Wstęp do rachunku prawdopodobieństwa
11. Prawa wielkich liczb. Centralne twierdzenie graniczne

Twierdzenie (Prawo Wielkich Liczb). *Niech $X_1, X_2, \ldots, X_n$ będzie ciągiem niezależnych zmiennych losowych o tym samym rozkładzie, dla których $\mathbb{E}X_1 = \mu$ oraz $VarX_1 = \sigma^2 < \infty$. Wtedy dla zmiennej losowej $S_n = X_1 + \ldots + X_n$ i dowolnej stałej $\varepsilon > 0$ zachodzi*

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right) = 0.$$

Przypomnijmy, że zmienna losowa X ma **rozkład normalny** z parametrami μ i σ^2, co zapisujemy jako $X \sim \mathcal{N}(\mu, \sigma^2)$, jeśli jej gęstość dana jest wzorem

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad \text{dla} \quad x \in (-\infty, \infty).$$

W tym wypadku $\mathbb{E}X = \mu$ oraz $\mathrm{Var}X = \sigma^2$.

Jeżeli $X \sim \mathcal{N}(0, 1)$, to mówimy, że X ma **standardowy rozkład normalny**. Dystrybuantę tego rozkładu będziemy oznaczać przez Φ. Funkcja Φ ma bardzo skomplikowaną postać, dlatego jej wartości będziemy odczytywać z tablic. Najczęściej interesować nas będą poniższe wartości dystrybuanty Φ:

$$\Phi(1{,}28) \approx 0{,}9, \;\; \Phi(1{,}64) \approx 0{,}95, \;\; \Phi(1{,}96) \approx 0{,}975, \;\; \Phi(2{,}33) \approx 0{,}99, \;\; \Phi(2{,}58) \approx 0{,}995.$$

Tabela przedstawiająca więcej wartości dystrybuanty Φ znajduje się na ostatniej stronie. Kalkulator rozkładu normalnego można znaleźć klikając w następujący link: **kalkulator**

Uwaga. Gęstość standardowego rozkładu normalnego jest funkcją parzystą. Oznacza to, że jej wykres jest symetryczny względem osi OY. Co więcej funkcja ta przyjmuje kształt dzwonu, stąd też nazywa się ją czasem krzywą dzwonową lub krzywą Gaussa (od nazwy rozkładu, a rozkład normalny – rozkładem Gaussa). W związku z powyższym dystrybuanta tego rozkładu spełnia zależność:

$$\Phi(-x) = 1 - \Phi(x),$$

dla wszystkich $x \in (-\infty, \infty)$. Dzięki temu przekształceniu, pomimo że mamy podane wartości dystrybuanty tylko dla dodatnich argumentów, możemy też odczytać wartości dla argumentów ujemnych.

Twierdzenie (Centralne Twierdzenie Graniczne). *Niech $X_1, X_2, \ldots, X_n$ będzie ciągiem niezależnych zmiennych losowych o tym samym rozkładzie, dla których $\mathbb{E}X_1 = \mu$, $VarX_1 = \sigma^2 < \infty$ i $\mathbb{E}(|X_1|^3) < \infty$. Wtedy dla zmiennej losowej $S_n = X_1 + \ldots + X_n$ i dowolnej stałej $x \in \mathbb{R}$ zachodzi*

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{S_n - \mu n}{\sigma\sqrt{n}} \leqslant x\right) = \Phi(x).$$

Uwaga. Centralne Twierdzenie Graniczne (w skrócie CTG) odnosi sie do sytuacji gdy n i $\mathbb{E}S_n$ są dostatecznie duże. Zwykle CTG dobrze przybliża rozkład S_n gdy $n \geqslant 30$ i $\mathbb{E}S_n \geqslant 10$. W tej sytuacji nie będziemy przejmować się błędem przybliżenia.

Jeśli stosujemy CTG do przybliżania zmiennej losowej S_n, która przyjmuje tylko wartości całkowite, należy pamiętać o uwzględnieniu „poprawki” wynoszącej $1/2$. Na przykład, chcąc użyć CTG do oszacowania prawdopodobieństwa $\mathbb{P}(S_n = k)$, należy je zastąpić przez prawdopodobieństwo

$$\mathbb{P}\left(k - 1/2 < S_n \leqslant k + 1/2\right).$$

Dodatek A. Zadania na ćwiczenia

Zadanie 1. Średnica sekwoi wzrasta od 1 do 2 cm rocznie, przy czym zakładamy, że przyrost ma rozkład jednostajny na odcinku $[1, 2]$, a przyrosty w kolejnych latach nie zależą od siebie. Oszacuj prawdopodobieństwo, że w ciągu 300 lat średnica zwiększy się o co najmniej 460 cm, stosując:

(a) nierówność Markowa,
(b) nierówność Czebyszewa,
(c) Centralne Twierdzenie Graniczne.

Zadanie 2. Rzucamy symetryczną kostką 100 razy. Oszacuj prawdopodobieństwo, że wyrzucimy szóstkę mniej niż 10 razy, stosując:

(a) twierdzenie Czebyszewa,
(b) CTG bez poprawki 1/2,
(c) CTG z poprawką 1/2.

Czy możemy oszacować to prawdopodobieństwo, stosując twierdzenie Markowa?

Stosując CTG, oszacuj prawdopodobieństwo wyrzucenia dokładnie 10 szóstek i porównaj je z rzeczywistą wartością tego prawdopodobieństwa.

Zadanie 3. Kierowniczka urzędu pocztowego twierdzi, że w jej placówce klient czeka w kolejce do okienka krócej niż 15 minut. Dociekliwa ekipa telewizyjna chce zweryfikować jej deklarację, mierząc średni czas obsługi 50 klientów. Jakie jest prawdopodobieństwo, że potwierdzi ona deklarację kierowniczki urzędu jeśli wiadomo, że rzeczywisty czas T stania w kolejce jest taką ciągłą zmienną losową, że $\mathbb{E}T = 14$ i $\mathrm{Var}T = 23$? Czy należy stosować w tym przypadku„poprawkę 1/2”?

Przypuśćmy, że ekipa telewizyjna zapytała uprzednio kierowniczkę, czas ilu kolejnych klientów ma zmierzyć. Co powinna odpowiedzieć, by mieć 99% pewności, że wyliczenia ekipy nie obalą jej twierdzenia?

Dodatek B. Zadania domowe

Zadanie 1. Oznaczmy przez S_{100} liczbę szóstek wyrzuconych w 100 rzutach symetryczną kostką. Użyj kalkulatora i tablic statystycznych do:

(a) wyznaczenia wartości $\mathbb{P}(S_{100} = 15)$;
(b) oszacowania $\mathbb{P}(S_{100} = 15)$ przy użyciu CTG.

Zadanie 2. Niech S_{400} oznacza liczbę orłów wyrzuconych w 400 rzutach monetą.

(a) Oszacuj, używając kalkulatora i tablic statystycznych, $\mathbb{P}(205 < S_{400} \leqslant 222)$.
(b) Znajdź przy pomocy CTG najmniejszą wartość k, dla której $\mathbb{P}(S_{400} \leqslant k) > 0{,}95$.
(c) Oszacuj prawdopodobieństwo z podpunktu (b), używając nierówności Czebyszewa.

Zadanie 3. Ile co najmniej razy musimy rzucić symetryczną kostką, aby prawdopodobieństwo tego, że szóstka wypadnie więcej niż w 20% rzutach było mniejsze niż 0,1?

(a) Oszacuj z góry potrzebną liczbę rzutów, korzystając z nierówności Czebyszewa.
(b) Rozwiąż to zadanie przy użyciu CTG.

Zadanie 4. Mierzony w minutach czas T, jaki spędza codziennie pan January czekając na autobus, ma rozkład wykładniczy z parametrem $\lambda = 0{,}1$. Używając CTG, oszacuj prawdopodobieństwo, że w ciągu ostatnich 100 dni pan January czekał na autobus w sumie dłużej niż 1200 minut (tzn. średnio dłużej niż 12 minut).

Dodatek C. Odpowiedzi do zadań domowych

B.1 (a) $\approx 0{,}10024$, (b) 0,09732

B.2 (a) bez poprawki 1/2: $0{,}29464$, z poprawką: 0,27894 (b) 217, (c) 245

B.3 (a) 1251, (b) 209

B.4 0,02275

Tablica 1. Dystrybuanta $\Phi(x)$ standardowego rozkładu normalnego $\mathcal{N}(0,1)$

x	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0	0,5	0,50399	0,50798	0,51197	0,51595	0,51994	0,52392	0,5279	0,53188	0,53586
0,1	0,53983	0,5438	0,54776	0,55172	0,55567	0,55962	0,56356	0,56749	0,57142	0,57535
0,2	0,57926	0,58317	0,58706	0,59095	0,59483	0,59871	0,60257	0,60642	0,61026	0,61409
0,3	0,61791	0,62172	0,62552	0,6293	0,63307	0,63683	0,64058	0,64431	0,64803	0,65173
0,4	0,65542	0,6591	0,66276	0,6664	0,67003	0,67364	0,67724	0,68082	0,68439	0,68793
0,5	0,69146	0,69497	0,69847	0,70194	0,7054	0,70884	0,71226	0,71566	0,71904	0,7224
0,6	0,72575	0,72907	0,73237	0,73565	0,73891	0,74215	0,74537	0,74857	0,75175	0,7549
0,7	0,75804	0,76115	0,76424	0,7673	0,77035	0,77337	0,77637	0,77935	0,7823	0,78524
0,8	0,78814	0,79103	0,79389	0,79673	0,79955	0,80234	0,80511	0,80785	0,81057	0,81327
0,9	0,81594	0,81859	0,82121	0,82381	0,82639	0,82894	0,83147	0,83398	0,83646	0,83891
1	0,84134	0,84375	0,84614	0,84849	0,85083	0,85314	0,85543	0,85769	0,85993	0,86214
1,1	0,86433	0,8665	0,86864	0,87076	0,87286	0,87493	0,87698	0,879	0,881	0,88298
1,2	0,88493	0,88686	0,88877	0,89065	0,89251	0,89435	0,89617	0,89796	0,89973	0,90147
1,3	0,9032	0,9049	0,90658	0,90824	0,90988	0,91149	0,91309	0,91466	0,91621	0,91774
1,4	0,91924	0,92073	0,9222	0,92364	0,92507	0,92647	0,92785	0,92922	0,93056	0,93189
1,5	0,93319	0,93448	0,93574	0,93699	0,93822	0,93943	0,94062	0,94179	0,94295	0,94408
1,6	0,9452	0,9463	0,94738	0,94845	0,9495	0,95053	0,95154	0,95254	0,95352	0,95449
1,7	0,95543	0,95637	0,95728	0,95818	0,95907	0,95994	0,9608	0,96164	0,96246	0,96327
1,8	0,96407	0,96485	0,96562	0,96638	0,96712	0,96784	0,96856	0,96926	0,96995	0,97062
1,9	0,97128	0,97193	0,97257	0,9732	0,97381	0,97441	0,975	0,97558	0,97615	0,9767
2	0,97725	0,97778	0,97831	0,97882	0,97932	0,97982	0,9803	0,98077	0,98124	0,98169
2,1	0,98214	0,98257	0,983	0,98341	0,98382	0,98422	0,98461	0,985	0,98537	0,98574
2,2	0,9861	0,98645	0,98679	0,98713	0,98745	0,98778	0,98809	0,9884	0,9887	0,98899
2,3	0,98928	0,98956	0,98983	0,9901	0,99036	0,99061	0,99086	0,99111	0,99134	0,99158
2,4	0,9918	0,99202	0,99224	0,99245	0,99266	0,99286	0,99305	0,99324	0,99343	0,99361
2,5	0,99379	0,99396	0,99413	0,9943	0,99446	0,99461	0,99477	0,99492	0,99506	0,9952
2,6	0,99534	0,99547	0,9956	0,99573	0,99585	0,99598	0,99609	0,99621	0,99632	0,99643
2,7	0,99653	0,99664	0,99674	0,99683	0,99693	0,99702	0,99711	0,9972	0,99728	0,99736
2,8	0,99744	0,99752	0,9976	0,99767	0,99774	0,99781	0,99788	0,99795	0,99801	0,99807
2,9	0,99813	0,99819	0,99825	0,99831	0,99836	0,99841	0,99846	0,99851	0,99856	0,99861
3	0,99865	0,99869	0,99874	0,99878	0,99882	0,99886	0,99889	0,99893	0,99896	0,999
3,1	0,99903	0,99906	0,9991	0,99913	0,99916	0,99918	0,99921	0,99924	0,99926	0,99929
3,2	0,99931	0,99934	0,99936	0,99938	0,9994	0,99942	0,99944	0,99946	0,99948	0,9995
3,3	0,99952	0,99953	0,99955	0,99957	0,99958	0,9996	0,99961	0,99962	0,99964	0,99965
3,4	0,99966	0,99968	0,99969	0,9997	0,99971	0,99972	0,99973	0,99974	0,99975	0,99976
3,5	0,99977	0,99978	0,99978	0,99979	0,9998	0,99981	0,99981	0,99982	0,99983	0,99983
3,6	0,99984	0,99985	0,99985	0,99986	0,99986	0,99987	0,99987	0,99988	0,99988	0,99989
3,7	0,99989	0,9999	0,9999	0,9999	0,99991	0,99991	0,99992	0,99992	0,99992	0,99992
3,8	0,99993	0,99993	0,99993	0,99994	0,99994	0,99994	0,99994	0,99995	0,99995	0,99995
3,9	0,99995	0,99995	0,99996	0,99996	0,99996	0,99996	0,99996	0,99996	0,99997	0,99997
4	0,99997	0,99997	0,99997	0,99997	0,99997	0,99997	0,99998	0,99998	0,99998	0,99998