Poprzednim razem

Wielowymiarowe zmienne losowe

- Łączny rozkład prawdopodobieństwa $P_{X,Y}(x_i, y_j)$
- Brzegowy rozkład prawdopodobieństwa $P_X(x_i) = \sum_j P_{X,Y}(x_i, y_j)$
- Łączna dystrybuanta $F_{X,Y}(x,y) = \sum_{x_i \le x} \sum_{y_j \le y} P_{X,Y}(x_i, y_j)$
- Łączna funkcja gęstości prawdopodobieństwa

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y} \qquad F_{X,Y}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(u,v)\, du\, dv$$

- Brzegowa funkcja gęstości prawdopodobieństwa

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,v)\, dv$$

Poprzedni razem cd

$$cov(X,Y) = E[(X - E(X))(Y - E(Y))] = \\ = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - E(X))(y - E(Y)) f_{X,Y}(x,y)\, dx\, dy$$

$$\rho_{X,Y} \equiv \rho(X,Y) \equiv corr(X,Y) = \frac{cov(X,Y)}{\sqrt{var(X) var(Y)}}$$

$$-1 \le \rho_{X,Y} \le 1$$

$$\rho_{X,Y} = 1 \quad dla \quad Y = aX + b \quad a > 0$$
$$\rho_{X,Y} = -1 \quad dla \quad Y = aX + b \quad a < 0$$

Współczynnik korelacji zmiennych losowych

4. **Współczynnik korelacji jest miarą zależności liniowej**

$$\rho_{X,Y} = 1 \quad dla \quad Y = aX + b \quad a > 0$$
$$\rho_{X,Y} = -1 \quad dla \quad Y = aX + b \quad a < 0$$

Dowód:

$$Y = aX + b$$
$$\to \quad E(Y) = E(aX + b) = aE(X) + b$$
$$\mathrm{var}(Y) = \mathrm{var}(aX + b) = a^2\, \mathrm{var}(X)$$
$$\mathrm{cov}(X,Y) = E\big[(X - E(X))(Y - E(Y))\big] = E\big[(X - E(X))(aX + b - aE(X) - b)\big] =$$
$$= E\big[(X - E(X))(aX - aE(X))\big] = aE\big[(X - E(X))(X - E(X))\big] =$$
$$= aE\Big[(X - E(X))^2\Big] = a\, \mathrm{var}(X)$$

$$\to \quad \rho_{X,Y} = \frac{\mathrm{cov}(X,Y)}{\sqrt{\mathrm{var}(X)\,\mathrm{var}(Y)}} = \frac{a\,\mathrm{var}(X)}{\sqrt{\mathrm{var}(X) \cdot a^2\, \mathrm{var}(X)}} = \frac{a\,\mathrm{var}(X)}{\sqrt{a^2 (\mathrm{var}(X))^2}} = \frac{a\,\mathrm{var}(X)}{|a|\,\mathrm{var}(X)} = \frac{a}{|a|}$$

Współczynnik korelacji zmiennych losowych

- 5. Jeżeli $\rho_{X,Y}=0$ to zmienne X i Y nie muszą być niezależne; nazywamy je wtedy **nieskorelowanymi.**

  Ważny wyjątek: jeżeli X i Y mają rozkłady normalne i $\rho_{X,Y}=0$ to zmienne X i Y są niezależne.
- 6. Jeżeli $\rho_{X,Y} \neq 0$ to zmienne X i Y nie są niezależne.
- **Macierz kowariancji**

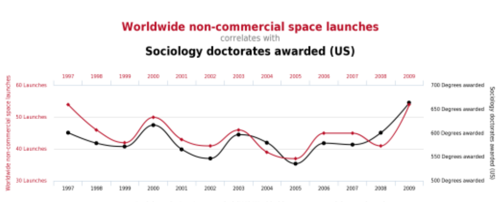
$$K_{X_1,\dots,X_n} = \begin{pmatrix} var(X_1) & cov(X_1,X_2) & \dots & cov(X_1,X_n) \\ cov(X_1,X_2) & var(X_2) & & \vdots \\ \vdots & & \ddots & \vdots \\ cov(X_1,X_n) & \dots & \dots & var(X_n) \end{pmatrix} \xrightarrow{2-dim} K_{X,Y} = \begin{pmatrix} var(X) & cov(X,Y) \\ cov(Y,X) & var(Y) \end{pmatrix}$$

Jest symetryczna, bo $cov(X_i,X_j)=cov(X_j,X_i)$.

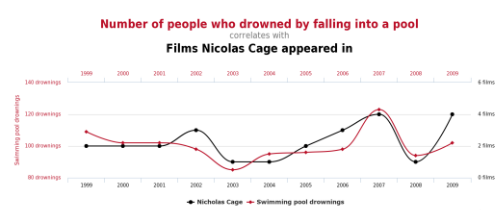
$$V_N = \frac{2}{N} \frac{\pi^{N/2}}{\Gamma(N/2)}$$

- Korelacja jako narzędzie testowania generatorów liczb pseudolosowych ($\rho_{X,Y}=0$), zadania kontrolne, np. N-wymiarowa kula o R=1.

http://tylervigen.com/spurious-correlations



http://tylervigen.com/spurious-correlations



Korelacje przypadkowe (fałszywe) czy prawdziwe?

Analiza ekspercka
Zmienne współzależne: X→Y, X→Z prowadzi do korelacji (Y,Z)
Zmienne pośredniczące: X→Y→Z prowadzi do korelacji (X,Z)
Analiza zebrania próby: zmienne kontrolne, badania z randomizacją, dobór próby
Analiza wyników: grupowanie, analiza regresji, analiza wariancji, sieci zależności

Wektor losowy – przypadek n>2

- Syt.: dany jest n-wymiarowy wektor losowy $\vec{X}$, z którego tworzymy wielowymiarową zmienną losową $\vec{Y}$:

$$\vec{Y} = \vec{Y}(\vec{X})$$

rozwijamy wokół Y=E(X), zostawiając człony liniowe

$$Y_i = \underbrace{Y_i(E(\vec{X}))}_{E(Y_i)} + \sum_j \left(\frac{\partial Y_i}{\partial X_j}\right)\Bigg|_{\vec{X}=E(\vec{X})} \left(X_j - E(X_j)\right) + \ldots$$

$$Y_i - E(Y_i) = \sum_j \left(\frac{\partial Y_i}{\partial X_j}\right)\Bigg|_{\vec{X}=E(\vec{X})} \left(X_j - E(X_j)\right)$$

Prawo przenoszenia błędów

$$cov(Y_k, Y_q) = E\left[(Y_k - E(Y_k))(Y_q - E(Y_q))\right] =$$

$$= E\left[\sum_l \left(\frac{\partial Y_k}{\partial X_l}\right)\Bigg|_{\vec{X}=E(\vec{X})} (X_l - E(X_l)) \sum_m \left(\frac{\partial Y_q}{\partial X_m}\right)\Bigg|_{\vec{X}=E(\vec{X})} (X_m - E(X_m))\right] =$$

$$= \sum_{l,m} \left(\frac{\partial Y_k}{\partial X_l}\right)\Bigg|_{\vec{X}=E(\vec{X})} \left(\frac{\partial Y_q}{\partial X_m}\right)\Bigg|_{\vec{X}=E(\vec{X})} E\left[(X_l - E(X_l))(X_m - E(X_m))\right] =$$

$$= \boxed{\sum_{l,m} \left(\frac{\partial Y_k}{\partial X_l}\right)\Bigg|_{\vec{X}=E(\vec{X})} \left(\frac{\partial Y_q}{\partial X_m}\right)\Bigg|_{\vec{X}=E(\vec{X})} cov(X_l, X_m)}$$

$$[C(X)]_{l,m} \equiv cov(X_l, X_m); \quad [C(Y)]_{k,q} \equiv cov(Y_k, Y_q); \quad T_{i,j} \equiv \left(\frac{\partial Y_i}{\partial X_j}\right)\Bigg|_{\vec{X}=E(\vec{X})}$$

$$[C(Y)]_{k,q} = \sum_{l,m} T_{k,l} T_{q,m} [C(X)]_{l,m} = \sum_{l,m} T_{k,l} [C(X)]_{l,m} T^T_{m,q}$$

$\boxed{C(Y) = TC(X)T^T}$ ⟵ macierzowo

Prawo przenoszenia błędów - przykład

$$cov(Y_k, Y_q) = \sum_{l,m} \left(\frac{\partial Y_k}{\partial X_l}\right)\Bigg|_{\vec{X}=E(\vec{X})} \left(\frac{\partial Y_q}{\partial X_m}\right)\Bigg|_{\vec{X}=E(\vec{X})} cov(X_l, X_m)$$

- Pomiar średniej prędkości polega na pomiarze dystansu i czasu w jakim ten dystans jest pokonywany

$s \pm \Delta s \to s \pm \frac{1}{\sqrt{3}}\Delta s \equiv s \pm \sigma(s)$ $\quad t \pm \sigma(t)$ $\quad v = \frac{s}{t}$

$$\mathrm{var}(v) = \mathrm{cov}(v,v) = \sum_{l,m=\{s,t\}} \left(\frac{\partial Y_k}{\partial X_l}\right)\Bigg|_{\vec{X}=E(\vec{X})} \left(\frac{\partial Y_q}{\partial X_m}\right)\Bigg|_{\vec{X}=E(\vec{X})} cov(X_l, X_m) =$$

$$= \left(\frac{\partial v}{\partial s}\right)\Bigg|_{\vec{X}=E(\vec{X})} \left(\frac{\partial v}{\partial s}\right)\Bigg|_{\vec{X}=E(\vec{X})} cov(s,s) + \left(\frac{\partial v}{\partial s}\right)\Bigg|_{\vec{X}=E(\vec{X})} \left(\frac{\partial v}{\partial t}\right)\Bigg|_{\vec{X}=E(\vec{X})} cov(s,t) + \left(\frac{\partial v}{\partial t}\right)\Bigg|_{\vec{X}=E(\vec{X})} \left(\frac{\partial v}{\partial s}\right)\Bigg|_{\vec{X}=E(\vec{X})} cov(t,s) +$$

$$+ \left(\frac{\partial v}{\partial t}\right)\Bigg|_{\vec{X}=E(\vec{X})} \left(\frac{\partial v}{\partial t}\right)\Bigg|_{\vec{X}=E(\vec{X})} cov(t,t) = \frac{1}{t}\cdot\frac{1}{t}\cdot \mathrm{var}(s) + 2\cdot\frac{1}{t}\cdot\frac{(-s)}{t^2}\cdot cov(s,t) + \frac{(-s)}{t^2}\cdot\frac{(-s)}{t^2}\cdot \mathrm{var}(t) =$$

$$= \frac{1}{t^2}\cdot(\sigma(s))^2 - \frac{2s}{t^3}\cdot cov(s,t) + \frac{s^2}{t^4}\cdot(\sigma(t))^2$$

- Zał: zmienne (s,t) są niezależne → cov(s,t)=0

$$\mathrm{var}(v) = \frac{1}{t^2}\cdot(\sigma(s))^2 + \frac{s^2}{t^4}\cdot(\sigma(t))^2 \quad \to \sigma(v) = \sqrt{\frac{1}{t^2}\cdot(\sigma(s))^2 + \frac{s^2}{t^4}\cdot(\sigma(t))^2}$$

Transformacje wektorów losowych

- Syt. $\vec{Z} = g(\vec{X})$
- Przypadek dyskretny: grupujemy te wartości $\vec{X}$, które dają te same wartości $\vec{Z}$ i dodajemy prawdopodobieństwa. Otrzymujemy rozkład prawdopodobieństwa wektora losowego $\vec{Z}$.
- Przypadek ciągły:

I sposób: (zał. Z jest jednowymiarowe czyli $Z = g(X_1, X_2, \ldots, X_n)$)

$$F_Z(z) = \underbrace{\int\ldots\int}_{\substack{\text{po wszystkich zdarzeniach} \\ \text{dla których } Z(A)\le z}} f_{\vec{X}}(x_1, x_2, \ldots, x_n)\, dx_1 \ldots dx_n$$

Znając dystrybuantę obliczamy (łączną) funkcję gęstości prawdopodobieństwa.

Dla Z n-wymiarowej całkujemy po wszystkich zdarzeniach dla których $Z_1(A) \le z_1 \wedge Z_2(A) \le z_2 \wedge \ldots \wedge Z_n(A) \le z_n$.

Transformacje wektorów losowych

- Przypadek ciągły:

II sposób:

zał.:

1. n=2 czyli (X,Y)→(V,W) czyli V=g_1(X,Y), W=g_2(X,Y)
2. istnieją jednoznaczne funkcje h_1: X=h_1(V,W) i h_2:Y=h_2(V,W)
3. dla każdego x,y funkcje g_1 i g_2 mają ciągłe pochodne cząstkowe
4. Jacobian

$$J_{(X,Y)} \equiv \begin{vmatrix} \frac{\partial g_1}{\partial x} & \frac{\partial g_1}{\partial y} \\ \frac{\partial g_2}{\partial x} & \frac{\partial g_2}{\partial y} \end{vmatrix} \neq 0$$

Wtedy

$$\boxed{f_{V,W}(v,w) = f_{X,Y}(x(v,w), y(v,w)) \left|J_{(X,Y)}\right|^{-1}}$$

Transformacje wektorów losowych - przykład

- Niezależne zmienne losowe X,Y mają rozkłady jednorodne na (0,1). Znaleźć łączną funkcję gęstości prawdopodobieństwa zmiennych $V \equiv \sqrt{-2\ln(X)}\cos(2\pi Y)$ oraz $W \equiv \sqrt{-2\ln(X)}\sin(2\pi Y)$
- Rozwiązanie:

Z treści zadania: $f_{X,Y}(x,y) = f_X(x)\cdot f_Y(y) = 1\cdot 1 = 1$

Potrzebujemy: $f_{V,W}(v,w) = f_{X,Y}(x(v,w), y(v,w)) \left|J_{(X,Y)}\right|^{-1}$

Liczymy

$$\tfrac{\partial V}{\partial x} = \tfrac{\partial}{\partial x}\left(\sqrt{-2\ln(X)}\cos(2\pi Y)\right) = \cos(2\pi Y)\tfrac{\partial}{\partial x}\left(\sqrt{-2\ln(X)}\right) = \tfrac{\cos(2\pi Y)}{2\sqrt{-2\ln(X)}}\cdot\tfrac{(-2)}{X}$$

$$\tfrac{\partial V}{\partial y} = \tfrac{\partial}{\partial y}\left(\sqrt{-2\ln(X)}\cos(2\pi Y)\right) = \sqrt{-2\ln(X)}\tfrac{\partial}{\partial y}\left(\cos(2\pi Y)\right) = \sqrt{-2\ln(X)}\cdot(-1)\sin(2\pi Y)\cdot 2\pi$$

$$\tfrac{\partial W}{\partial x} = \tfrac{\partial}{\partial x}\left(\sqrt{-2\ln(X)}\sin(2\pi Y)\right) = \sin(2\pi Y)\tfrac{\partial}{\partial x}\left(\sqrt{-2\ln(X)}\right) = \tfrac{\sin(2\pi Y)}{2\sqrt{-2\ln(X)}}\cdot\tfrac{(-2)}{X}$$

$$\tfrac{\partial W}{\partial y} = \tfrac{\partial}{\partial y}\left(\sqrt{-2\ln(X)}\sin(2\pi Y)\right) = \sqrt{-2\ln(X)}\tfrac{\partial}{\partial y}\left(\sin(2\pi Y)\right) = \sqrt{-2\ln(X)}\cdot\cos(2\pi Y)\cdot 2\pi$$

Transformacje wektorów losowych - przykład

$$V \equiv \sqrt{-2\ln(X)}\cos(2\pi Y)$$
$$W \equiv \sqrt{-2\ln(X)}\sin(2\pi Y)$$

$$J_{(X,Y)} = \det\begin{pmatrix} \frac{\partial V}{\partial x} & \frac{\partial V}{\partial y} \\ \frac{\partial W}{\partial x} & \frac{\partial W}{\partial y} \end{pmatrix} = \frac{\partial V}{\partial x}\cdot\frac{\partial W}{\partial y} - \frac{\partial V}{\partial y}\cdot\frac{\partial W}{\partial x} = \left(\cos(2\pi Y)\right)^2\cdot\tfrac{(-2\pi)}{X} - \left(\sin(2\pi Y)\right)^2\cdot\tfrac{(2\pi)}{X} =$$

$$= \tfrac{(-2\pi)}{X}\cdot\left(\left(\cos(2\pi Y)\right)^2 + \left(\sin(2\pi Y)\right)^2\right) = \tfrac{(-2\pi)}{X}$$

$$\left|J_{(X,Y)}\right|^{-1} = \left|\tfrac{(-2\pi)}{X}\right|^{-1} = \left(\tfrac{2\pi}{X}\right)^{-1} = \tfrac{X}{2\pi}$$

$$V^2 + W^2 = \left(\sqrt{-2\ln(X)}\cos(2\pi Y)\right)^2 + \left(\sqrt{-2\ln(X)}\sin(2\pi Y)\right)^2 = \left(\sqrt{-2\ln(X)}\right)^2 = -2\ln(X)$$

$$\rightarrow \quad X = e^{\frac{-1}{2}\left(V^2+W^2\right)} \quad \rightarrow \quad \left|J_{(X,Y)}\right|^{-1} = \tfrac{1}{2\pi}e^{\frac{-1}{2}\left(V^2+W^2\right)}$$

$$\rightarrow \quad f_{V,W}(v,w) = f_{X,Y}(x(v,w),y(v,w))\left|J_{(X,Y)}\right|^{-1} = 1\cdot\tfrac{1}{2\pi}e^{\frac{-1}{2}\left(V^2+W^2\right)} = \tfrac{1}{2\pi}e^{\frac{-1}{2}\left(V^2+W^2\right)}$$

$$\rightarrow \quad f_{V,W}(v,w) = \tfrac{1}{\sqrt{2\pi}}e^{\frac{-1}{2}V^2}\cdot\tfrac{1}{\sqrt{2\pi}}e^{\frac{-1}{2}W^2} = N_V(0,1)\cdot N_W(0,1) = f_V(v)\cdot f_W(w)$$

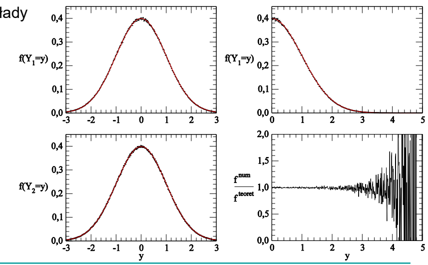
Dwa niezależne rozkłady N(0,1)

Transformacja Box-Mullera

- X_1, X_2 mają rozkłady jednorodne na przedziale (0,1); wtedy

$$Y_1 = \sqrt{-2ln(X_1)}\,cos(2\pi X_2) \qquad i \qquad Y_2 = \sqrt{-2ln(X_1)}\,sin(2\pi X_2)$$

mają rozkłady N(0,1).



Wielowymiarowy rozkład normalny

- Łączna gęstość prawdopodobieństwa

$$\boxed{f_{\vec{X}}(\vec{x}) = \left(\frac{\det(K^{-1})}{(2\pi)^n}\right)^{1/2} e^{\frac{-1}{2}(\vec{x}-\vec{a})^T K^{-1}(\vec{x}-\vec{a})}}$$

$\vec{X}$ – wektor n zmiennych losowych
K – macierz kowariancji nxn
[redacted] – stały wektor n liczb
(E(X[redacted]a)

- dla n=2

$$K^{-1} = \frac{1}{var(X_1)var(X_2) - (cov(X_1,X_2))^2}\begin{pmatrix} var(X_2) & -cov(X_1,X_2) \\ -cov(X_1,X_2) & var(X_1) \end{pmatrix}$$

- Standaryzacja $U_i = \frac{X_i - a_i}{\sigma(X_i)}$ dla n=2 prowadzi do

$$f_{\vec{U}}(u_1,u_2) = \left(\frac{\det(K^{-1})}{(2\pi)^2}\right)^{1/2} e^{\frac{-1}{2}\vec{u}^T K^{-1}\vec{u}}$$

ρ to współczynnik korelacji u_1 i u_2

$$K^{-1} = \frac{1}{1-\rho^2}\begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix}$$

Wielowymiarowy rozkład normalny - elipsa kowariancji

- Łączna gęstość prawdopodobieństwa

$$f_{\vec{U}}(u_1,u_2) = const \rightarrow \vec{u}^T K^{-1}\vec{u} = const \rightarrow$$

$$\frac{1}{1-\rho^2}(u_1^2 + u_2^2 - 2u_1u_2\rho) = const \rightarrow$$

$$\frac{1}{1-\rho^2}\left(\frac{(x_1-a_1)^2}{(\sigma(X_1))^2} + \frac{(x_2-a_2)^2}{(\sigma(X_2))^2} - 2\rho\frac{(x_1-a_1)}{\sigma(X_1)}\frac{(x_2-a_2)}{\sigma(X_2)}\right) = const$$

- Jest to równanie elipsy zwanej dla const=1 elipsą kowariancji. Prawdopodobieństwo wystąpienia zmiennych losowych wewnątrz elipsy kowariancji jest niezależne od $\sigma(X_1)$,$\sigma(X_2)$ i ρ.

Dystrybuanta warunkowa

- Syt. Szukamy **dystrybuanty warunkowej zmiennej X** czyli dystrybuanty zmiennej X pod warunkiem, że zmienna losowa Y=y. Ale dla zmiennej ciągłej P(Y=y)=0 dlatego rozpatrujemy y<Y<y+dy i przejdziemy do granicy dy→0.

$$\int_{-\infty}^{x}\int_{y}^{y+dy} f_{X,Y}(u,v)\,du\,dv = F_{X|Y}\left(x\,|\,Y\in(y,y+dy)\right)\cdot\int_{y}^{y+dy} f_Y(v)\,dv$$

$$\int_{-\infty}^{x} f_{X,Y}(u,y)\,du\,dy = F_{X|Y}\left(x\,|\,Y\in(y,y+dy)\right)\cdot f_Y(y)\,dy \qquad /:dy \quad /dy\rightarrow 0$$

$$\int_{-\infty}^{x} f_{X,Y}(u,y)\,du = F_{X|Y}\left(x\,|\,Y=y)\right)\cdot f_Y(y)$$

$$\boxed{F_{X|Y}(x\,|\,y) = \frac{\int_{-\infty}^{x} f_{X,Y}(u,y)\,du}{f_Y(y)}}$$

Warunkowa funkcja gęstości prawdopodobieństwa

$$\boxed{f_{X|Y}(x\,|\,y) = \frac{d}{dx}F_{X|Y}(x\,|\,y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}}$$

- Definicja ta prowadzi do równoważnej definicji niezależności dwóch zmiennych losowych poprzez warunek:

$$f_{X|Y}(x\,|\,y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

$$f_{X|Y}(x\,|\,y)\cdot f_Y(y) = f_{X,Y}(x,y)$$

$$f_{X|Y}(x\,|\,y)\cdot f_Y(y) = f_X(x)\cdot f_Y(y)$$

$$\boxed{\begin{matrix} f_{X|Y}(x\,|\,y) = f_X(x) \\ F_{X|Y}(v\,|\,y) = F_X(v) \end{matrix}} \qquad / \int_{-\infty}^{v} dx$$

Warunkowa wartość oczekiwana

$$\boxed{E(X \mid Y=y)=\int_{-\infty}^{+\infty} x f_{X|Y}(x \mid y)dx}$$

- Jest to funkcja zmiennej losowej Y, zatem E(X|Y) też jest zmienną losową

$$E[E(X \mid Y)]=\int_{-\infty}^{+\infty} E(X \mid Y=y)\, f_Y(y)\, dy=$$

$$=\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{+\infty} x\, f_{X|Y}(x \mid y)\, dx\right] f_Y(y)\, dy=\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{+\infty} x \frac{f_{X,Y}(x,y)}{f_Y(y)}\, dx\right] f_Y(y)\, dy=$$

$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x\, f_{X,Y}(x,y)\, dx\, dy=\int_{-\infty}^{+\infty} x\, f_X(x)\, dx=E(X)$$

ogólnie dla dowolnej funkcji h(X) $\boxed{E(h(X))=E[E(h(X) \mid Y)]}$

Twierdzenia graniczne

Związki pomiędzy różnymi charakterystykami rozkładów najczęściej badane w jakiejś granicy.

Przykład:

- **Nierówność Czebyszewa-Bienayme**

  Zał. Dla zmiennej losowej X istnieje E(X):=μ i var(X)=σ^2<+∞

$$\boxed{\forall a>0: \quad P(|X-\mu| \geq a) \leq \frac{\sigma^2}{a^2}}$$

Powyższe twierdzenie nie zależy od rozkładu → ograniczenia nie są silne

Twierdzenia graniczne

- **Prawo wielkich liczb Bernoulliego (słabe)**

  Zał. S_n – liczba sukcesów (o prawdopodobieństwie *p*) w n próbach

$$\boxed{\forall \varepsilon>0: \quad \lim_{n\to\infty} P\left(\left|\frac{S_n}{n}-p\right| \leq \varepsilon\right)=1}$$

- Słabe prawo „Jest bardzo prawdopodobne, że częstotliwość będzie równa p dla ustalonego, dużego n".
- Jest to „zbieżność według prawdopodobieństwa"
- **Mocne prawo wielkich liczb Bernoulliego**

  Zał. S_n – liczba sukcesów (o prawdopodobieństwie *p*) w n próbach

$$\boxed{\forall \varepsilon>0: \quad P\left(\lim_{n\to\infty}\left|\frac{S_n}{n}-p\right| \leq \varepsilon\right)=1} \qquad \boxed{P\left(\lim_{n\to\infty} \frac{S_n}{n}=p\right)=1}$$

- Mocne prawo „Częstotliwość musi być w pobliżu p gdy n rośnie".
- Jest to „zbieżność prawie na pewno"

Całkowanie metodą Monte Carlo

- **Mocne prawo wielkich liczb Kołmogorowa**
- X_n – ciąg niezależnych zmiennych losowych o jednakowym rozkładzie, E(|X_1|) jest skończona, $S_n=X_1+X_2+\ldots+X_n$

$$\boxed{P\left(\lim_{n\to\infty} \frac{S_n}{n}=E(X_1)\right)=1}$$

- Uzasadnia to definicję częstotliwościową prawdopodobieństwa
- Problem: całkowanie numeryczne f(x) w przedziale (a,b)

$$I=\int_a^b f(x)dx=\int_a^b f(x)\frac{g(x)}{g(x)}dx=E\left(\frac{f(x)}{g(x)}\right)=\frac{1}{n}\sum_{i=1}^{n}\frac{f(x_i)}{g(x_i)}$$

$$\boxed{g(x)>0 \qquad \int_a^b g(x)dx=1}$$

MPWL

Całkowanie metodą Monte Carlo

- Zatem $I=\int_a^b f(x)\,dx=\frac{1}{n}\sum_{i=1}^{n}\frac{f(x_i)}{g(x_i)}$
- Niech g(x_i) – jednostajna na przedziale (a,b).

$$\int_a^b f(x)\,dx=\frac{(b-a)}{n}\sum_{i=1}^{n} f(x_i)$$

- Zał. Znamy *var*(f(x_i)). Wtedy *var*(I)=*var*(f(x_i))/n
- Poprawę jakości można uzyskać min przez:
  1) dobierając g(x) jako podobną do f(x)
  2) losowanie warstwowe (rozbijamy całkę na przedziały, gdzie f(x) jest w przybliżeniu stała
  3) metodą zmiennych kontrolnych – h(x) podobne do f(x) ale o znanej całce. $\int_a^b f(x)\,dx=\int_a^b (f(x)-h(x))dx+\int_a^b h(x)\,dx$

Całkowanie metodą Monte Carlo

- Całkowanie metodą Monte Carlo stosujemy do całek wielokrotnych.
- Np. całka dziesięciokrotna z $f(x_1,x_2,\dots,x_{10})$, przy czym jednokrotne wyliczenie $f(x_1,x_2,\dots,x_{10})$ zajmuje 0.0001 sekundy.
- Metoda prostokątów wymaga N^{10} wyliczeń funkcji
  Dla N=10 czas wykonania programu $10^{10}/10^4=10^6$ sekund
  (około 12 dób)
- Metoda Monte Carlo
  Niepewność $\sigma(f(x_i))\ n^{(-0.5)}$
  Dla $\sigma(f(x_i))=0.01f(x_i)$ (1%) i $n=10^6$ mamy
  dokładność rzędu 0.00001 (0.001%)
  czas wykonania programu $10^6/10^4=100$ sekund

Twierdzenia graniczne

- **Centralne twierdzenie graniczne**
  Zał. Zmienne losowe $X_1,X_2,\dots,X_n$ są niezależnymi zmiennymi o tej samej funkcji gęstości prawdopodobieństwa z $E(X_i):=\mu$ i $var(X_i)=\sigma^2>0$

$$\lim_{n\to\infty} P\left(\left|\frac{x_1+\ldots+x_n-n\mu}{\sigma\sqrt{n}}\right|\le a\right)=F_{N(0,1)}(a)-F_{N(0,1)}(-a)$$

  gdzie $F_{N(0,1)}(a)$ oznacza dystrybuantę rozkładu N(0,1).
- Wniosek: zmienna będąca sumą innych zmiennych losowych dąży do rozkładu $N(n\mu,n\sigma^2)$.
- Ogólniejsza wersja twierdzenia mówi, że suma zmiennych o dowolnych funkcjach gęstości o dodatnich wariancjach, w której żaden ze składników nie dominuje, ma rozkład normalny.