RACHUNEK PRAWDOPODOBIEŃSTWA

WYKŁAD 8.

ZBIEŻNOŚĆ CIĄGU ZMIENNYCH LOSOWYCH.

TWIERDZENIA GRANICZNE

Zbieżność ciągu zmiennych losowych z prawdopodobieństwem 1 (prawie napewno)

Ciąg zmiennych losowych (X_n) jest zbieżny do zmiennej losowej X z prawdopodobieństwem 1 jeśli

$$P\left(\left\{\omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\right\}\right) = 1$$

Średniokwadratowa zbieżność ciągu zmiennych losowych

Ciąg zmiennych losowych (X_n) jest średniokwadratowo zbieżny do zmiennej losowej X jeśli

$$\lim_{n\to\infty} E\left(\left|X_n - X\right|^2\right) = 0$$

Rozpatrując ten rodzaj zbieżności zakładamy, że dla występujących tu zmiennych losowych (X_n), X istnieje skończony moment rzędu 2.

Niekiedy stosuje się zapis $\text{l.i.m. } X_n = X$ (skrót od „limit in mean”).

Stochastyczna zbieżność ciągu zmiennych losowych

Ciąg zmiennych losowych (X_n) jest stochastycznie (wg prawdopodobieństwa) zbieżny do zmiennej losowej X jeśli

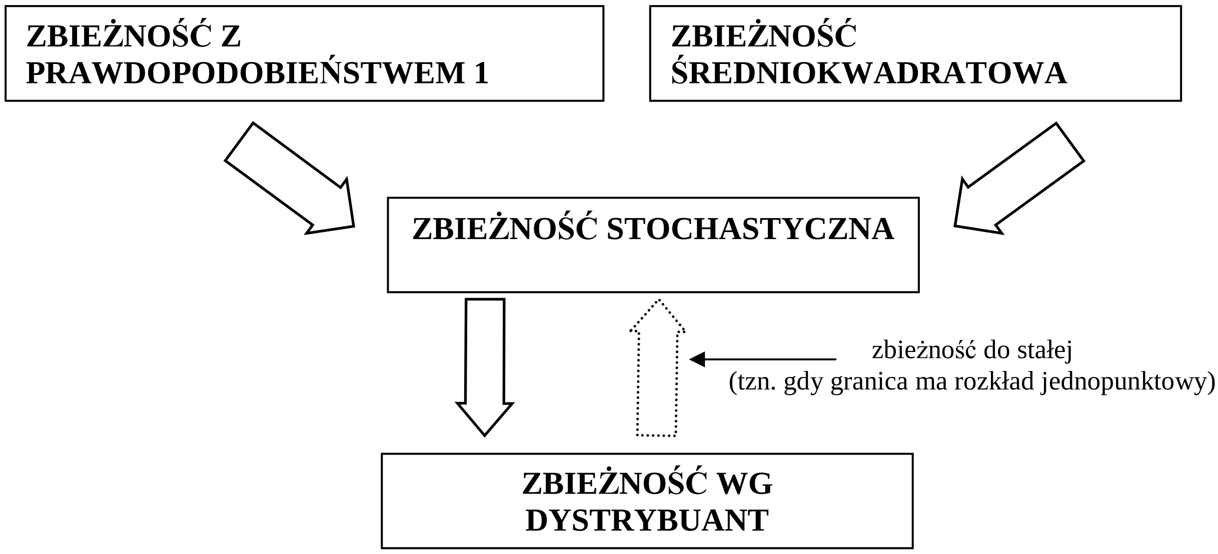
$$\bigwedge_{\varepsilon>0} \lim_{n\to\infty} P\left(\left|X_n - X\right| < \varepsilon\right) = 1$$

lub równoważnie

$$\bigwedge_{\varepsilon>0} \lim_{n\to\infty} P\left(\left|X_n - X\right| \geq \varepsilon\right) = 0$$

Zbieżność ciągu zmiennych losowych wg dystrybuant (wg rozkładu)

Ciąg zmiennych losowych (X_n) jest zbieżny do zmiennej losowej X wg dystrybuant jeśli ciąg ich dystrybuant F_n jest zbieżny do dystrybuanty F w każdym punkcie jej ciągłości (F jest dystrybuantą zmiennej losowej X).


ZBIEŻNOŚĆ Z PRAWDOPODOBIEŃSTWEM 1
ZBIEŻNOŚĆ ŚREDNIOKWADRATOWA
ZBIEŻNOŚĆ STOCHASTYCZNA
zbieżność do stałej
(tzn. gdy granica ma rozkład jednopunktowy)
ZBIEŻNOŚĆ WG DYSTRYBUANT

Przykład.
Rozpatrzmy ciąg zmiennych losowych skokowych określonych na przedziale [0, 1) w następujący sposób

$$X_{kn}(\omega) = \begin{cases} 1 & \text{gdy } \omega \in \left[\frac{k}{n}; \frac{k+1}{n}\right) \\ 0 & \text{gdy } \omega \in [0,1) - \left[\frac{k}{n}; \frac{k+1}{n}\right) \end{cases}$$

$$P(X_{kn} = 1) = \frac{1}{n}; \qquad P(X_{kn} = 0) = 1 - \frac{1}{n}$$

Ciąg X_{01}, X_{02}, X_{12}, X_{03}, X_{13}, X_{23}, ... jest zbieżny stochastycznie do zera bo

$$\bigwedge_{0<\varepsilon<1} \lim_{n\to\infty} P\left(\left|X_n\right| \geq \varepsilon\right) = \lim_{n\to\infty} \frac{1}{n} = 0$$

Natomiast ciąg ten nie jest zbieżny w żadnym punkcie przedziale [0, 1) bowiem dla każdego ustalonego punktu otrzymujemy rozbieżny ciąg zer i jedynek (zera i jedynki występują na dowolnie dalekich miejscach).

Przykład.

Ciąg zmiennych losowych X_n ciągłych o rozkładach jednostajnych na przedziałach (0, 1/n) jest zbieżny do rozkładu jednopunktowego X ($P(X = 0) = 1$) wg dystrybuant.

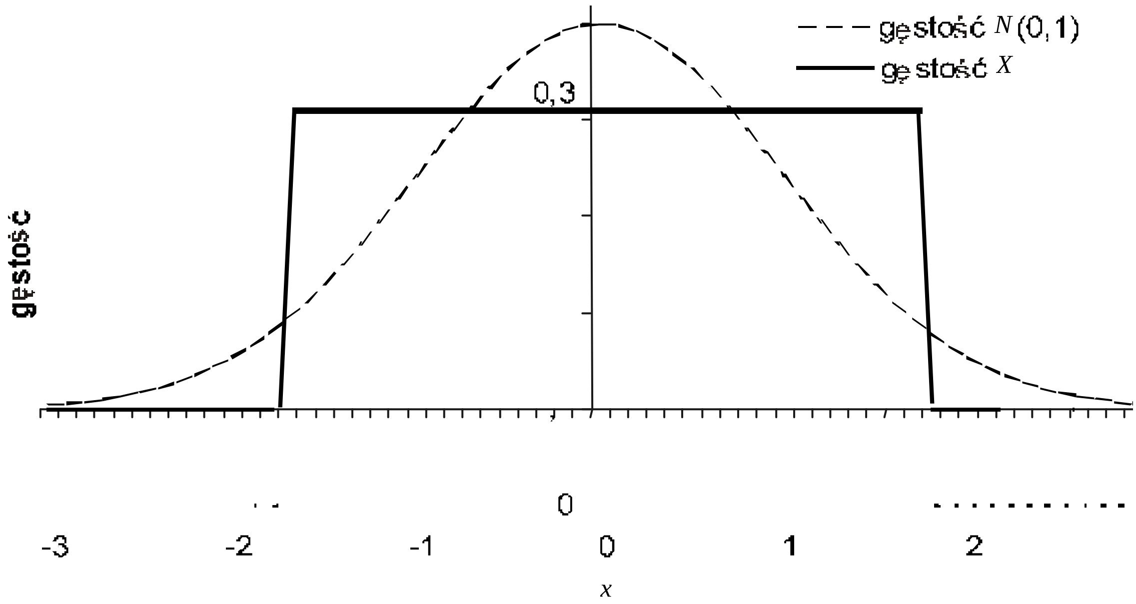
Centralne twierdzenie graniczne Lindeberga – Levy'ego

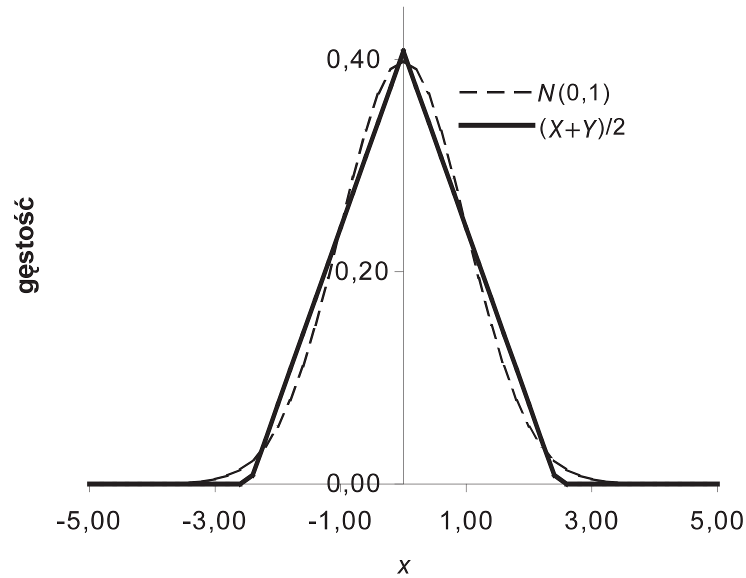
Jeśli niezależne zmienne losowe X_i (i = 1, 2, ..., n) mają taki sam rozkład oraz istnieje $E(X_n) = m$ i $D^2(X_n) = \sigma^2 > 0$ to ciąg dystrybuant (F_n) standaryzowanych średnich arytmetycznych $\overline{X}_n$ (lub standaryzowanych sum $\sum_{i=1}^{n} X_i$)

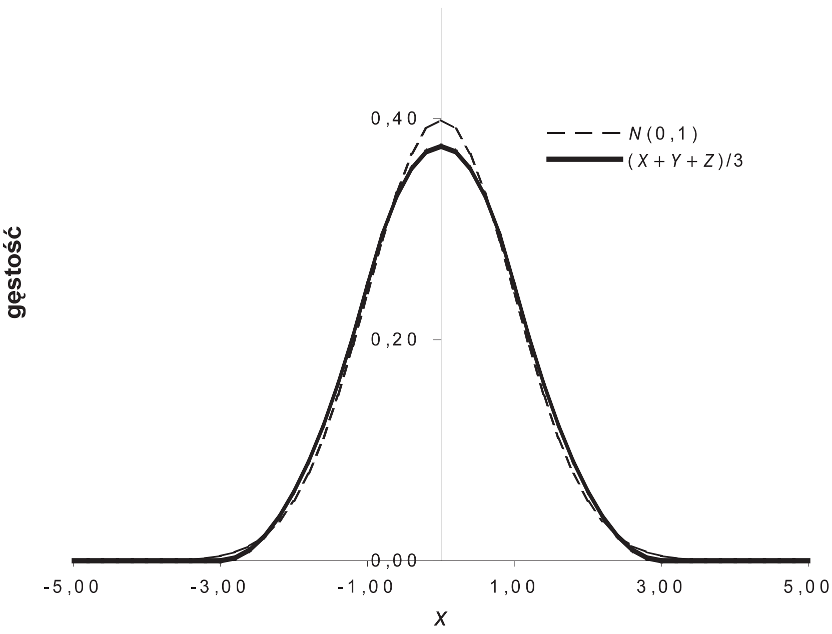
$$Y_n = \frac{\overline{X}_n - m}{\sigma / \sqrt{n}} = \frac{\sum_{i=1}^{n} X_n - mn}{\sigma \sqrt{n}}$$

jest zbieżny do dystrybuanty Φ rozkładu N(0, 1).

Aby się przekonać, że suma niezależnych zmiennych losowych o takim samym rozkładzie może dążyć do rozkładu $N(0, 1)$ porównajmy rozkład $N(0, 1)$ i standaryzowane rozkłady X, $(X + Y)/2$, $(X + Y + Z)/3$, gdzie X, Y, Z niezależne zmienne losowe o rozkładzie jednostajnym w przedziale $[-0{,}5; 0{,}5]$.


gęstość $N(0,1)$
gęstość X
0,3
gęstość
0
-3
-2
-1
0
1
2
x


0,40
$N(0,1)$
$(X+Y)/2$
gęstość
0,20
0,00
-5,00
-3,00
-1,00
1,00
3,00
5,00
x


gęstość
0,40
0,20
0,00
N(0,1)
(X + Y + Z)/3
-5,00
-3,00
-1,00
1,00
3,00
5,00
x

Wniosek

Dla dużych n (w praktyce $n \geq 30$)

$$P\left(a \leq \frac{\sum_{i=1}^{n} X_i - nm}{\sigma\sqrt{n}} < b \right) \cong \Phi(b) - \Phi(a)$$

W przypadku szczególnym gdy X_i ($i = 1, 2, ..., n$) maja rozkład zerojedynkowy to powyższe twierdzenie nazywamy **twierdzeniem Moivre'a-Laplace'a** (zmienne losowe $Y_n = \sum_{i=1}^{n} X_i$ maja rozkład dwumianowy).

Wniosek z twierdzenia Moivre'a-Laplace'a:

$$P\left(a \le \frac{Y_i - np}{\sqrt{npq}} < b \right) \cong \Phi(b) - \Phi(a)$$

Uwaga. Powyższe twierdzenia wskazują na ważną rolę rozkładu normalnego.

Przykład

Wadliwość partii żarówek wynosi 0,01. Z tej partii żarówek wylosowano 625 żarówek. Obliczyć prawdopodobieństwo, że wśród wylosowanych żarówek będzie

a) mniej niż 10 wadliwych,

b) najwyżej 10 wadliwych.

Rozwiązanie.

Y_n – liczba wadliwych żarówek wśród wylosowanych,

Ad a)

$$P(Y_i < 10) = P\left(\frac{Y_i - 625 \cdot 0{,}01}{\sqrt{625 \cdot 0{,}01 \cdot 0{,}99}} < \frac{10 - 625 \cdot 0{,}01}{\sqrt{625 \cdot 0{,}01 \cdot 0{,}99}}\right) \cong$$

$$\cong \Phi(1{,}51) = 0{,}93448$$

Ad b)

$$P(Y_i \leq 10) = P(Y_i < 10) + P(Y_i = 10) = P(Y_i < 11) =$$

$$= P\left(\frac{Y_i - 625\,0{,}01}{\sqrt{625\,0{,}01 \cdot 0{,}99}} < \frac{11 - 625\,0{,}01}{\sqrt{625\,0{,}01 \cdot 0{,}99}} \right) \cong \Phi(1{,}91) = 0{,}97193$$

Prawo wielkich liczb Chinczyna

(X_i) – ciąg niezależnych zmiennych losowych o takim samym rozkładzie oraz niech istnieje $E(X_i) = m$.

Wtedy ciąg $Y_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ jest zbieżny stochastycznie do m.

Wniosek

Dla dużych n jeśli istnieje $D^2(X_n) = \sigma^2 > 0$ to

$$\bigwedge_{\varepsilon>0} \quad P\left(\left|Y_n - m\right| < \varepsilon\right) \cong 2\Phi\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right) - 1$$

Przypadek szczególny – **prawo wielkich liczb Bernoulliego:**

(X_i) – ciąg niezależnych zmiennych losowych o rozkładzie dwumianowym wtedy ciąg $\frac{X_n}{n}$ jest stochastycznie zbieżny do p.

Wniosek

Dla dużych *n*:

$$\bigwedge_{\varepsilon>0} \quad P\left(\left|\frac{X_n}{n}-p\right|<\varepsilon\right)\cong 2\Phi\left(\frac{\varepsilon\sqrt{n}}{\sqrt{pq}}\right)-1$$

Przykład

Wadliwość partii żarówek wynosi 0,1. Z tej partii żarówek losujemy *n* żarówek. Ile żarówek należy wylosować aby prawdopodobieństwo, że średnia liczba wadliwych żarówek różniła się co do wartości bezwzględnej od wadliwości partii o mniej niż 0,025 było co najmniej równe 0,95.

Rozwiązanie

Y_n – liczba wadliwych żarówek wśród wylosowanych

$$P\left(\left|\frac{Y_n}{n}-0{,}1\right|<0{,}025\right)\cong 2\Phi\left(\frac{0{,}025\sqrt{n}}{\sqrt{0{,}1\cdot 0{,}9}}\right)-1\geq 0{,}95$$

stąd

$$\Phi\left(\frac{0{,}025\sqrt{n}}{\sqrt{0{,}1\cdot 0{,}9}}\right)\geq 0{,}975$$

oraz $\dfrac{0{,}025\sqrt{n}}{\sqrt{0{,}1\cdot 0{,}9}}\geq 1{,}96$

zatem $\sqrt{n}\geq 23{,}52$ i $n > 553$.

Przybliżenia lokalne.
Przybliżenie lokalne Poissona

dla dużych n (praktycznie $n \geq 30$) i małych p (praktycznie $p \leq 0{,}2$) mamy

$$\binom{n}{k} p^k q^{n-k} \approx \frac{\lambda^k}{k!} e^{-\lambda} \qquad \text{gdzie } \lambda = n \cdot p$$

Przybliżenie lokalne Moivre'a-Laplace'a

dla dużych n mamy

$$\binom{n}{k} p^k q^{n-k} \approx \frac{1}{\sqrt{npq}} f\left(\frac{k-np}{\sqrt{npq}}\right)$$

Gdzie f- gęstość rozkładu N(0, 1).

Ocenę odchylenia wartości zmiennej losowej od jej wartości oczekiwanej daje **nierówność Czebyszewa**:

X – zmienna losowa oraz istnieje E(*X*) = *m* i $D^2(X) = \sigma^2 > 0$ wtedy

$$\bigwedge_{\varepsilon>0} P\left(\left|X - m\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{\varepsilon^2}$$

Z nierównością Czebyszewa związane są inne nierówności np.

1) **nierówność Markowa**

$$\bigwedge_{\varepsilon>0} \bigwedge_{p>0} P\left(|X| \geq \varepsilon\right) \leq \frac{E|X|^p}{\varepsilon^p}$$

2) **nierówność Czebyszewa II**

$$\bigwedge_{\varepsilon>0} P\left(X \geq \varepsilon\right) \leq \frac{EX}{\varepsilon}$$

3) **nierówność Czebyszewa III (wykładnicza)**

jeśli $Ee^{\lambda X} < \infty$ $\quad \bigwedge_{\varepsilon>0} P\left(X \geq \varepsilon\right) \leq \frac{Ee^{\lambda X}}{e^{\lambda\varepsilon}}$

4)**nierówność Bernsteina**

jeśli S_n – liczba sukcesów w n próbach Bernoulliego z prawdopodobieństwem sukcesu p to

$$\bigwedge_{\varepsilon>0} P\left(\left|\frac{S_n}{n} - p\right| \geq \varepsilon\right) \leq 2e^{-2n\varepsilon^2}$$