Wielowymiarowe zmienne losowe (wektory losowe)

- Syt. Eksperyment losowy E, związana z nim przestrzeń próbek S
- **n-wymiarowy wektor losowy** $\vec{X}$ to funkcja, która każdemu elementowi s należącemu zbioru S przyporządkowuje wektor n liczb rzeczywistych ($X_1(s)$, $X_2(s)$,…,$X_n(s)$) (n zmiennych losowych),
- Np. Losujemy punkt na płaszczyźnie $X_1(s)=x, X_2(s)=y$
- Nie ma związku pomiędzy wymiarem elementu w S (czyli liczbą zmiennych potrzebnych do opisania wyniku eksperymentu) a wymiarem wektora losowego $\vec{X}$.

  np. $\vec{X}=(X_1,X_2,X_3,X_4)$:

  $$X_1(s)=x_a, X_2(s)=y_a, X_3(s)=x_a+y_a, X_4(s)=sign(x_a)$$
- W dalszym ciągu ograniczymy się do n=2

Łączny rozkład prawdopodobieństwa

Syt. $\vec{Z}=(X,Y)$ $\quad S_{\vec{Z}}=\{(x_i,y_j)\}$ $\quad i=1,2,\ldots$ $\quad j=1,2,\ldots$

- **Łącznym rozkładem prawdopodobieństwa** nazywamy (dwuargumentową) funkcję, która

  $\forall i,j:\quad P_{X,Y}(x_i,y_j)=P[\{X=x_i\}\cap\{Y=y_j\}]\equiv P[X=x_i,Y=y_j]$
- Własności:

  $\forall i,j \quad P_{X,Y}(x_i,y_j)\geq 0$

  $\sum_i\sum_j P_{X,Y}(x_i,y_j)=1$

  $P(A)=\sum_{(x_i,y_j)\in A}P_{X,Y}(x_i,y_j)$
- **Łączna dystrybuanta**

  $F_{X,Y}(x,y)=P[X\leq x,Y\leq y]=\sum_{x_i\leq x}\sum_{y_j\leq y}P_{X,Y}(x_i,y_j)$

Łączna funkcja gęstości prawdopodobieństwa

Syt. $\vec{Z}=(X,Y)$, X,Y – ciągłe zmienne losowe

- **Łączną funkcją gęstości prawdopodobieństwa pary (X,Y)** nazywamy funkcję, która wiąże się z prawdopodobieństwem zdarzenia A poprzez:

  $P(A)=\iint_A f_{X,Y}(x,y)\,dx\,dy$
- Własności:

  $\forall x,y \quad f_{X,Y}(x,y)\geq 0$

  $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_{X,Y}(x,y)\,dx\,dy=1$

  $P(x<X\leq x+dx, y<Y\leq y+dy)=f_{X,Y}(x,y)\,dx\,dy$

Kowariancja i korelacja zmiennych losowych

- Syt. Z=g(X,Y)

  $$E(Z)=\begin{cases}\sum_i\sum_j g(x_i,y_j)P_{X,Y}(x_i,y_j)\\ \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)f(x,y)\,dx\,dy\end{cases}$$
- Wnioski: 1. $E(aX+bY)=aE(X)+bE(Y)$
- Dla zmiennych niezależnych

  2. $E(g_1(X)\cdot g_2(Y))=E(g_1(X))\cdot E(g_2(Y))$

  w szczególności dla $g_1(X)=X$ i $g_2(Y)=Y$ $\quad E(XY)=E(X)\cdot E(Y)$
- Dla zmiennych zależnych

  $E(XY)\neq E(X)\cdot E(Y)$

  $cov(X,Y)=E(XY)-E(X)\cdot E(Y)$

Kowariancja i korelacja zmiennych losowych

- Wzór $cov(X,Y)=E(XY)-E(X)\cdot E(Y)$ jest równoważny

  $$\boxed{\begin{aligned}&cov(X,Y)=E[(X-E(X))(Y-E(Y))]=\\&=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-E(X))(y-E(Y))f_{X,Y}(x,y)\,dx\,dy\end{aligned}}$$
- Własności:

  1. $cov(X,X)=var(X)\geq 0$
  2. Kowariancja nie musi być nieujemna.
  3. Dla zmiennych niezależnych cov(X,Y)=0
  4. E(XY) nazywamy **korelacją zmiennych** X i Y

Współczynnik korelacji zmiennych losowych

- **Współczynnik korelacji** zmiennych losowych X,Y (unormowana kowariancja) to

  $$\boxed{\rho_{X,Y}\equiv corr(X,Y)=\frac{cov(X,Y)}{\sqrt{var(X)var(Y)}}}$$
- Własności:

  1. Współczynnik korelacji jest bezwymiarowy.
  2. $-1\leq\rho_{X,Y}\leq 1$
  3. **Współczynnik korelacji jest miarą zależności liniowej**

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

Współczynnik korelacji zmiennych losowych

- Własności:
  4. Dla X,Y niezależnych $\rho_{X,Y}=0$.
  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**

$$K_{X,Y}=\begin{pmatrix} var(X) & cov(X,Y) \\ cov(Y,X) & var(Y) \end{pmatrix}$$

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

Jest symetryczna, bo *cov*(X,Y)=*cov*(Y,X).

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

Twierdzenia graniczne

- **Nierówność Czebyszewa**

Zał. Zmienna losowa X ma wartości nieujemne

$$\boxed{\forall a>0:\quad P(X\geq a)\leq \frac{E(X)}{a}}$$

Dowód:

$$E(X)=\sum_{x_k} x_k P_X(x_k)=\sum_{x_k<a} x_k P_X(x_k)+\sum_{x_k\geq a} x_k P_X(x_k)\geq$$

$$\geq \sum_{x_k\geq a} x_k P_X(x_k)\geq \sum_{x_k\geq a} aP_X(x_k)=a\sum_{x_k\geq a} P_X(x_k)=aP_X(X\geq a)$$

$$\Rightarrow\quad E(X)\geq aP_X(X\geq a)\quad\Rightarrow\quad P_X(X\geq a)\leq\frac{E(X)}{a}$$

Twierdzenia graniczne

- **Nierówność Czebyszewa-Bienayme**

Zał. Dla zmiennej losowej X istnieje E(X):=μ i var(X)=$\sigma^2<\infty$

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

Dowód:

$(X-\mu)^2\geq 0$ $\Rightarrow$ (z nierówności Czebyszewa)

$$P\left((X-\mu)^2\geq a^2\right)\leq\frac{E((X-\mu)^2)}{a^2}$$

$$P\left(|X-\mu|\geq a\right)\leq\frac{var(X)}{a^2}=\frac{\sigma^2}{a^2}$$

Oba powyższe twierdzenia nie zależą od rozkładu.

Przykład: losowanie kul

Twierdzenia graniczne

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

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

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

Dowód (rozpatrujemy zdarzenie przeciwne):

$$X\equiv\frac{S_n}{n}\quad\Rightarrow\quad E(X)=E\left(\frac{S_n}{n}\right)=\frac{1}{n}E(S_n)=\frac{1}{n}np=p$$

$$var(X)=var\left(\frac{S_n}{n}\right)=\frac{1}{n^2}var(S_n)=\frac{1}{n^2}np(1-p)=\frac{p(1-p)}{n}$$

(z nierówności Czebyszewa-Bienayme)

$$P(|X-\mu|\geq a)\leq\frac{\sigma^2}{a^2}\quad\Rightarrow\quad P\left(\left|\frac{S_n}{n}-p\right|>\varepsilon\right)\leq\frac{p(1-p)}{n\varepsilon^2}\xrightarrow{\lim_{n\to\infty}}$$

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

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 Kołomogorowa**
- 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.
  $$\int_a^b f(x)\,dx = \int_a^b \left(f(x)-h(x)\right)dx + \int_a^b h(x)\,dx$$

Całkowanie metodą Monte Carlo

- Całkowanie metodą Monte Carlo stosujemy do całek wielokrotnych.
- Np. całka dziesięciokrotna z f(x_1,x_2,…,x_{10}), przy czym jednokrotne wyliczenie f(x_1,x_2,…,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 σ(f(x_i)) $n^{(-0.5)}$
  Dla σ(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 $X_1, X_2, \ldots, X_n$ są niezależnymi zmiennymi o tej samej funkcji gęstości prawdopodobieństwa z $E(X_i):=\mu$ i $var(X_i)=\sigma^2>0$

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

  gdzie $F_{N(0,1)}(a)$ oznacza dystrybuantę rozkładu N(0,1).
- Wniosek: zmienna będąca sumą innych zmiennych losowych dąży do rozkładu N(nμ,nσ^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.