6 TWIERDZENIA GRANICZNE

Niech (X_n) będzie ciągiem zmiennych losowych (ciągiem losowym). Rozważmy odpowiadające mu: 1) ciąg funkcji prawdopodobieństwa $P_i(x)=P(X_i=x)$, $i \in N$, w przypadku zmiennych losowych dyskretnych albo ciąg gęstości $(f_n(x))$ w przypadku zmiennych losowych typu ciągłego oraz 2) ciąg dystrybuant $(F_n(x))$. Twierdzenia graniczne, a więc przy $n \to \infty$, dotyczące zbieżności grupy 1) nazywamy *twierdzeniami granicznymi lokalnymi*, a grupy 2) – *twierdzeniami granicznymi integralnymi.*

Przykładem granicznego twierdzenia lokalnego jest twierdzenie (2.7.18) o zbieżności funkcji prawdopodobieństwa rozkładu Bernoulliego do funkcji prawdopodobieństwa rozkładu Poissona.

6.1. CENTRALNE TWIERDZENIA GRANICZNE

Ważniejsze znaczenie – szczególnie w zastosowaniach – mają integralne twierdzenia graniczne, do których należą prawa wielkich liczb rozważone poniżej oraz szereg twierdzeń granicznych, z których do najważniejszych należy:

CENTRALNE TWIERDZENIE GRANICZNE LINDEBERGA-LEVY'EGO. *Jeżeli $\{X_n\}$ jest losowym ciągiem niezależnych zmiennych o jednakowym rozkładzie, o wartości przeciętnej α_1 i skończonej wariancji $\sigma^2>0$, to ciąg (F_n) dystrybuant standaryzowanych średnich arytmetycznych $\bar{X}_n$ (albo – co na jedno wychodzi – standaryzowanych sum $\sum_{i=1}^{n} X_i$)*

$$Y_n=\frac{\bar{X}_n-\alpha_1}{\frac{\sigma}{\sqrt{n}}}=\frac{\sum_{i=1}^{n} X_i-n\alpha_1}{\sigma\sqrt{n}} \tag{6.1.1}$$

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

$$\lim_{n\to\infty} F_n(y)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} \exp(-\tfrac{1}{2}t^2)\,dt \equiv \Phi(y). \tag{6.1.2}$$

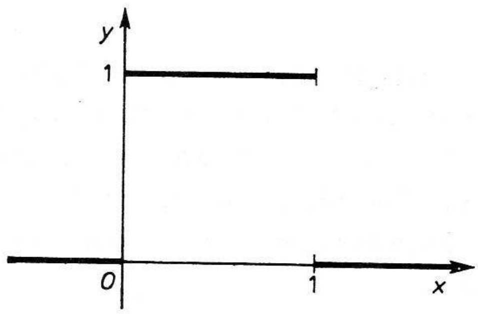
Dowód można znaleźć np. w [12]. Należy tutaj podkreślić fakt, że zmienne losowe mogą mieć zarówno rozkład dyskretny jak i typu ciągłego.

Ze wzoru (6.1.2) wynika, że dla dużych n (w praktyce rzędu kilkunastu, co oczywiście zależy od żądanego przybliżenia) można stosować wzór przybliżony

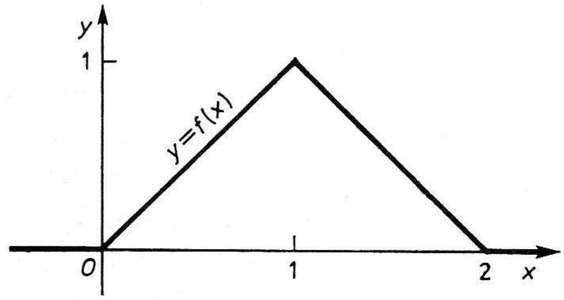
$$P(y_1 < Y_n \leqslant y_2) \approx \Phi(y_2) - \Phi(y_1), \tag{6.1.3}$$

gdzie Φ oznacza dystrybuantę rozkładu $N(0, 1)$.

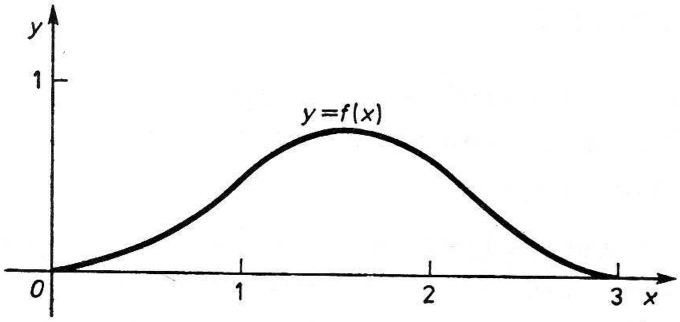
Rysunki 6.1 - 6.4 przedstawiają wykresy gęstości sumy n niezależnych zmiennych losowych o jednakowym rozkładzie równomiernym na przedziale $\langle 0, 1\rangle$ dla $n = 1, 2, 3, 4$.



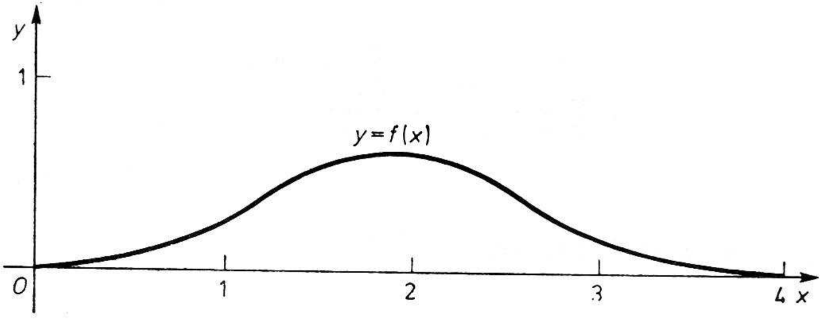
Rys. 6.1. Gęstość zmiennej losowej Y_1 o rozkładzie równomiernym skoncentrowanym na $\langle 0, 1\rangle$



Rys. 6.2. Gęstość f sumy $Y_2 = X_1 + X_2$ niezależnych zmiennych losowych o tym samym rozkładzie równomiernym na $\langle 0, 1\rangle$



Rys. 6.3. Gęstość f sumy $Y_3 = X_1 + X_2 + X_3$ niezależnych zmiennych losowych o tym samym rozkładzie równomiernym na $\langle 0, 1\rangle$



Rys. 6.4. Gęstość f sumy $Y = X_1 + X_2 + X_3 + X_4$ niezależnych zmiennych losowych o tym samym rozkładzie równomiernym na $\langle 0, 1\rangle$

ZADANIE 6.1. Losowy błąd pomiaru pewnej wielkości ma rozkład o wartości przeciętnej $\alpha_1 = 0$ (brak błędu systematycznego) i odchyleniu standardowym 0,08. Obliczyć prawdopodobieństwo, że błąd średniej arytmetycznej 100 pomiarów nie przekroczy (co do wartości bezwzględnej) 0, 1.

Rozwiązanie. Oznaczamy losowy błąd przez X_i, $i=1, \ldots, 100$. Na podstawie przybliżonego wzoru (6.1.3) mamy

$$P(|\bar{X}_{100}|<0{,}1)=P\left(-\frac{0{,}1}{0{,}08}<Y_{100}<\frac{0{,}1}{0{,}08}\right)\approx 2\Phi(1{,}25)-1=0{,}7888.$$

Stosując twierdzenie Lindeberga-Levy'ego do ciągu (X_n) niezależnych zmiennych losowych o tym samym rozkładzie zero-jedynkowym (2.7.5) i oznaczając $S_n=\sum_{k=1}^{n} X_k$, otrzymujemy:

INTEGRALNE TWIERDZENIE MOIVRE'A-LAPLACE'A. *Jeśli (S_n) jest ciągiem zmiennych losowych o rozkładzie dwumianowym (2.7.7) z parametrami (n, p), $0<p<1$ (a więc o wartości przeciętnej $ES_n=np$ i wariancji $\operatorname{Var} S_n=npq$) oraz Y_n jest ciągiem standaryzowanych zmiennych losowych:*

$$Y_n=\frac{S_n-np}{\sqrt{npq}},$$

to dla każdej pary wartości $y_1<y_2$ zachodzi wzór:

$$\lim_{n\to\infty} P\left(y_1<\frac{S_n-np}{\sqrt{npq}}<y_2\right)=\Phi(y_2)-\Phi(y_1). \tag{6.1.4}$$

6.1.1. Zadania rozwiązane.

ZADANIE 6.2. Prawdopodobieństwo, że w czasie T przestanie świecić jedna żarówka jest równe 0,1. Obliczyć prawdopodobieństwo, że w czasie T spośród 100 przestanie świecić od 7 do 19 żarówek przy założeniu, że żarówki przepalają się niezależnie.

Rozwiązanie. Niech S_{100} będzie liczbą żarówek spośród stu, które w ciągu czasu T przestały świecić: $ES_{100}=np=10$, $D^2S_{100}=npq=10\cdot 0{,}9=9$.

Korzystamy ze wzoru (6.1.4). Ponieważ jednak dla zmiennej losowej ciągłej X zachodzi np. równość $P(a<X<b)=P(a\leqslant X\leqslant b)$, która dla zmiennych skokowych ogólnie nie zachodzi, więc w przypadku rozkładu dwumianowego (a więc skokowego) przy a, b całkowitych nieujemnych postępuje się zazwyczaj następująco:

$$P(a\leqslant S_n\leqslant b)=P(a-0{,}5<S_n<b+0{,}5)=P\left(\frac{a-0{,}5-np}{\sqrt{npq}}<Y_n<\frac{b+0{,}5-np}{\sqrt{npq}}\right)=$$

$$=\Phi\left(\frac{b+0{,}5-np}{\sqrt{npq}}\right)-\Phi\left(\frac{a-0{,}5-np}{\sqrt{npq}}\right).$$

W zadaniu mamy:

$$P(7 \leqslant S_{100} \leqslant 19) = P(6{,}5 < S_{100} < 19{,}5) = P\left(\frac{6{,}5-10}{3} < Y_{100} < \frac{19{,}5-10}{3}\right) \approx$$

$$\approx \Phi\left(\frac{9{,}5}{3}\right) - \Phi\left(-\frac{3{,}5}{3}\right) = \Phi(3{,}17) + \Phi(1{,}17) - 1 = 0{,}8783.$$

ZADANIE 6.3. W centrali telefonicznej znajduje się n linii działających niezależnie. Prawdopodobieństwo, że dowolna ustalona linia jest zajęta, jest równe 0,1. Jakie powinno być n, aby prawdopodobieństwo tego, że co najmniej 7 linii jest zajętych było równe 0,95?

Rozwiązanie. Liczba linii zajętych jest zmienną losową S_n o rozkładzie dwumianowym z parametrami: n, $p=0{,}1$. Korzystając z tw. Moivre'a-Laplace'a dobieramy n tak, aby zachodziła równość

$$P(S_n \geqslant 0{,}07n) = 0{,}95.$$

Zauważmy, że $ES_n = 0{,}1n$ oraz $D^2X_n = 0{,}1 \cdot 0{,}9n$, skąd odchylenie standardowe: $\sigma = 0{,}3\sqrt{n}$. Standaryzując zmienną losową S_n, otrzymamy

$$P\left(Y_n \geqslant \frac{0{,}07n - 0{,}1n}{0{,}3\sqrt{n}}\right) = 0{,}95,$$

a po przejściu do granicznego rozkładu normalnego:

$$1 - \Phi(-0{,}1\sqrt{n}) \approx 0{,}95, \qquad \Phi(0{,}1\sqrt{n}) \approx 0{,}95.$$

Dla wartości dystrybuanty rozkładu $N(0, 1)$ równej 0,95 odczytujemy z tablic liczbową wartość argumentu

$$0{,}1\sqrt{n} \approx 1{,}64, \qquad n \approx 268{,}96.$$

W centrali telefonicznej powinny być co najmniej 269 linii.

6.2. PRAWA WIELKICH LICZB

Niech (X_n) będzie ciągiem zmiennych losowych, dla których $EX_i = \mu_i < \infty$ dla $i \in N$ oraz

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad E\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} EX_i.$$

Jeżeli dla losowego ciągu (X_n) i dla dowolnego $\varepsilon > 0$

$$\lim_{n\to\infty} P(|\bar{X}_n - E\bar{X}_n| \geqslant \varepsilon) = 0, \tag{6.2.1}$$

to mówimy, że dla tego ciągu zachodzi *słabe prawo wielkich liczb*. Mówimy również, że $\bar{X}_n - E\bar{X}_n \to 0$ *według prawdopodobieństwa* (*stochastycznie, według miary P*), co zapisujemy $\bar{X}_n - E\bar{X}_n \xrightarrow{P} 0$. W przypadku, gdy przy tym samym założeniu

$$P[\lim_{n\to\infty} (\bar{X}_n - E\bar{X}_n) = 0] = 1, \tag{6.2.2}$$

wtedy mówimy, że dla losowego ciągu (X_n) zachodzi *mocne prawo wielkich liczb*, a zbieżność $\bar{X}_n - E\bar{X}_n \to 0$ wyrażoną wzorem (6.2.2) nazywamy *zbieżnością z prawdopodobieństwem* 1 (*prawie na pewno, prawie wszędzie P*), co będziemy zapisywali $\bar{X}_n - E\bar{X}_n \xrightarrow{p=1} 0$.

Pierwsze prawo wielkich liczb (słabe) dla ciągu (X_n) zmiennych losowych o tych samych rozkładach zero-jedynkowych udowodnił Bernoulli (1713 r.). Znacznie silniejsze wyniki osiągnął Kołmogorow: Mocne prawo wielkich liczb Kołmogorowa (I). Dla losowego ciągu (X_n) o wspólnie ograniczonych wariancjach $(D^2X_i \leqslant C,\ i \in N)$ zachodzi mocne prawo wielkich liczb.

ZADANIE 6.4. Wykazać, że jeśli w ciągu n niezależnych doświadczeń prawdopodobieństwo zajścia zdarzenia A w i-tym doświadczeniu jest równe p_i, to

$$P\left[\lim_{n\to\infty}\left(\frac{S_n}{n} - \frac{1}{n}\sum_{i=1}^{n} p_i\right) = 0\right] = 1,$$

gdzie S_n oznacza liczbę zajść zdarzenia A w pierwszych n doświadczeniach.

Rozwiązanie. $S_n = \sum_{i=1}^{n} X_i$, gdzie $P(X_i=1)=p_i$, $P(X_i=0)=1-p_i$, skąd $EX_i = p_i$, $D^2X_i = p_i(1-p_i) \leqslant \frac{1}{4}$, a więc są spełnione założenia twierdzenia Kołmogorowa, prawdziwy jest więc wzór (6.2.2). Wykazaliśmy więc, że mocne prawo wielkich liczb zachodzi m. in. dla schematu Poissona.

W dalszych badaniach okazało się, że mocne prawo wielkich liczb może zachodzić bez założenia o ograniczoności wariancji, a nawet bez jej istnienia, przy przyjęciu jednak założenia o jednakowym rozkładzie i niezależności zmiennych.

MOCNE PRAWO WIELKICH LICZB KOŁMOGOROWA (II). *Warunkiem koniecznym i wystarczającym zachodzenia mocnego prawa wielkich liczb dla ciągu (X_n) niezależnych zmiennych losowych o jednakowym rozkładzie jest istnienie skończonej wartości oczekiwanej $E(X_i) = m$ dla $i \in N$.*

Dowód znajduje się w [9].

ZADANIE 6.5. Niezależne zmienne losowe X_i, $i \in N$, mają ten sam rozkład prawdopodobieństwa $P(X_i = 2^k) = 0{,}8 \cdot 0{,}2^k$ dla $k \in N_0$, $i \in N$. Czy dla tego ciągu zachodzi mocne prawo wielkich liczb Kołmogorowa?

Rozwiązanie. Obliczamy $EX_i = \sum_{k=0}^{\infty} 2^k \cdot 0{,}8 \cdot 0{,}2^k = 0{,}8 \sum_{k=0}^{\infty} 0{,}4^k = 0{,}8 \dfrac{1}{1-0{,}4} = \frac{4}{3}$. A więc zachodzi wzór

$$P[\lim_{n\to\infty}(\bar{X}_n - \tfrac{4}{3}) = 0] = 1.$$

6.3. ZADANIA DO ROZWIĄZANIA

6.6. W pewnym magazynie znajduje się towar o przeciętnej wadliwości 0,1. Korzystając z twierdzenia Moivre'a Laplace'a, obliczyć prawdopodobieństwo, że wśród losowo wybranych 100 sztuk towaru procent sztuk wadliwych różni się od 10 o co najwyżej 0,15.

6.7. W urnie znajduje się 36 kul białych i 64 czarnych. Losujemy kule po jednej ze zwracaniem. Ile losowań należy dokonać, aby prawdopodobieństwo tego, że częstość otrzymywania kuli białej różni się od 0,36 o co najmniej 0,12 było równe 0,1?

6.8. W pewnej grupie ludzi co dziesiąty człowiek jest daltonistą. Obliczyć prawdopodobieństwo, że wśród 100 losowo wybranych ludzi będzie od 5 do 12 daltonistów.

6.9. W zajezdni znajduje się 200 autobusów. Prawdopodobieństwo, że losowo wybrany autobus jest sprawny do jazdy wynosi 0,7. Obliczyć prawdopodobieństwo, że w losowo wybranej chwili co najmniej 160 autobusów jest sprawnych.

6.10. W pewnej szkole uczy się 500 dzieci. Prawdopodobieństwo, że losowo wybrany uczeń ma co najmniej jedną dwójkę jest równe 0,1. Obliczyć prawdopodobieństwo, że w tej szkole liczba dzieci, które mają co najmniej jedną dwójkę różni się od 50 o co najwyżej 10.

6.11. Urządzenie składa się z n elementów. Urządzenie pracuje, jeśli co najmniej 70% elementów jest sprawnych. Prawdopodobieństwo awarii jednego elementu jest równe 0,2. Jak duża powinna być liczba elementów, aby z prawdopodobieństwem 0,95 urządzenie pracowało?

6.12. Strzelec trafia do celu z prawdopodobieństwem 0,5. Jaką liczbę strzałów musi oddać, aby prawdopodobieństwo tego, że częstość trafienia do celu różni się od 0,5 o co najwyżej 0,1 było równe 0,95?

6.13. Niech $X_1, \ldots, X_{100}$ będą niezależnymi zmiennymi losowymi o tym samym rozkładzie beta o gęstości $f(x) = 12x(1-x)^2$ dla $0<x<1$. Obliczyć prawdopodobieństwo: $P(20 \leqslant \sum\limits_{i=1}^{100} X_i \leqslant 30)$.

6.14. Niech $X_1, \ldots, X_{200}$ będą niezależnymi zmiennymi losowymi o rozkładzie geometrycznym: $P(X_k = i) = (\frac{1}{2})^i$, gdzie $i \in N, k = 1, \ldots, 200$. Obliczyć prawdopodobieństwo, że średnia arytmetyczna tych zmiennych przyjmuje wartości z przedziału $\langle 1, 4 \rangle$.

6.15. Niech $X_1, \ldots, X_k, \ldots$ będą niezależnymi zmiennymi losowymi o jednakowym rozkładzie: $P(X_k = i) = \frac{2}{3}(\frac{1}{3})^{i-1}$, $i \in N$, $k \in N$. Wykazać że dla zmiennej losowej $Y_n = \frac{1}{n} \sum\limits_{k=1}^{n} X_k$ zachodzi mocne prawo wielkich liczb.

6.16. Niezależne zmienne losowe $X_1, \ldots, X_k, \ldots$. podlegają temu samemu rozkładowi prawdopodobieństwa: $P\left(X_k = \frac{(-1)^i}{i}\right) = \frac{1}{2^i}$, $i, k \in N$. Sprawdzić czy w tym przypadku zachodzi twierdzenie Kołmogorowa.

6.17. Niech $X_1, \ldots, X_k, \ldots$ będą niezależnymi zmiennymi losowymi o rozkładzie Poissona: $P(X_k = i) = \frac{e^{-1}}{i!}$, $i \in N_0$. Zbadać czy dla średniej arytmetycznej $Y_n = \frac{1}{n} \sum\limits_{k=1}^{n} X_k$ zachodzi a) mocne prawo wielkich liczb, b) centralne twierdzenie graniczne.

6.18. Dodajemy 10 000 liczb, każda z nich jest zaokrąglona z dokładnością do 10^{-m}. Zakładając, że błędy zaokrąglenia są niezależnymi zmiennymi losowymi o rozkładzie równomiernym w przedziale $(-\frac{1}{2} \cdot 10^{-m}, \frac{1}{2} \cdot 10^{-m})$, wyznaczyć przedział, w którym z prawdopodobieństwem 0,99 będzie się zawierał błąd sumy.

6.19. Partia towaru zawiera 20% braków. Obliczyć prawdopodobieństwo, że w próbce o liczności: a) $n=100$ sztuk, b) $n=400$ sztuk, c) $n=1600$ sztuk. Stosunek $\frac{k}{n}$ (gdzie k jest liczbą braków) różni się od wadliwości p partii nie więcej niż o 0,02. Przedstawić obliczone prawdopodobieństwo za pomocą wykresów odpowiednich gęstości rozkładu normalnego.

Odpowiedzi

6.6. $P\left(\left|\dfrac{k}{n}-0{,}1\right|\leqslant 0{,}15\right)\approx 2\Phi(5)-1\approx 1.$

6.7. Należy wykonać co najmniej $n\geqslant 40$ losowań.

6.8. $P(5\leqslant X\leqslant 12)\approx\Phi(0{,}67)+\Phi(1{,}67)-1\approx 0{,}7011.$

6.9. $P(X\geqslant 160)\approx 1-P(X<160)=1-\Phi\left(\dfrac{20}{\sqrt{42}}\right)\approx 0{,}00097.$

6.10. $P(|X-50|\leqslant 10)\approx 2\Phi\left(\dfrac{10}{\sqrt{45}}\right)-1\approx 0{,}8788.$

6.11. Urządzenie powinno się składać z co najmniej $n=44$ elementów.

6.12. $n\geqslant 96.$

6.13. $P(20\leqslant \sum\limits_{i=1}^{100} X_i\leqslant 30)\approx 0,\quad \bigwedge\limits_{i=1,\dots,100} EX_i=\frac{2}{5},\quad D^2X_i=\frac{1}{25}$ ((2.8.21)).

6.14. $P\left(1\leqslant\dfrac{1}{200}\sum\limits_{i=1}^{200}X_i\leqslant 4\right)\approx 1,\quad \bigwedge\limits_{i=1,\dots,200} EX_i=2,\quad D^2X_i=2$ ((2.6.30) i (2.6.31)).

6.15. $E(X_k)=\frac{3}{2}<\infty$ dla $k\in N$.

6.16. Tak, ponieważ $E(X_k)=\sum\limits_{i=1}^{\infty}\dfrac{(-1)^i}{i2^i}=-\ln\frac{3}{2}$ dla $k\in N$.

6.17. a) Tak, ponieważ $\bigwedge\limits_{k\in N}E(X_k)=1$;

b) tak, ponieważ $\bigwedge\limits_{k\in N}E(X_k)=1$ i $D^2(X_k)=1$.

6.18. Niech X_k, $k=1,\dots,10\,000$ będą błędami zaokrągleń poszczególnych liczb, wtedy dla rozkładu równomiernego w $(-\frac{1}{2}\cdot 10^{-m}, \frac{1}{2}\cdot 10^{-m})$ $EX_k=0$, $D^2X_k=\dfrac{10^{-2m}}{12}$ ((2.8.6)). Oznaczamy $Y=\sum\limits_{k=1}^{10000}X_k$, wówczas $EY=0$, $D^2Y=10000\cdot\dfrac{10^{-2m}}{12}$, a odchylenie standardowe $\sigma_Y=\dfrac{100\cdot 10^{-m}}{2\sqrt{3}}$, skąd i z centralnego tw. granicznego mamy: $P\left(\left|\dfrac{X-0}{\sigma_Y}\right|<y\right)=0{,}99$, z tablic rozkładu $N(0,1)$ odczytujemy $y=2{,}58$, a więc błąd sumy $Y\in\left(-2{,}58\dfrac{10^{2-m}}{2\sqrt{3}},\ +2{,}58\dfrac{10^{2-m}}{2\sqrt{3}}\right)$, a w przybliżeniu $Y\in(-0{,}75\cdot 10^{2-m}, 0{,}75\cdot 10^{2-m})$.