

VI. TWIERDZENIA GRANICZNE

6.1. Wprowadzenie

Twierdzenia graniczne dotyczą własności granicznych ciągów zmiennych losowych i dzielą się na:

- *twierdzenia lokalne* – opisują zbieżności ciągu funkcji prawdopodobieństwa w przypadku ciągu $\{X_n\}$ zmiennych losowych typu skokowego lub zbieżności ciągu gęstości prawdopodobieństwa w przypadku ciągu $\{X_n\}$ zmiennych losowych typu ciągłego,
- *twierdzenia integralne* – dotyczą zbieżności ciągu dystrybuant zmiennych losowych X_n do pewnej dystrybuanty granicznej (jest to tzw. *zbieżność słaba*).

Definicja 6.1. Mówimy, że ciąg $\{F_n(x)\}$ dystrybuant zmiennych losowych X_n jest *słabo zbieżny* do funkcji $F(x)$, jeżeli funkcją tą jest taka dystrybuanta, która w każdym swym punkcie ciągłości spełnia warunek

$$\lim_{n \rightarrow \infty} F_n(x) = F(x).$$

Uwagi:

- 1) Funkcją graniczną może być tylko dystrybuanta.
- 2) Nie wymaga się, aby ciąg dystrybuant $\{F_n(x)\}$ był zbieżny do funkcji $F(x)$ w punktach nieciągłości tej dystrybuanty.

Przykład 6.1. Rozpatrzmy ciąg zmiennych losowych o rozkładzie jednopunktowym

$$P(X_n = k) = \begin{cases} 1 & \text{dla } k = n, \\ 0 & \text{dla } k \neq n. \end{cases}$$

Zbadajmy zbieżność ciągu dystrybuant tych zmiennych losowych.

Dystrybuanty mają postać

$$F_n(x) = \begin{cases} 0 & \text{dla } x \leq n, \\ 1 & \text{dla } x > n. \end{cases}$$

Widać, że $\lim_{n \rightarrow \infty} F_n(x) = 0$ dla dowolnego x . Ponieważ funkcja graniczna jest

stałe równa zero, więc nie może być dystrybuantą, a tym samym podany ciąg dystrybuant $\{F_n(x)\}$ nie jest zbieżny w rozpatrywanym sensie.

Twierdzenia integralne rozważające ciągi sum niezależnych zmiennych losowych o dystrybuancie, będącej dystrybuantą rozkładu normalnego, nazywamy *twierdzeniami centralnymi*.

6.2. Twierdzenia lokalne

Jedno z twierdzeń lokalnych już poznaliśmy przy okazji omawiania schematu Poissona. Było ono sformułowane w języku zdarzeń i dotyczyło serii doświadczeń według schematu Bernoulliego. Podamy to twierdzenie raz jeszcze używając pojęcia zmiennej losowej.

Twierdzenie 6.1. *Jeżeli ciąg $\{X_n\}$ jest ciągiem zmiennych losowych o rozkładzie Bernoulliego*

$$P(X_n = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad 0 < p < 1, \quad k = 0, 1, \dots, n,$$

to przy zachowaniu warunku $\lim_{\substack{n \rightarrow \infty \\ p \rightarrow 0}} np = \lambda > 0$ jest spełniona równość

$$\lim_{n \rightarrow \infty} P(X_n = k) = P(X = k),$$

gdzie $P(X = k)$ oznacza funkcję prawdopodobieństwa zmiennej losowej o rozkładzie Poissona z parametrem λ .

Wiemy już, że na podstawie tego twierdzenia można stosować wzór przybliżony

$$P(X_n = k) \approx P(X = k),$$

o ile tylko liczba n jest dostatecznie duża, a liczba λ dostatecznie mała (w praktyce $n \geq 100$ oraz $\lambda \leq 20$). Jeżeli warunek na parametr λ nie jest spełniony, to stosujemy inny wzór przybliżony:

$$P(X_n = k) \approx \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{k-m}{\sigma}\right)^2\right],$$

gdzie $m = np$, $\sigma = \sqrt{npq}$, $q = 1 - p$. Stosowanie tego wzoru wynika z poniższego twierdzenia.

Twierdzenie 6.2 (Moivre'a-Laplace'a). *Jeżeli ciąg $\{Y_n\}$ jest ciągiem standaryzowanych zmiennych losowych X_n o rozkładzie dwumianowym (Bernoulliego), to*

$$\lim_{n \rightarrow \infty} \sqrt{npq} P(Y_n = y_n) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}y^2\right),$$

gdzie $y_n = \frac{k - np}{\sqrt{npq}}$ oznacza wartość, jaką przyjmuje zmienna losowa Y_n , gdy

zmienna losowa X_n przyjmuje wartość k taką, że gdy $n \rightarrow \infty$, to również $k \rightarrow \infty$ w taki sposób, że $(n - k) \rightarrow \infty$ oraz

$$\lim_{n \rightarrow \infty} Y_n = Y.$$

Dowód tego twierdzenia (żmudny) pomijamy. ■

Przykład 6.2. Rzucamy 144 razy kostką. Obliczyć prawdopodobieństwo, że otrzymamy 25 razy piątkę.

Mamy $p = 1/6$, $q = 1 - p = 5/6$, $k = 25$, $n = 144$, skąd $np = 24$, $npq = 20$. Zatem

$$\begin{aligned}
 P(X_{144} = 25) &\approx \frac{1}{\sqrt{2\pi \cdot 20}} \exp \left[-\frac{1}{2} \left(\frac{25-24}{\sqrt{20}} \right)^2 \right] = \frac{1}{\sqrt{20}} \cdot \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2} (0,2236)^2 \right] \\
 &= \frac{1}{\sqrt{20}} \cdot f(0,2236),
 \end{aligned}$$

gdzie $f(x)$ oznacza gęstość rozkładu normalnego $N(0, 1)$. Z tablic otrzymujemy $f(0,2236) \approx 0,3894$, skąd

$$P(X_{144} = 25) \approx \frac{1}{\sqrt{20}} 0,3894 \approx 0,0871.$$

6.3. Twierdzenia integralne

Twierdzenie 6.3 (twierdzenie centralne Lindeberga-Lévy'ego). Niech $\{X_i\}$ oznacza ciąg zmiennych losowych, które są

- niezależne,
- o tym samym rozkładzie,
- o skończonej wartości oczekiwanej $E(X_i) = m < \infty$,
- o skończonej i większej od zera wariancji $0 < D^2(X_i) = \sigma^2 < \infty$.

Niech zmienne losowe Y_n będą sumami zmiennych losowych X_i , tj.

$$Y_n = \sum_{i=1}^n X_i.$$

Jeżeli $\{Z_n\}$ oznacza ciąg standaryzowanych zmiennych losowych Y_n , tj.

$$Z_n = \frac{Y_n - E(Y_n)}{D(Y_n)},$$

to jest spełniona równość

$$\lim_{n \rightarrow \infty} F_n(z) = \Phi(z),$$

gdzie z oznacza dowolną liczbę rzeczywistą, funkcja $\Phi(z)$ jest dystrybuantą rozkładu normalnego, a ciąg $\{F_n(z)\}$ jest ciągiem dystrybuant odpowiadających zmiennym losowym Z_n .

Dowód tego twierdzenia pomijamy (jest trochę skomplikowany). ■

Z wzoru podanego w powyższym twierdzeniu wynika przydatna w zastosowaniach zależność przybliżona. Ponieważ dla dowolnych liczb rzeczywistych z_1 i z_2 mamy

$$\lim_{n \rightarrow \infty} P(z_1 \leq Z_n < z_2) = \lim_{n \rightarrow \infty} [F_n(z_2) - F_n(z_1)] = \Phi(z_2) - \Phi(z_1),$$

więc dla dużych wartości n otrzymujemy

$$P(z_1 \leq Z_n \leq z_2) \approx \Phi(z_2) - \Phi(z_1). \quad (*)$$

Nierówność $z_1 \leq Z_n < z_2$ jest równoważna nierówności

$$z_1 \leq \frac{Y_n - nm}{\sigma\sqrt{n}} < z_2,$$

a ta z kolei – nierówności $nm + z_1\sigma\sqrt{n} \leq Y_n < nm + z_2\sigma\sqrt{n}$. Wzór (*) możemy zatem zapisać w postaci

$$P(y_1 \leq Y_n \leq y_2) \approx \Phi(z_2) - \Phi(z_1),$$

gdzie

$$y_1 = nm + z_1\sigma\sqrt{n}, \quad y_2 = nm + z_2\sigma\sqrt{n}.$$

Jeżeli przedział $[y_1, y_2]$ wartości zmiennej losowej Y jest symetryczny ze względu na wartość nm , tzn. gdy $z_1 = -z_2$, to

$$y_1 = nm - z_2\sigma\sqrt{n}, \quad y_2 = nm + z_2\sigma\sqrt{n}$$

i powyższy wzór przyjmie następującą postać:

$$\begin{aligned} P(y_1 \leq Y_n \leq y_2) &= P\left(\frac{y_1 - nm}{\sigma\sqrt{n}} \leq \frac{Y_n - nm}{\sigma\sqrt{n}} \leq \frac{y_2 - nm}{\sigma\sqrt{n}}\right) \\ &= P(-z_2 \leq Z_n \leq z_2) \approx \Phi(z_2) - \Phi(-z_2) = 2\Phi(z_2) - 1. \end{aligned}$$

Przykład 6.3. Zmienne losowe X_i ($i = 1, 2, \dots$) są niezależne i mają jednakowe rozkłady

$$P(X_i = k) = 0,2 \text{ dla } k = 1, 2, 3, 4, 5.$$

Znaleźć prawdopodobieństwo, że zmienna $Y_{100} = \sum_{i=1}^{100} X_i$ przyjmie wartość większą od 320.

Zauważmy, że zmienne losowe X_i mają jednakowe wartości oczekiwane

$$m = E(X_i) = \frac{1}{5} \sum_{k=1}^5 k = 3$$

oraz wariancje

$$\sigma^2 = D^2(X_i) = \frac{1}{5} \sum_{k=1}^5 k^2 - m^2 = 11 - 9 = 2.$$

Na podstawie twierdzenia Lindberga-Lévy'ego zmienna losowa

$$Z_{100} = \frac{\sum_{i=1}^{100} (X_i - m)}{\sqrt{\sum_{i=1}^{100} D^2(X_i)}} = \frac{Y_{100} - 300}{10\sqrt{2}}$$

ma rozkład asymptotycznie normalny, a więc

$$\begin{aligned}
 P(Y_{100} > 320) &= P(320 < Y_{100} < \infty) = P\left(\frac{320-300}{10\sqrt{2}} < Z_{100} < \infty\right) \\
 &= P(\sqrt{2} < Z_{100} < \infty) \approx 1 - \Phi(\sqrt{2}) \approx 0,0788.
 \end{aligned}$$

Przykład 6.4. Zmienne losowe X_i ($i = 1, 2, \dots$) są niezależne i mają jednakowe rozkłady o gęstości

$$f(x) = \begin{cases} \frac{1}{2} & \text{dla } 0 \leq x \leq 2, \\ 0 & \text{dla innych } x. \end{cases}$$

Obliczyć prawdopodobieństwo, że wartość sumy

$$Y_{48} = \sum_{i=1}^{48} X_i$$

będzie mniejsza od 40.

Obliczamy wartość oczekiwaną i wariancję zmiennych X_i :

$$\begin{aligned}
 E(X_i) &= \int_0^2 \frac{1}{2} x dx = \frac{1}{2} \int_0^2 x dx = \frac{1}{2} \cdot \frac{x^2}{2} \Big|_0^2 = 1, \\
 E(X_i^2) &= \int_0^2 \frac{1}{2} x^2 dx = \frac{1}{2} \int_0^2 x^2 dx = \frac{1}{2} \cdot \frac{x^3}{3} \Big|_0^2 = \frac{4}{3}, \\
 D^2(X_i) &= E(X_i^2) - (E(X_i))^2 = \frac{4}{3} - 1 = \frac{1}{3}.
 \end{aligned}$$

Zmienne X_i spełniają założenia twierdzenia Lindeberga-Lévy'ego i na jego podstawie mamy

$$\begin{aligned}
 P(Y_{48} < 40) &= P(-\infty < Y_{48} < 40) = P\left(-\infty < \frac{Y_{48} - 48 \cdot 1}{\sqrt{\frac{1}{3} \cdot 48}} < \frac{40 - 48}{4}\right) \\
 &= P(-\infty < Z_{48} < -2) \approx \Phi(-2) = 1 - \Phi(2) \approx 0,0227.
 \end{aligned}$$

Niech zmienne losowe X_i ($i = 1, 2, \dots$) będą niezależne i niech każda z nich ma rozkład zero-jedynkowy, tj.

$$P(X_i = 1) = p, \quad P(X_i = 0) = q, \quad 0 < p < 1, \quad q = 1 - p.$$

Suma

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

ma rozkład Bernoulliego o wartości oczekiwanej $E(Y_n) = np$ i wariancji $D^2(Y_n) = npq$. W tym szczególnym przypadku twierdzenie Lindeberga-Lévy'ego nazywa się *integralnym twierdzeniem Moivre'a-Laplace'a* i jest następujące:

Twierdzenie 6.4. *Jeżeli ciąg $\{F_n(z)\}$ jest ciągiem dystrybuant zmiennych losowych*

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

gdzie zmienne Y_n mają rozkład dwumianowy (Bernoulliego), to dla każdej wartości rzeczywistej z jest spełniona równość

$$\lim_{n \rightarrow \infty} F_n(z) = \Phi(z).$$

W zadaniach, w których obliczamy prawdopodobieństwo, że zmienna losowa Y_n podlegająca rozkładowi Bernoulliego przyjmuje wartości z określonego przedziału można zatem dla dużych wartości n korzystać z przybliżonego wzoru

$$P(y_1 \leq Y_n \leq y_2) \approx \Phi(z_2) - \Phi(z_1),$$

gdzie liczby

$$z_1 = \frac{y_1 - np}{\sqrt{npq}} \text{ i } z_2 = \frac{y_2 - np}{\sqrt{npq}}$$

są wartościami zmiennej losowej Z_n .

W szczególności, gdy przedział $[y_1, y_2]$ wartości zmiennej losowej Y_n jest symetryczny ze względu na wartość oczekiwaną np , tj. gdy

$$y_1 = np - z_2 \sqrt{npq}, \quad y_2 = np + z_2 \sqrt{npq},$$

to $z_1 = -z_2$ i powyższy wzór można napisać w postaci

$$\begin{aligned} P(y_1 \leq Y_n \leq y_2) &= P\left(\frac{y_1 - np}{\sqrt{npq}} \leq \frac{Y_n - np}{\sqrt{npq}} \leq \frac{y_2 - np}{\sqrt{npq}}\right) \\ &= P(-z_2 \leq Z_n \leq z_2) \approx \Phi(z_2) - \Phi(-z_2) = 2\Phi(z_2) - 1. \end{aligned}$$

Przykład 6.5. Rzucono 720 razy kostką do gry. Jakie jest prawdopodobieństwo, że liczba wyrzuconych czwórek będzie zawarta w przedziale od 100 do 150?

W jednym rzucie mamy $p = 1/6$ i $q = 5/6$. Poszczególne rzuty są niezależne, a ich liczba jest wystarczająco duża. Jeżeli przez Y oznaczymy zmienną losową przyjmującą wartości równe liczbie wyrzuconych oczek, to pytamy o prawdopodobieństwo $P(100 \leq Y \leq 150)$. Ponieważ $np = 120$ i $\sqrt{npq} = 10$, więc mamy

$$\begin{aligned} P(100 \leq Y \leq 150) &= P\left(\frac{100 - 120}{10} \leq \frac{Y - 120}{10} \leq \frac{150 - 120}{10}\right) \\ &= P(-2 \leq Z \leq 3) \approx \Phi(3) - \Phi(-2) = \Phi(3) - (1 - \Phi(2)) \\ &= \Phi(3) + \Phi(2) - 1 \approx 0,9987 + 0,9772 - 1 = 0,9759. \end{aligned}$$

Przykład 6.6. Partia towaru ma wadliwość 5%. Ilu elementową próbkę należy pobrać, aby z prawdopodobieństwem 0,99 można było twierdzić, że liczba sztuk wadliwych w próbce będzie zawarta w granicach od 4% do 6%?

Prawdopodobieństwo otrzymania sztuki wadliwej wynosi $p = 0,05$, a prawdopodobieństwo sztuki dobrej $q = 0,95$. Jeśli przez n oznaczmy liczbę elementów próby, to mamy $np = 0,05n$, $npq = 0,05 \cdot 0,95n = 0,0475n$, skąd

$$\sqrt{npq} = \sqrt{0,0475n}.$$

Niech Y oznacza zmienną losową przyjmującą wartości równe liczbie sztuk wadliwych w partii towaru. Pytamy o prawdopodobieństwo liczby sztuk wadliwych w partii towaru w granicach od $y_1 = 0,04n$ do $y_2 = 0,06n$. Oznaczając

$$Z = \frac{Y - np}{\sqrt{npq}}$$

mamy

$$\begin{aligned} P(0,04n \leq Y \leq 0,06n) &= P\left(\frac{0,04n - 0,05n}{\sqrt{0,0475n}} \leq Z \leq \frac{0,06n - 0,05n}{\sqrt{0,0475n}}\right) \\ &= P\left(-\frac{0,01\sqrt{n}}{\sqrt{0,0475}} \leq Z \leq \frac{0,01\sqrt{n}}{\sqrt{0,0475}}\right) \approx 2\Phi\left(\frac{0,01\sqrt{n}}{\sqrt{0,0475}}\right) - 1. \end{aligned}$$

Z założenia prawdopodobieństwo to jest równe 0,99, czyli

$$2\Phi\left(\frac{0,01\sqrt{n}}{\sqrt{0,0475}}\right) - 1 = 0,99,$$

a więc

$$\Phi\left(\frac{0,01\sqrt{n}}{\sqrt{0,0475}}\right) = 0,995.$$

Z tablic odczytujemy, że $\frac{0,01\sqrt{n}}{\sqrt{0,0475}} \approx 2,57$, skąd po wykonaniu działań mamy $n = 3137$.

Zadania

1. Wykonano 100 serii niezależnych doświadczeń. W każdej serii wartość oczekiwana uzyskania sukcesu była równa 2, a odchylenie standardowe wynosiło 1,4. Znaleźć przybliżoną wartość prawdopodobieństwa, że w tych 100 seriach doświadczeń liczba sukcesów będzie zawarta w przedziale od 186 do 214.
2. Pewien towar ma wadliwość 10%. Zakupiono 900 sztuk tego towaru. Obliczyć prawdopodobieństwo, że liczba znalezionych w tej partii sztuk wadliwych będzie zawierać się w granicach od 9% do 12%.

6.4. Zbieżność stochastyczna i prawa wielkich liczb

Definicja 6.2. Mówimy, że ciąg zmiennych losowych $\{X_n\}$ jest *zbieżny stochastycznie* lub *zbieżny według prawdopodobieństwa* do zmiennej losowej X , co zapisujemy

$$X_n \xrightarrow{st.} X \text{ lub } X_n \xrightarrow{w.pr.} X,$$

jeśli ciąg prawdopodobieństw bezwzględnego odchylenia zmiennej losowej X_n od zmiennej losowej X o nie mniej niż wielkość $\varepsilon > 0$ ma granicę równą zero, tj.

$$\lim_{n \rightarrow \infty} P(|X_n - X| \geq \varepsilon) = 0,$$

co inaczej można zapisać w postaci

$$\lim_{n \rightarrow \infty} P(|X_n - X| < \varepsilon) = 1.$$

Definicja 6.3. Mówimy, że ciąg zmiennych losowych jest *stochastycznie zbieżny do zera*, co zapisujemy

$$X_n \xrightarrow{st.} 0,$$

jeśli dla dowolnego $\varepsilon > 0$ jest spełniona zależność

$$\lim_{n \rightarrow \infty} P(|X_n| \geq \varepsilon) = 0$$

lub inaczej

$$\lim_{n \rightarrow \infty} P(|X_n| < \varepsilon) = 1.$$

Uwaga: Zbieżności stochastycznej nie należy mylić ze zbieżnością zwykłą. W przypadku zbieżności stochastycznej do zera ciągu zmiennych losowych $\{X_n\}$ nie oznacza to, że dla dowolnego $\varepsilon > 0$ istnieje taka liczba N , że dla każdej wartości $n > N$ zachodzi nierówność $|X_n| < \varepsilon$. Zbieżność stochastyczna do zera pozwala tylko stwierdzić, że prawdopodobieństwo zajścia zdarzenia $|X_n| < \varepsilon$ rośnie do jedności, gdy $n \rightarrow \infty$.

Przykład 6.7. Ciąg $\{X_n\}$ ($n = 0, 1, 2, \dots$) Zmiennych losowych przyjmuje wartości $a > 0$ na odcinku $[0, 1/2^n]$ i 0 na odcinku $(1/2^n, 1]$ z prawdopodobieństwami równymi długości odcinków. Wykazać, że ciąg tych zmiennych losowych jest stochastycznie zbieżny do zera.

Z definicji ciągu $\{X_n\}$ wynika, że

$$P(X_n = a) = \frac{\frac{1}{2^n}}{1} = \frac{1}{2^n}$$

(stosunek długości odcinka, na którym zmienna losowa przyjmuje wartość a do długości całego odcinka). Zauważmy, że

$$0 \leq P(X_n \geq \varepsilon) \leq P(X_n = a),$$

bo dla $\varepsilon > a$ mamy $P(X_n \geq \varepsilon) = 0$, gdyż zmienna losowa X_n nie przyjmuje wartości większych od a , a dla $\varepsilon \leq a$ jest $P(X_n \geq \varepsilon) = P(X_n = a)$, gdyż przyjęcie przez zmienną losową X_n wartości nie większej od tak przyjętego ε jest równoważne z przyjęciem przez nią wartości równej a . Przechodząc w podanych nierównościach do granicy mamy

$$0 \leq \lim_{n \rightarrow \infty} P(X_n \geq \varepsilon) \leq \lim_{n \rightarrow \infty} P(X_n = a) = \lim_{n \rightarrow \infty} \frac{1}{2^n} = 0.$$

Zatem

$$\lim_{n \rightarrow \infty} P(X_n \geq \varepsilon) = 0,$$

co w myśl definicji 6.3 oznacza, że ciąg $\{X_n\}$ zmiennych losowych jest stochastycznie zbieżny do zera.

Definicja 6.4. *Prawem wielkich liczb* nazywamy każde twierdzenie, które rozważa zbieżność stochastyczną do zera pewnego ciągu zmiennych losowych.

Prawa wielkich liczb zalicza się twierdzeń integralnych. Dalej podajemy najważniejsze z nich (dowody pomijamy).

Definicja 6.5. Mówimy, że ciąg zmiennych losowych $\{X_n\}$ spełnia *warunek Markowa*, jeżeli zmienne te mają skończone wariancje σ_n^2 oraz

$$\lim_{n \rightarrow \infty} \sigma_n^2 = 0.$$

Twierdzenie 6.5 (Czebyszewa). *Jeżeli ciąg $\{X_n\}$ jest ciągiem zmiennych losowych, dla którego jest spełniony warunek Markowa, to ciąg zmiennych losowych $\{X_n - E(X_n)\}$ jest stochastycznie zbieżny do zera, tzn.*

$$\lim_{n \rightarrow \infty} P(|X_n - E(X_n)| \geq \varepsilon) = 0.$$

Twierdzenie 6.6 (Chinczyna). *Jeżeli ciąg $\{X_i\}$ jest ciągiem niezależnych zmiennych losowych o jednakowym rozkładzie i skończonej wartości oczekiwanej $E(X_i) = m$, to ciąg zmiennych losowych*

$$U_n = \frac{1}{n} \sum_{i=1}^n X_i$$

jest stochastycznie zbieżny do m .

Twierdzenie 6.7 (prawo wielkich liczb Bernoulliego). *Jeżeli ciąg $\{X_i\}$ jest ciągiem niezależnych zmiennych losowych o wspólnym rozkładzie zero-jedynkowym, tzn.*

$$P(X_i = 1) = p, \quad P(X_i = 0) = 1 - p$$

oraz $Y_n = \sum_{i=1}^n X_i$, to ciąg $\left\{U_n = \frac{Y_n}{n}\right\}$ jest stochastycznie zbieżny do p .

Przykład 6.8. Dany jest ciąg zmiennych losowych $\{X_n\}$, przy czym zmienna X_n może przyjmować wartości $-na$, 0 oraz na z prawdopodobieństwami równymi odpowiednio

$$\frac{1}{2^n}, 1 - \frac{1}{2^{n-1}} \text{ oraz } \frac{1}{2^n}.$$

Zbadać, czy do tego ciągu zmiennych losowych stosuje się prawo wielkich liczb (twierdzenie) Czebyszewa.

Należy sprawdzić założenia tego twierdzenia, tzn. czy podane zmienne mają skończone wariancje i czy ciąg wariancji ma granicę równą zero. Mamy

$$\begin{aligned} E(X_n) &= -na \cdot \frac{1}{2^n} + 0 \cdot \left(1 - \frac{1}{2^{n-1}}\right) + na \cdot \frac{1}{2^n} = 0, \\ E(X_n^2) &= (-na)^2 \cdot \frac{1}{2^n} + 0^2 \cdot \left(1 - \frac{1}{2^{n-1}}\right) + (na)^2 \cdot \frac{1}{2^n} = \frac{n^2 a^2}{2^{n-1}}, \end{aligned}$$

skąd

$$\sigma_n^2 = D^2(X_n) = E(X_n^2) - (E(X_n))^2 = \frac{n^2 a^2}{2^{n-1}}.$$

Oznacza to, że zmienne X_n mają skończone wariancje. Ponadto mamy

$$\lim_{n \rightarrow \infty} \sigma_n^2 = \lim_{n \rightarrow \infty} \frac{n^2 a^2}{2^{n-1}} = 0,$$

więc na mocy wspomnianego twierdzenia ciąg zmiennych losowych $\{X_n\}$ jest stochastycznie zbieżny do zera, tzn.

$$\lim_{n \rightarrow \infty} P(|X_n| \geq \varepsilon) = 0.$$

Definicja 6.6. *Mocnym prawem wielkich liczb* nazywamy każde twierdzenie, które z prawdopodobieństwem równym jedności ocenia zbieżność (w zwykłym sensie) pewnego ciągu zmiennych losowych

$$P\left(\lim_{n \rightarrow \infty} U_n = U\right) = 1.$$

Równość podaną w powyższej definicji można też zapisać następująco:

$$P\left(\lim_{n \rightarrow \infty} U_n \neq U\right) = 0,$$

czyli prawdopodobieństwo, że ciąg nie jest zbieżny do U jest równe zero.

Jednym z podstawowych mocnych praw wielkich liczb jest twierdzenie Borela-Cantelliego, zwane też *mocnym prawem wielkich liczb Bernoulliego*, gdyż odnosi się do ciągu zmiennych losowych w schemacie Bernoulliego. Podajemy je bez dowodu.

Twierdzenie 6.8. *Niech $\{Y_n\}$ oznacza ciąg zmiennych losowych podlegających rozkładowi Bernoulliego*

$$P(Y_n = k) = \binom{n}{k} p^k q^{n-k}, \quad 0 < p < 1, \quad q = 1 - p, \quad k = 0, 1, \dots, n.$$

Wówczas

$$P\left(\lim_{n \rightarrow \infty} \frac{Y_n}{n} = p\right) = 1.$$

Ponieważ wielkość $\frac{Y_n}{n} = \frac{k}{n}$ jest częstością względną sukcesów w serii n doświadczeń, więc twierdzenie to mówi, że z prawdopodobieństwem równym jedności przy liczbie doświadczeń dążących do nieskończoności częstość sukcesów w schemacie Bernoulliego dąży do prawdopodobieństwa uzyskania sukcesu w pojedynczym doświadczeniu.

Wykłady opracowano na podstawie podręczników

T. Gerstenkorn, T. Śródka, *Kombinatoryka i rachunek prawdopodobieństwa*, Państwowe Wydawnictwo Naukowe, Warszawa

oraz

R. Zieliński, *Rachunek prawdopodobieństwa z elementami statystyki matematycznej*, Państwowe Zakłady Wydawnictw Szkolnych, Warszawa.