik – \bar{X}_2
- zaczynamy więc pracować razem, znowu wybieramy 100 8-latków i dostajemy trzeci wynik – \bar{X}_3
- ale przecież jest tylko **jeden prawdziwy** średni wzrost 8-latek w całej populacji!
- ponieważ **średnia z próby jest również zmienną losową**, możemy wykonać wielokrotnie próbę losową i dostać wiele średnich → **otrzymujemy rozkład wartości średniej z próby**
- jeśli mamy dużo prób losowych → **rozkład wartości średniej z prób dąży do rozkładu normalnego (CTG):** $N(\mu, \sigma/\sqrt{n})$

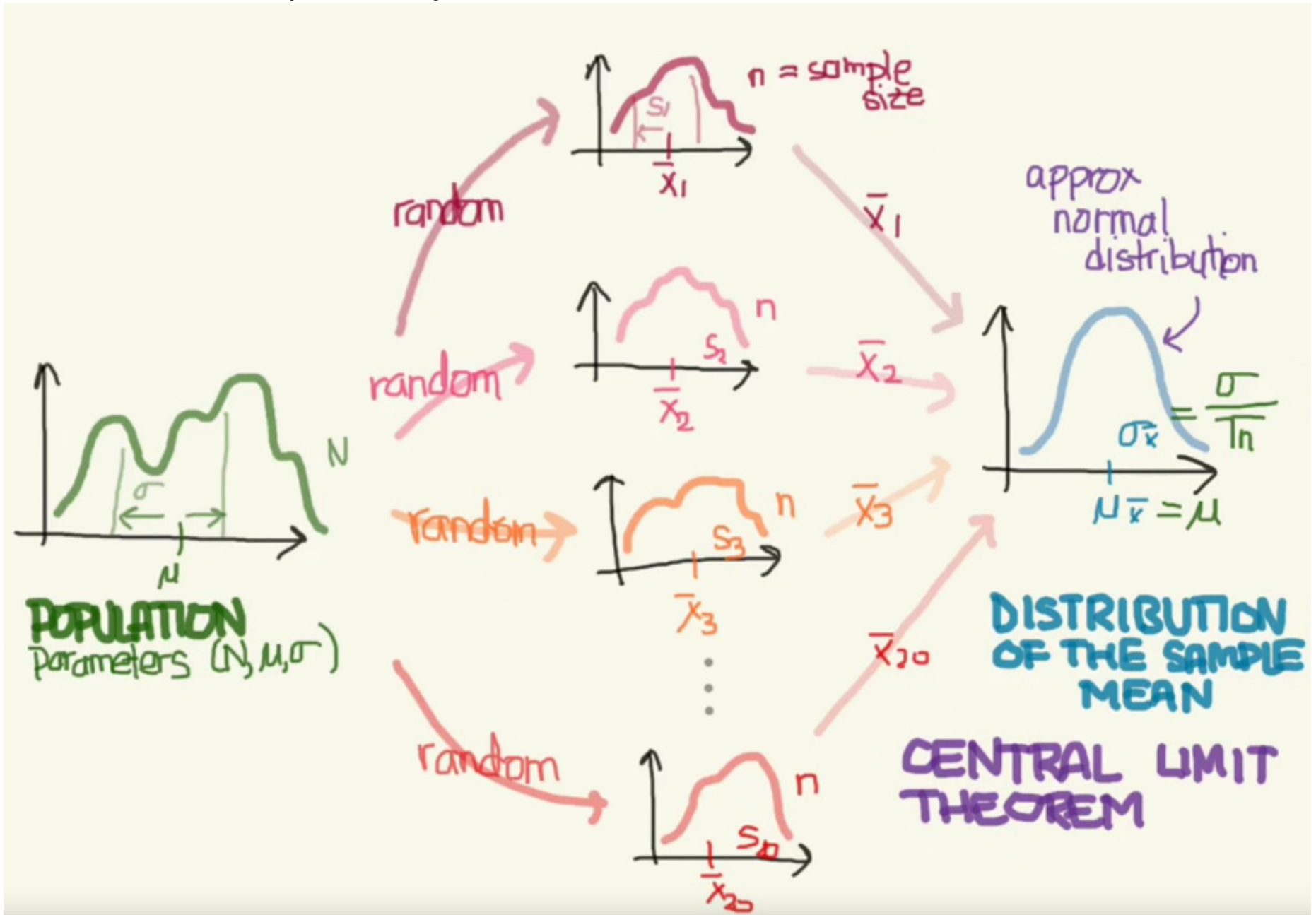
Centralne twierdzenie graniczne - przykład 2

<https://www.youtube.com/watch?v=3SKwerKHbRk>



Centralne twierdzenie graniczne - przykład 2

<https://www.youtube.com/watch?v=3SKwerKHbRk>



Centralne twierdzenie graniczne - przykład 2

- Przykład:
 - wyobraźmy sobie, że szacujemy wzrost w **populacji** ośmioletnich dzieci w Polsce. Rozkład populacji ma parametry: μ, σ
 - wybieramy losowo 100 8-latków i liczymy średnią wartość z próby losowej – \bar{X}_1
 - nasz kolega wykonuje analogiczne doświadczenie → dostaje inny wynik – \bar{X}_2
 - zaczynamy więc pracować razem, znowu wybieramy 100 8-latków i dostajemy trzeci wynik – \bar{X}_3
 - ale przecież jest tylko **jeden prawdziwy** średni wzrost 8-latek w całej populacji!
 - ponieważ **średnia z próby jest również zmienną losową**, możemy wykonać wielokrotnie próbę losową i dostać wiele średnich → **otrzymujemy rozkład wartości średniej z próby**
 - jeśli mamy dużo prób losowych → **rozkład wartości średniej z prób dąży do rozkładu normalnego (CTG):** $N(\mu, \sigma/\sqrt{n})$



Model Laplace'a niepewności pomiarowych

Model Laplace'a niepewności pomiarowych

- W 1783 roku Laplace zaproponował następującą interpretację niepewności pomiarowych:
 - niech m_0 będzie wartością prawdziwą rozważanej wielkości mierzonej
 - pomiar jest zakłócany przez dużą liczbę n niezależnych czynników, z których każdy powoduje odchylenie rzędu ε
 - każde zakłócenie powoduje równe prawdopodobieństwo wywołania zmiany mierzonej wartości zarówno o $+\varepsilon$ i $-\varepsilon$
 - niepewność pomiaru jest zatem sumą poszczególnych zakłóceń
 - rozkład niepewności opisany jest w takim przypadku rozkładem dwumianowym

Liczba zakłóceń

Odchylenie od wartości prawdziwej

	-3ε	-2ε	$-\varepsilon$	0	ε	2ε	3ε
0				1			
1			$\frac{1}{2}$		$\frac{1}{2}$		
2		$\frac{1}{4}$		$\frac{1}{2}$		$\frac{1}{4}$	
3	$\frac{1}{8}$		$\frac{3}{8}$		$\frac{3}{8}$		$\frac{1}{8}$

Model Laplace'a niepewności pomiarowych

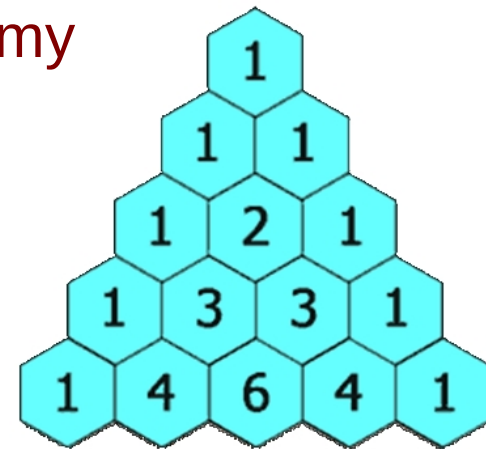
- przy braku zakłóceń prawdopodobieństwo uzyskania m_0 będzie oczywiście wynosić 1
- przy jednym zakłóceniu prawdopodobieństwo dzieli się po równo na dwie możliwości $m_0+\epsilon$ oraz $m_0-\epsilon$
- tak samo się dzieje przy każdym kolejnym zakłóceniu
- oczywiście, prawdopodobieństwa prowadzące do tego samego wyniku pomiarowego się sumują
- jeśli $p=q=1/2$, to model zachowuje się identycznie jak tzw. trójkąt Pascala (obrazek po prawej)

- analogicznie jak w przykładzie 2 z CTG, wprowadzamy zmienną standardową:

$$U = \frac{2 \left(\sum_{i=1}^{\infty} \epsilon X_i - n \frac{\epsilon}{2} \right)}{\sqrt{n} \epsilon}$$

- w granicy $n \rightarrow \infty$ rozkład jest normalny z wartością oczekiwaną 0 i odchyleniem standardowym $\sqrt{n} \epsilon / 2$

- **Wniosek: niepewności opisane rozkładem Gaussa są wynikiem wielu małych niezależnych zaburzeń**





Funkcja charakterystyczna rozkładu

Funkcja charakterystyczna rozkładu

- Dotychczas zajmowaliśmy się tylko zmiennymi losowymi rzeczywistymi (każdej realizacji zdarzenia losowego można przypisać liczbę rzeczywistą – zmienną losową)
- Definicję można uogólnić na zmienne losowe zespolone, składające się z dwóch zmiennych losowych rzeczywistych: $Z = X + iY$
- Wartość oczekiwana (z własności wart. ocz.): $E(Z) = E(X) + i \cdot E(Y)$
- Analogicznie, zmienne losowe zespolone są niezależne, jeżeli odpowiednio ich części rzeczywiste i urojone są niezależne
- **Po co nam to wszystko? Do zdefiniowania funkcji charakterystycznej rozkładu:**
 - X jest zmienną losową rzeczywistą o rozkładzie $f(x)$ i dystrybuancie $F(x)$
 - **Funkcja charakterystyczna** zdefiniowana jest jako wartość oczekiwana:

$$\phi(t) = E(\exp(itX))$$

Funkcja charakterystyczna rozkładu

- zatem dla ciągłej zmiennej losowej jest to transformata Fouriera:

$$\phi(t) = \int_{-\infty}^{\infty} \exp(itx) f(x) dx$$

- dla rozkładów dyskretnych: $\phi(t) = \sum_{i=1}^n \exp(itx_i) P(X = x_i)$

- jeśli zdefiniujemy momenty: $\lambda_n = E(X^n) = \int_{-\infty}^{\infty} x^n f(x) dx$

- widać, że można je otrzymać przez n -krotne różniczkowanie funkcji charakterystycznej w punkcie $t=0$:

$$\phi^{(n)}(t) = \frac{d^n \phi(t)}{dt^n} = i^n \int_{-\infty}^{\infty} x^n \exp(itx) f(x) dx \Rightarrow \phi^{(n)}(0) = i^n \lambda_n$$

- dla zmiennej losowej przesuniętej o wartość oczekiwaną: $Y = X - \hat{x}$

- funkcja charakterystyczna zmiennej Y dana jest jako:

$$\phi_y(t) = \int_{-\infty}^{\infty} \exp(it(x - \hat{x})) f(x) dx = \phi(t) \exp(-it \hat{x}) \quad \phi(t) = \int_{-\infty}^{\infty} \exp(itx) f(x) dx$$

- wówczas n -ta pochodna związana jest z momentem zmiennej X względem wartości oczekiwanej:

$$\phi_y^{(n)}(0) = i^n \mu_n = i^n E((X - \hat{x})^n)$$

- w szczególności wariancja: $\sigma^2(X) = -\phi_y''(0)$

Funkcja charakterystyczna rozkładu

- odwracając transformatę Fouriera możemy z funkcji charakterystycznej otrzymać funkcję gęstości:

$$f(x) = \frac{1}{2} \int_{-\infty}^{\infty} \exp(-itx) \phi(t) dt$$

- istnieje ściśle **jednoznaczny** związek między dystrybuantą a funkcją charakterystyczną, nawet wtedy, gdy mamy do czynienia z rozkładem dyskretnym, wtedy:

$$F(b) - F(a) = \frac{i}{2\pi} \int_{-\infty}^{\infty} \frac{\exp(itb) - \exp(ita)}{t} \phi(t) dt$$

- **funkcji charakterystycznej i dystrybuanty można używać zamiennie i przechodzić z jednej do drugiej w miarę potrzeb**
- funkcja charakterystyczna sumy dwóch niezależnych zmiennych:

$$\begin{aligned} W = X + Y \Rightarrow \phi_w(t) &= E(\exp(it(X+Y))) = E(\exp(itX) \exp(itY)) = \\ &= E(\exp(itX)) E(\exp(itY)) = \phi_x(t) \phi_y(t) \end{aligned}$$

Funkcja charakterystyczna rozkładu

- Przykładowe własności funkcji charakterystycznej dla wybranych rozkładów:

- rozkład Poissona: $\phi(t) = \exp(\lambda(e^{it} - 1))$

- suma rozkładów Poissona jest również rozkładem Poissona:

$$\phi(t) = \exp((\lambda_1 + \lambda_2)(e^{it} - 1))$$

- rozkład normalny: $\phi(t) = \exp(it \hat{x}) \exp\left(\frac{-\sigma^2 t^2}{2}\right)$

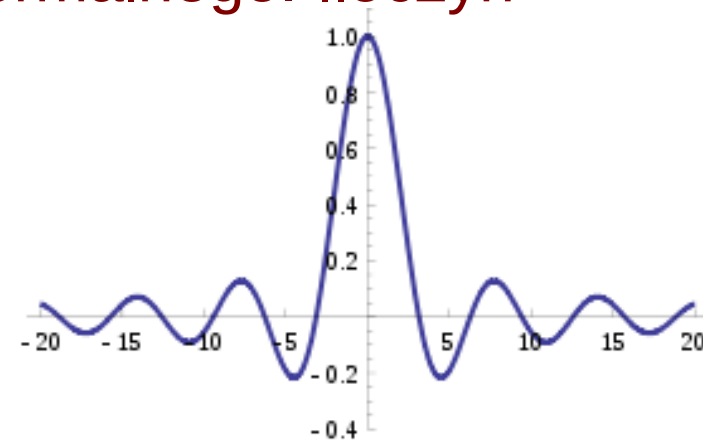
- jeżeli $a=0$, wówczas funkcja charakterystyczna o średniej równej 0 ma postać (z dokładnością do czynnika normalizacyjnego) gęstości prawdopodobieństwa rozkładu normalnego. Iloczyn wariancji obu rozkładów jest równy 1

- rozkład jednorodny: $f(x) = \frac{1}{b-a}, a \leq x \leq b$

$$\phi(t) = \frac{2}{(b-a)t} \sin\left(\frac{1}{2}(b-a)t\right) e^{i(a+b)t/2}$$

- suma rozkładów Gaussa:

$$\begin{aligned} \phi_u(t) &= \phi_x \cdot \phi_y = \exp(it \hat{x}) \exp(-\sigma_x^2 t^2 / 2) \exp(it \hat{y}) \exp(-\sigma_y^2 t^2 / 2) \\ &= \exp(it(\hat{x} + \hat{y})) \exp(-(\sigma_x^2 + \sigma_y^2) t^2 / 2) \end{aligned}$$





Wielowymiarowy rozkład Gausa

Wielowymiarowy rozkład Gaussa

- Rozważmy wektor zmiennych losowych: $\mathbf{X} = (X_1, X_2, \dots, X_n)$
- Gęstość prawdopodobieństwa wielowymiarowego rozkładu normalnego:
$$\phi(\mathbf{x}) = k \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{a})^T B(\mathbf{x} - \mathbf{a})\right) = k \exp\left(-\frac{1}{2}g(\mathbf{x})\right) \quad g(\mathbf{x}) = (\mathbf{x} - \mathbf{a})^T B(\mathbf{x} - \mathbf{a})$$
- gdzie \mathbf{a} jest n -wymiarowym wektorem wart. oczekiwanych, natomiast B jest dodatnio określoną macierzą symetryczną o wymiarze $n \times n$
- Z symetrii rozkładu normalnego:
$$\phi(\mathbf{x}) = E(\mathbf{X} - \mathbf{a}) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (\mathbf{x} - \mathbf{a}) \phi(\mathbf{x}) dx_1 \dots dx_n = 0 \quad E(\mathbf{X}) = \mathbf{a}$$
- Jeśli zróżniczkujemy to wyrażenie względem \mathbf{a} :
$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} [\mathbf{I} - (\mathbf{x} - \mathbf{a})(\mathbf{x} - \mathbf{a})^T B] \phi(\mathbf{x}) dx_1 \dots dx_n = 0$$
- Z def. wartości oczekiwanej zawartość nawiasu kwadratowego wynosi 0:
$$0: \quad E(\mathbf{I} - (\mathbf{X} - \mathbf{a})(\mathbf{X} - \mathbf{a})^T B) = 0 \Rightarrow E((\mathbf{X} - \mathbf{a})(\mathbf{X} - \mathbf{a})^T) B = \mathbf{I}$$
- Czyli: $C = E((\mathbf{X} - \mathbf{a})(\mathbf{X} - \mathbf{a})^T) = B^{-1}$
- Macierz C jest macierzą kowariancji zmiennych losowych \mathbf{X}

Wielowymiarowy rozkład Gaussa

- Przedyskutujmy rozkład dwóch zmiennych: $X = (X_1, X_2)$

- Macierz C ma wtedy następującą postać:

$$C = B^{-1} = \begin{pmatrix} \sigma_1^2 & \text{cov}(X_1, X_2) \\ \text{cov}(X_1, X_2) & \sigma_2^2 \end{pmatrix}$$

- Odwracając macierz C otrzymamy:

$$B = \frac{1}{\sigma_1^2 \sigma_2^2 - \text{cov}(X_1, X_2)^2} \begin{pmatrix} \sigma_2^2 & -\text{cov}(X_1, X_2) \\ -\text{cov}(X_1, X_2) & \sigma_1^2 \end{pmatrix}$$

- W przypadku zmiennych niezależnych (kowariancje wynoszą 0):

$$B_0 = \begin{pmatrix} 1/\sigma_1^2 & 0 \\ 0 & 1/\sigma_2^2 \end{pmatrix}$$

- Wstawiając B_0 do ogólnego wzoru otrzymamy łączną gęstość dwóch niezależnych zmiennych losowych jako iloczyn dwóch rozkładów 1D:

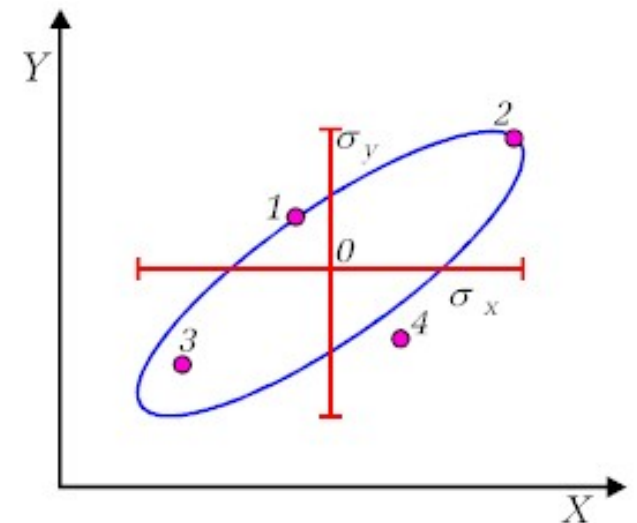
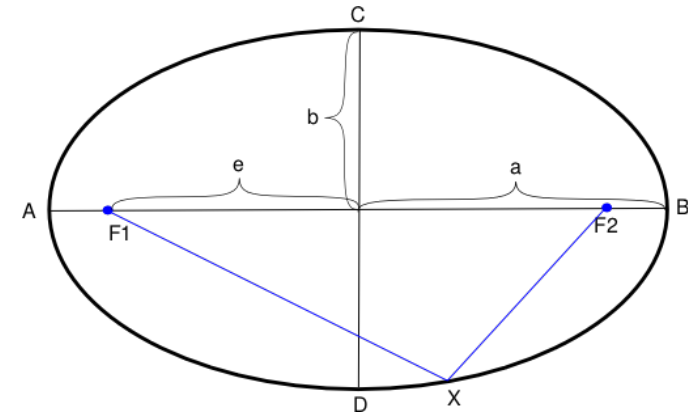
$$\phi(x_1, x_2) = k \exp\left(-\frac{1}{2} \frac{(x_1 - a_1)^2}{\sigma_1^2}\right) \exp\left(-\frac{1}{2} \frac{(x_2 - a_2)^2}{\sigma_2^2}\right), \quad k = \frac{1}{2\pi\sigma_1\sigma_2}$$

Wielowymiarowy rozkład Gaussa

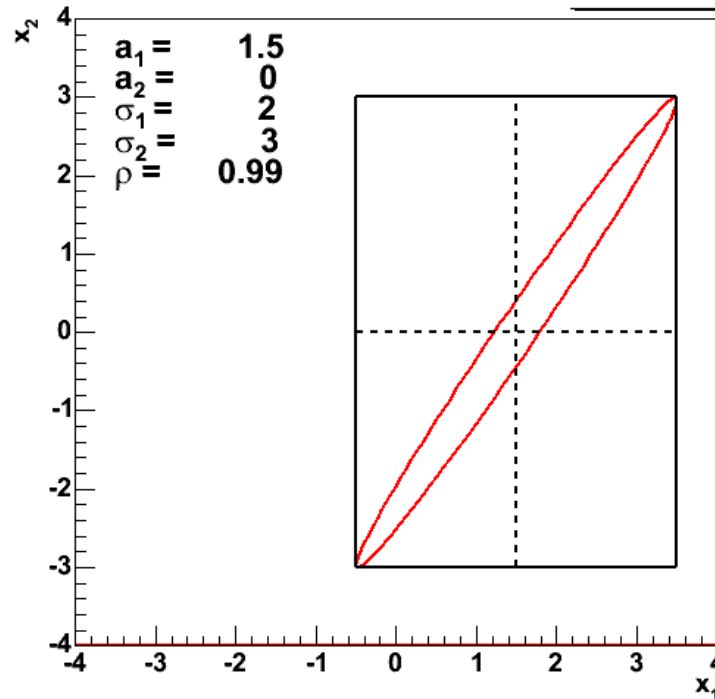
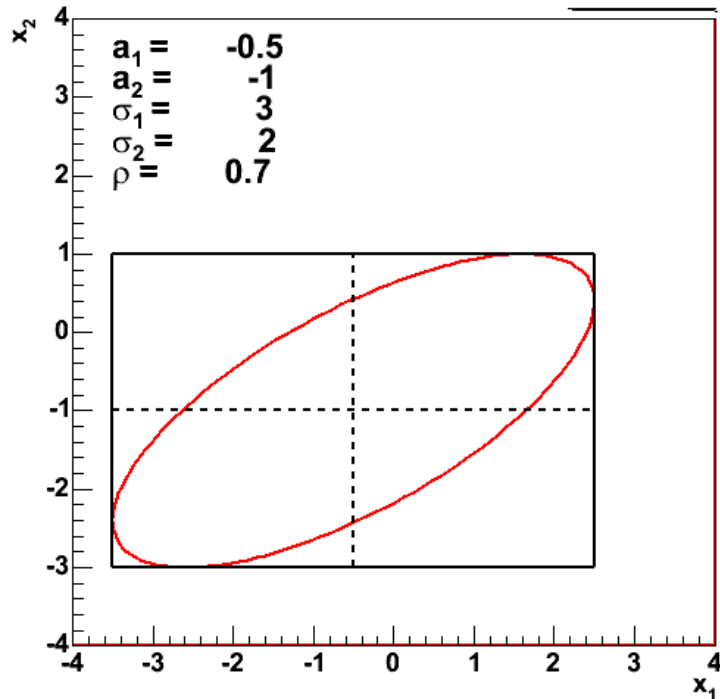
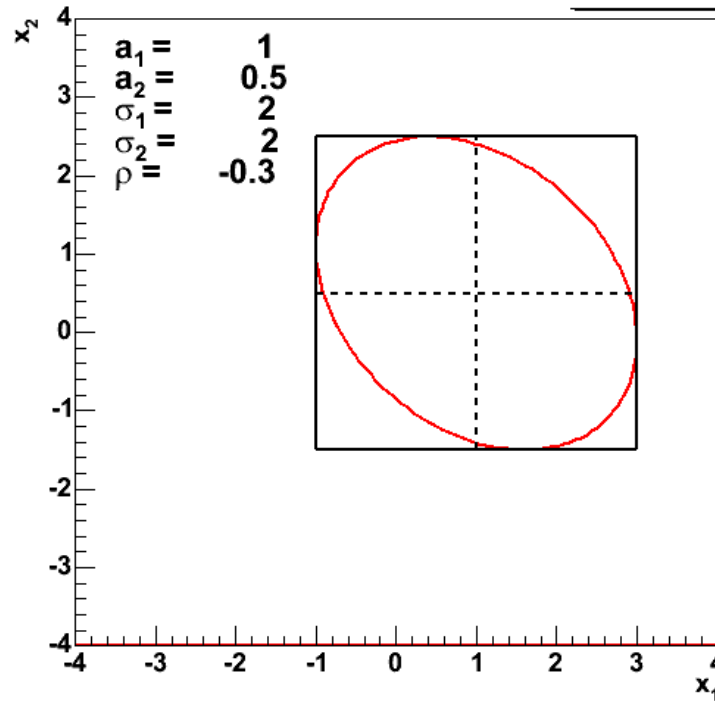
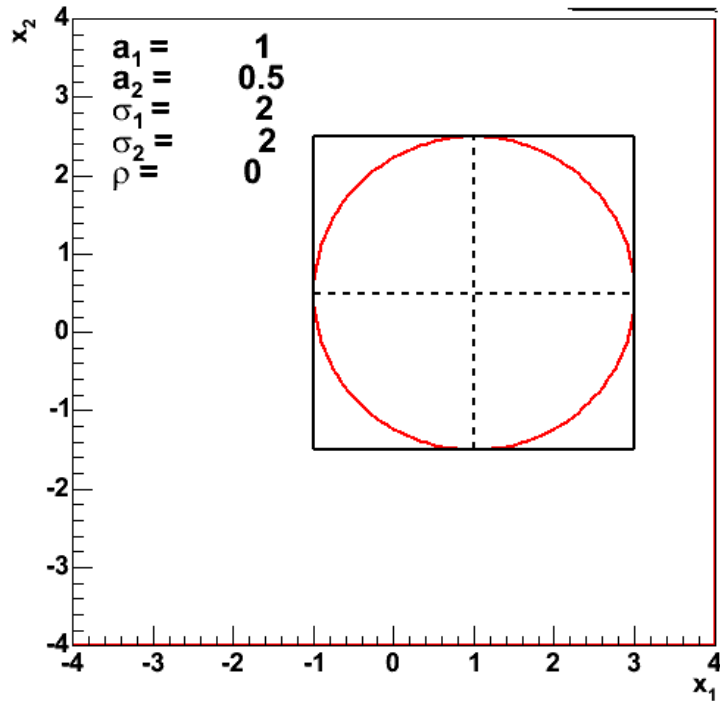
- Współczynnik k w ogólnym przypadku: $k = \left(\frac{\det B}{(2\pi)^n} \right)^{\frac{1}{2}}$
- Wprowadźmy teraz zmienne zredukowane: $U_i = \frac{X_i - a_i}{\sigma_i}, i=1,2$
- I współczynnik korelacji: $\rho = \frac{\text{cov}(X_1, X_2)}{\sigma_1 \sigma_2} = \text{cov}(U_1, U_2)$
- Wtedy gęstość prawdopodobieństwa:
$$\phi(u_1, u_2) = k \exp\left(-\frac{1}{2} \mathbf{u}^T B \mathbf{u}\right) = k \exp\left(-\frac{1}{2} g(\mathbf{u})\right), \quad B = \frac{1}{1-\rho^2} \begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix}$$
- Szukamy linii stałej gęstości prawdopodobieństwa poprzez przyrównanie wykładnika eksponensa do wartości stałej:
$$-\frac{1}{2} \mathbf{u}^T B \mathbf{u} = -\frac{1}{2} \frac{1}{(1-\rho^2)} (u_1^2 + u_2^2 + 2u_1 u_2 \rho) = -\frac{1}{2} g(\mathbf{u}) = \text{const}$$
- Jeśli na moment przyjmiemy, że $g(\mathbf{u})=1$ i wstawimy pierwotne zmienne x_1, x_2
- Otrzymamy równanie elipsy (**elipsy kowariancji**) o środku w (a_1, a_2) , której osie główne tworzą kąt α z osiami głównymi x_1, x_2 :
$$\frac{(x_1 - a_1)^2}{\sigma_1^2} - 2\rho \frac{x_1 - a_1}{\sigma_1} \frac{x_2 - a_2}{\sigma_2} + \frac{(x_2 - a_2)^2}{\sigma_2^2} = 1 - \rho^2 \quad \text{tg } 2\alpha = \frac{2\rho\sigma_1\sigma_2}{\sigma_1^2 - \sigma_2^2}$$

Elipsa kowariancji

- Wzór ogólny na elipsę (nieobróconą): $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$
- a i b to wielka i mała półoś elipsy
- W naszym przypadku elipsa jest dodatkowo obrócona o kąt α , zależny od wsp. korelacji
- Elipsa kowariancji zawsze zdefiniowana jest wewnątrz prostokąta środka w (a_1, a_2) oraz bokach σ_1, σ_2
- Jeżeli współczynnik korelacji wynosi $\rho = \pm 1$ to elipsa kowariancji degeneruje się do prostej pokrywającej się z jedną z przekątnych prostokąta
- Elipsa kowariancji jest linią stałego prawdopodob.
- Rysunek po prawej:
 - punkty 1 i 2 mają takie samo prawdopodobieństwo
 - prawdopodobieństwo punktu 3 jest większe niż 4



Elipsa kowariancji



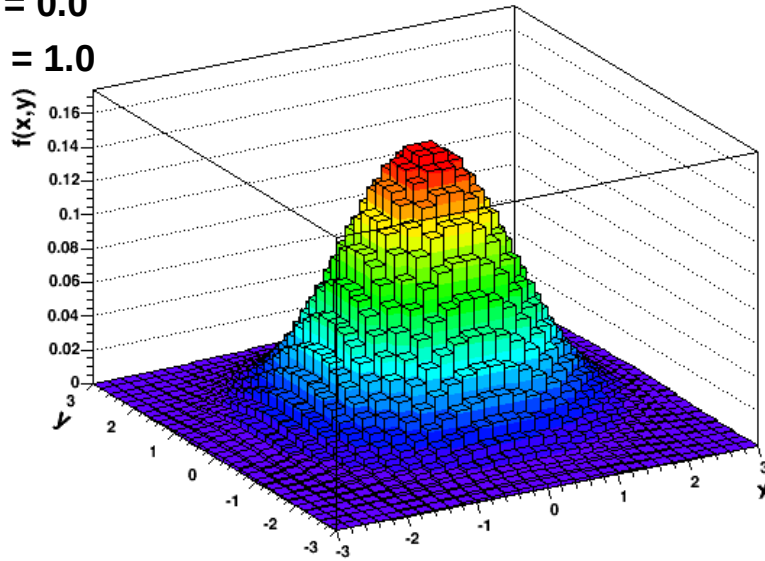
- Korelacja wydłuża i obraca elipsę
- Rozmiar elipsy zależy od wariancji
- Elipsa kowariancji zawiera **pełną** informację o macierzy kowariancji (w przypadku 2D)
- W 3D – elipsoida kowariancji
- W nD – hiperelipsoida kowariancji

Elipsa kowariancji

$$\text{cov}(X_1, X_2) = 0.0$$

$$a_1 = a_2 = 0.0$$

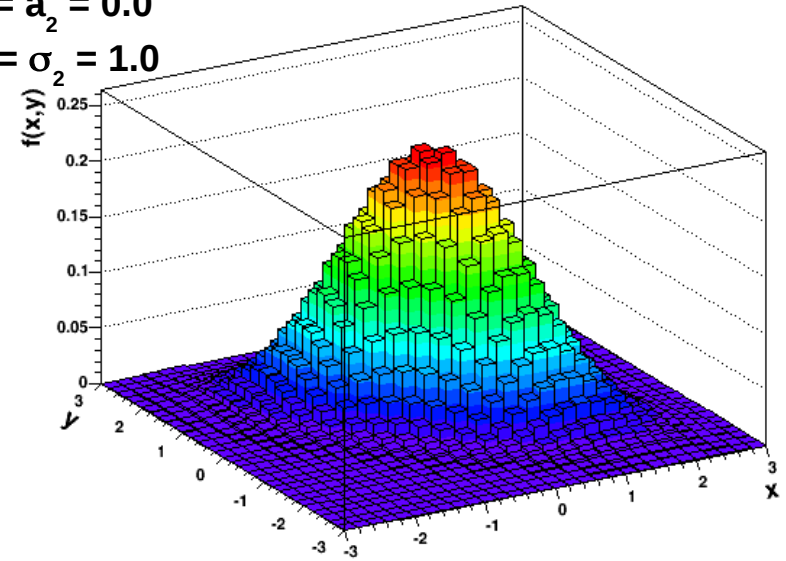
$$\sigma_1 = \sigma_2 = 1.0$$



$$\text{cov}(X_1, X_2) = 0.75$$

$$a_1 = a_2 = 0.0$$

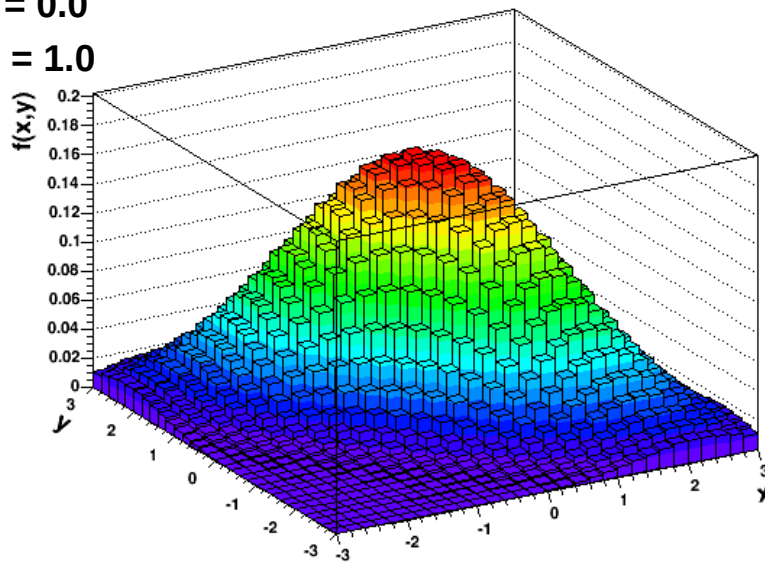
$$\sigma_1 = \sigma_2 = 1.0$$



$$\text{cov}(X_1, X_2) = 0.5$$

$$a_1 = a_2 = 0.0$$

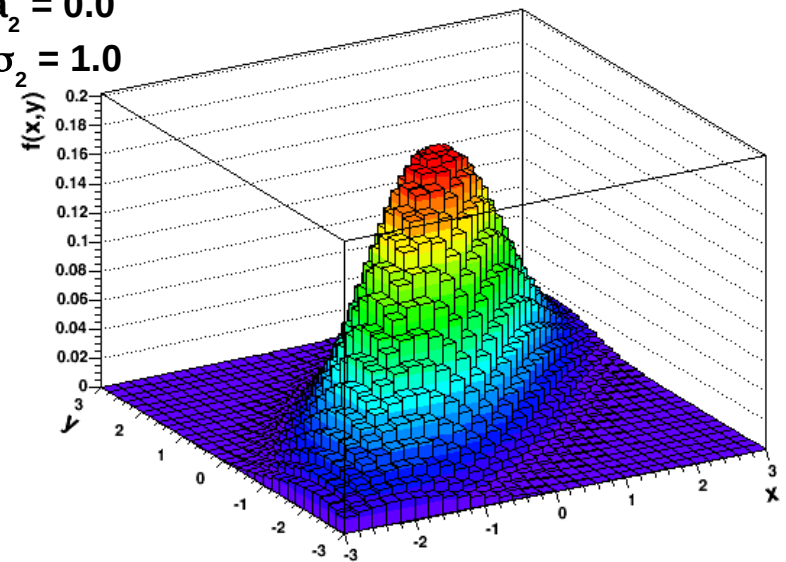
$$\sigma_1 = \sigma_2 = 1.0$$



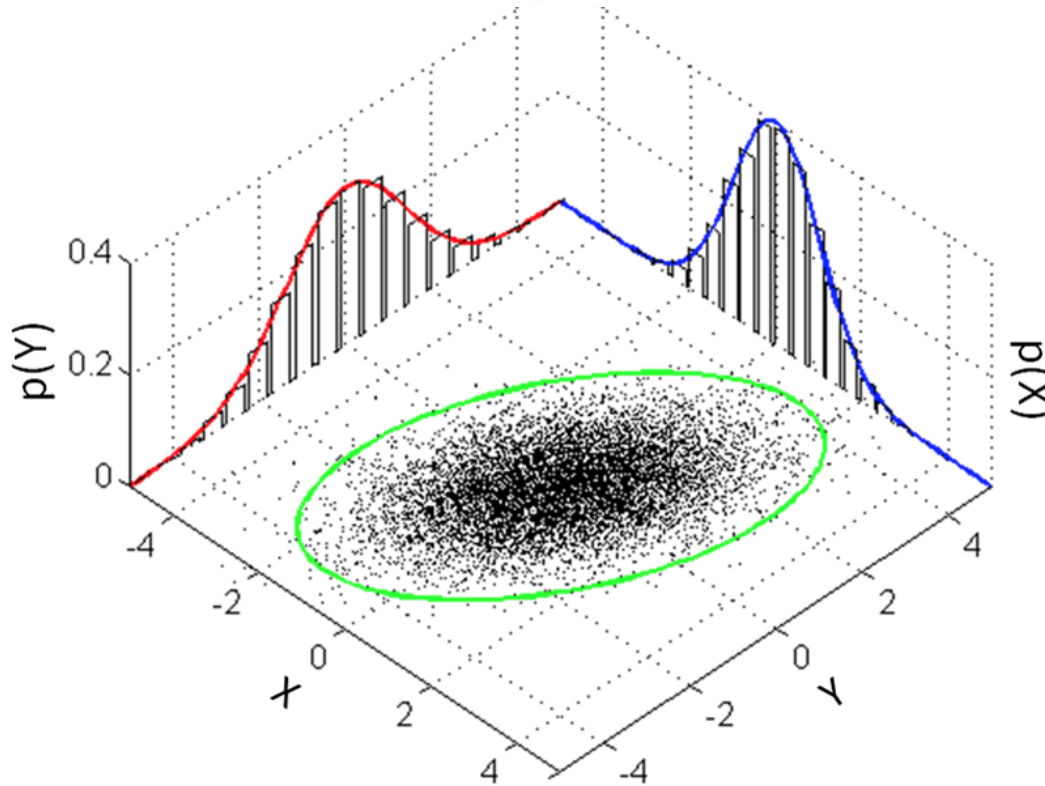
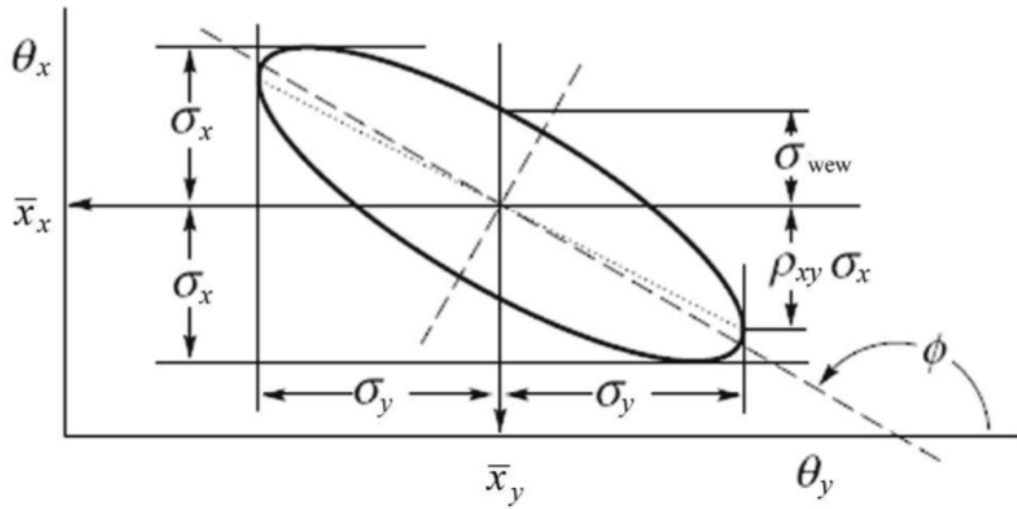
$$\text{cov}(X_1, X_2) = -0.5$$

$$a_1 = a_2 = 0.0$$

$$\sigma_1 = \sigma_2 = 1.0$$



Elipsa kowariancji



- Możemy wyznaczać również inne wartości: $g(\mathbf{u}) = \text{const} \neq 1$
- Każda elipsa kowariancji określa obszar prawdopodobieństwa analogicznie jak w przypadku 1D:

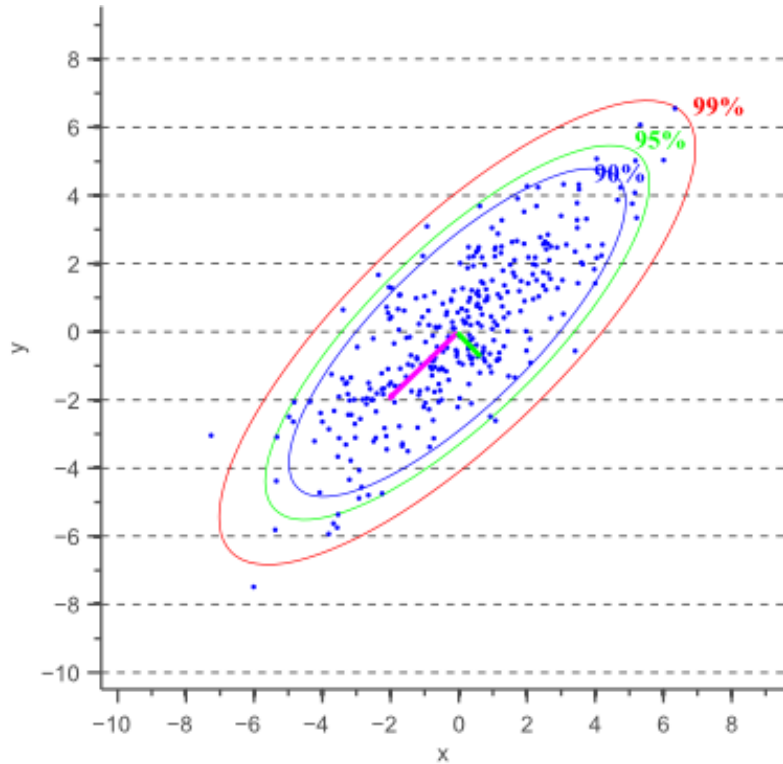
$$P(|X - a| \leq \sigma) = 68,3\%$$

- Wartość prawdopodobieństwa wewnątrz elipsy zależy od ilości wymiarów, w 2D (dla elipsy 1σ):

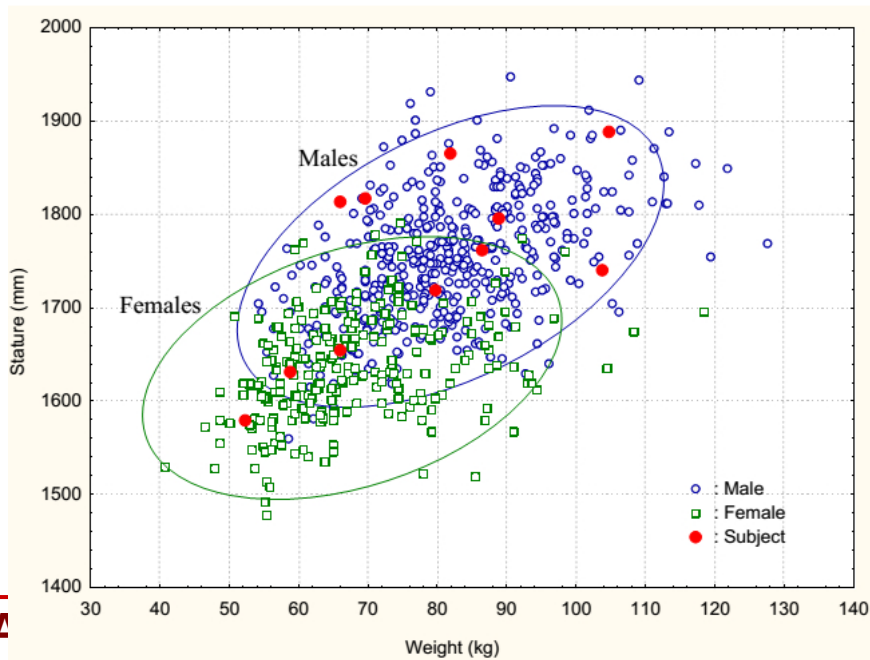
$$P = 39,3\%$$

- Inne linie stałego prawdopodobieństwa (elipsy) wyznaczają inne wartości prawdopodobieństwa

Elipsa kowariancji - wykorzystanie

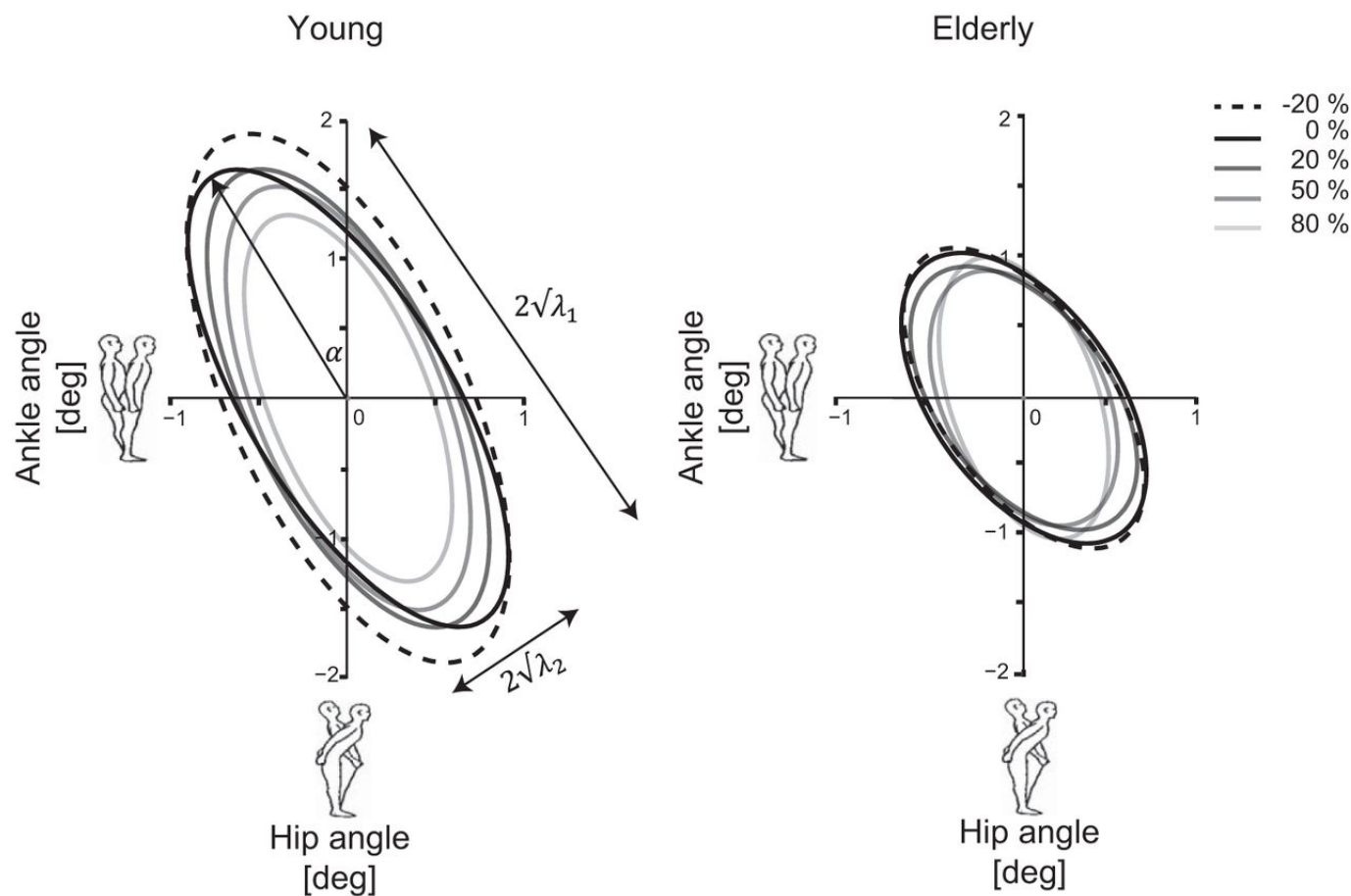


- Elipsy stałego prawdopodobieństwa mają ścisłe powiązanie z przedziałami ufności (o nich w przyszłości)
- Np. najczęściej określa się elipsę zawierającą prawdopodobieństwo 95% z wyników danych
- Przykład – korelacja wzrostu (*stature*) - wagi (*weight*) człowieka
- Analizy tego typu (dwóch lub więcej zmiennych jednocześnie) nazywamy analizą (statystyką) wielowymiarową (*multivariate analysis, statistics*)

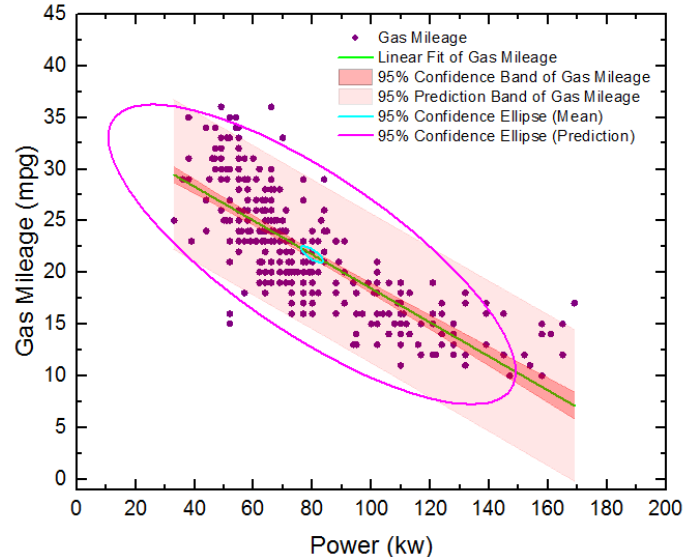
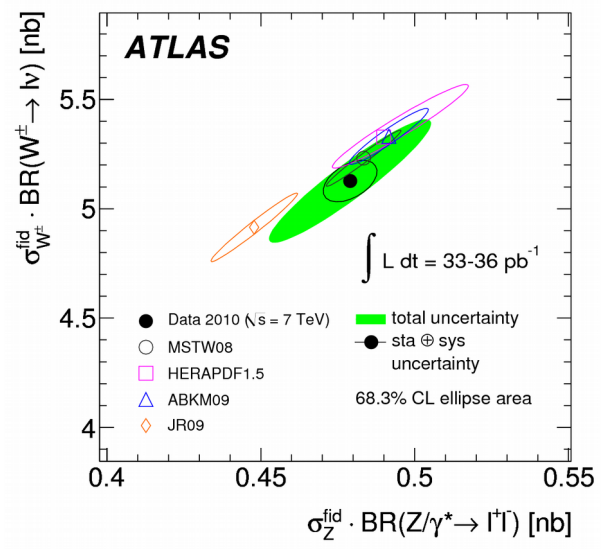
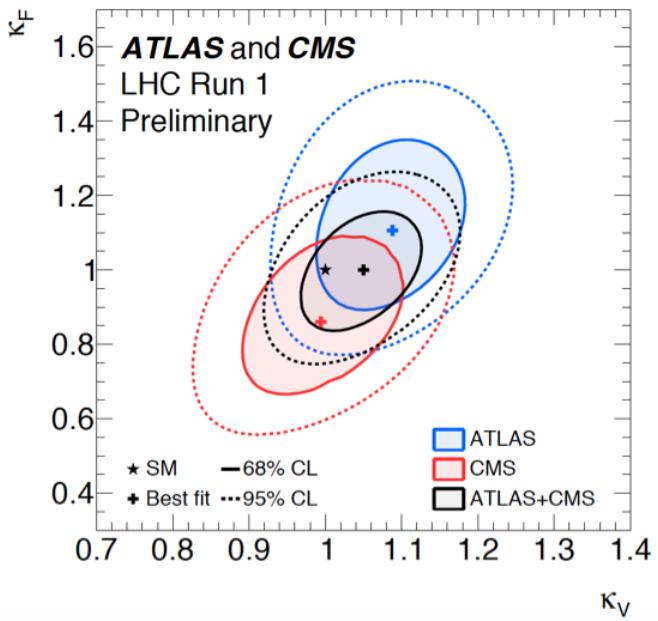
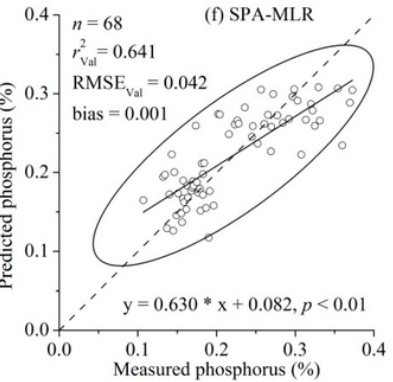
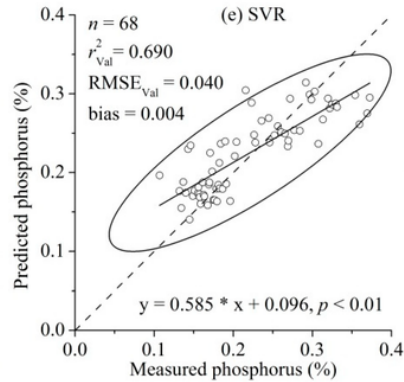
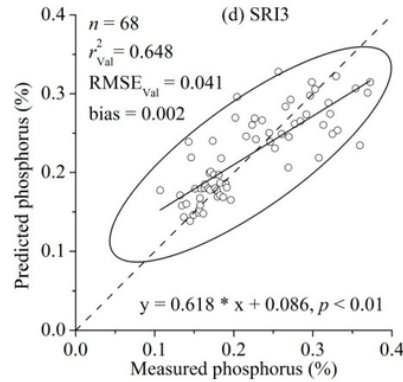
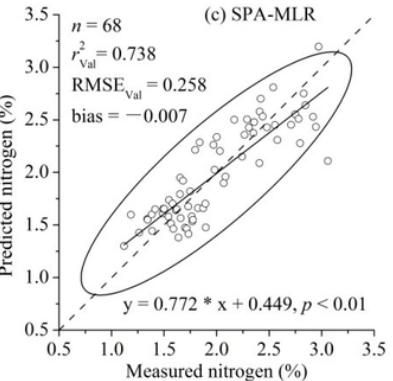
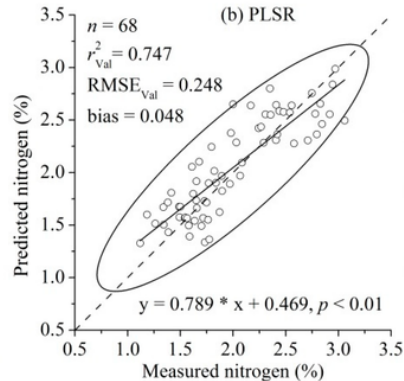
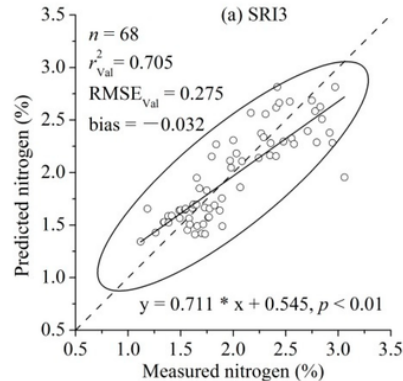
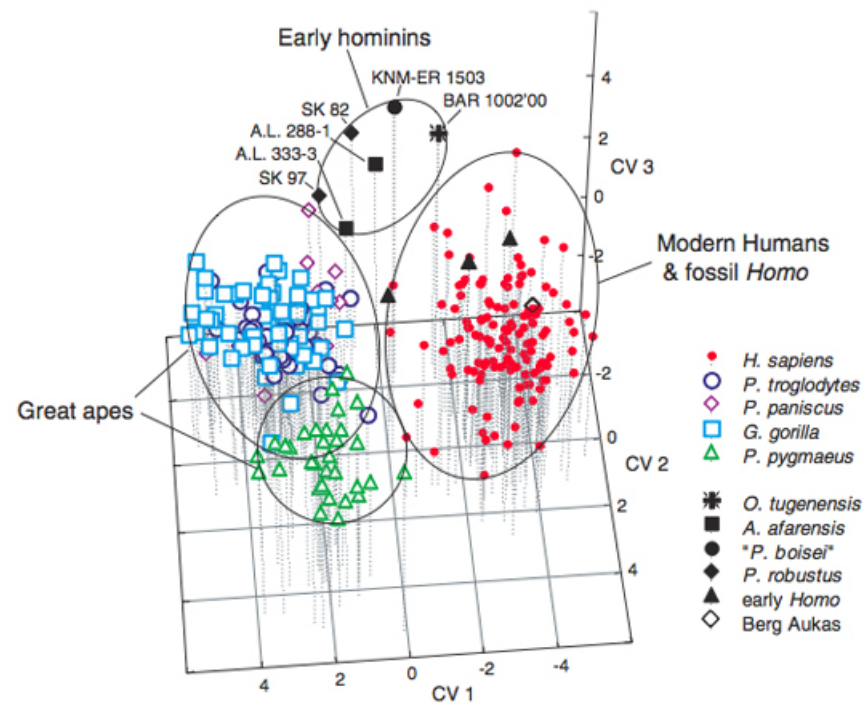


Elipsa kowariancji - wykorzystanie

- Zależność kąta zgięcia w kostce od kąta zgięcia w biodrze u młodszych i starszych osób



Elipsa kowariancji - wykorzystanie





KONIEC