Rachunek Prawdopodobieństwa MAP1181

Wydział Matematyki, Matematyka Stosowana

Lista 8. Rozwiązanie zadania 8.3 - dodatkowy podpunkt (f)

Opracowanie: Piotr Janus

Zadanie **8.3**

(f) Na ulicy stoi sprzedawca hot dogów. Załóżmy, że każdy z mijających go przechodniów kupuje jednego hot doga z prawdopodobieństwem 0.05. Niech X oznacza ilość ludzi mijających sprzedawcę aż do chwili, gdy sprzeda on 50 porcji hot dogów. Znajdź dokładny i asymptotyczny rozkład zmiennej losowej X. Oblicz prawdopodobieństwo, że zanim sprzedawca sprzeda 50 hot dogów, minie go nie więcej niż 1100 osób. Porównaj wyniki otrzymane na podstawie rozkładów dokładnego i asymptotycznego.

Rozwiązanie:

- Model: schemat Bernoulliego. Jako sukces przyjmijmy sytuację, gdy mijający przechodzień kupi hot-doga, $p = 0.05$.
- X, czyli liczba osób mijających stoisko do momentu sprzedaży 50-tego hot doga, to czas oczekiwania na 50-ty sukces.

Rozkład dokładny zmiennej losowej X:

- Wiadomo, że czas oczekiwania na 50-ty sukces ma rozkład Pascala $\mathcal{NB}(m = 50, p = 0.05)$, tzn.
$$P(X = k) = \binom{k-1}{50-1} 0.05^{50} 0.95^{k-50}, \text{ gdzie } k = 50, 51, 52 \dots$$
- Szukane prawdopodobieństwo, że zanim sprzedawca sprzeda 50 hot dogów, minie go nie więcej niż 1100 osób, wynosi
$$P(X \leqslant 1100) = \sum_{k=50}^{1100} \binom{k-1}{50-1} \cdot 0.05^{50} \cdot (1-0.05)^{k-50} \approx 0.7659.$$

Rozkład asymptotyczny zmiennej losowej X:

- Zauważmy, że $X = \sum_{i=1}^{50} X_i$, gdzie X_1 to czas oczekiwania na sprzedaż pierwszego hot doga, natomiast X_i, gdzie $i = 2, 3, \dots, 50$, to czas oczekiwania (gdzie oczekiwanie rozpoczynamy tuż po sprzedaży $(i-1)$-szego hot doga) na pierwszy sukces w schemacie Bernoulliego (czyli na sprzedaż i-tego hot doga).
- Ponieważ modelem jest schemat Bernoulliego, zmienne losowe $X_1, X_2, \dots, X_{50}$ są niezależne i mają taki sam rozkład geometryczny $\mathcal{G}eo(p = 0.05)$.
- $X_i \sim \mathcal{G}eo(p = 0.05)$, więc $\mathrm{E}X_i = \frac{1}{p} = 20 = m$, a $\mathrm{D}^2 X_i = \frac{1-p}{p^2} = 380 = \sigma^2$.
- Ponieważ wariancja $\sigma^2 = 380$ jest skończona i większa od 0, a $n = 50$ jest wystarczająco duże, możemy skorzystać z CTG Lindeberga–Lévy'ego. Otrzymujemy, że zmienna losowa $\dfrac{X - nm}{\sigma\sqrt{n}}$ ma asymptotycznie standardowy rozkład normalny $\mathcal{N}(0, 1)$, a stąd X ma asymptotycznie rozkład normalny $\mathcal{N}(nm, \sigma\sqrt{n}) = \mathcal{N}(1000, \sqrt{19000})$.

- Oznacza to, że

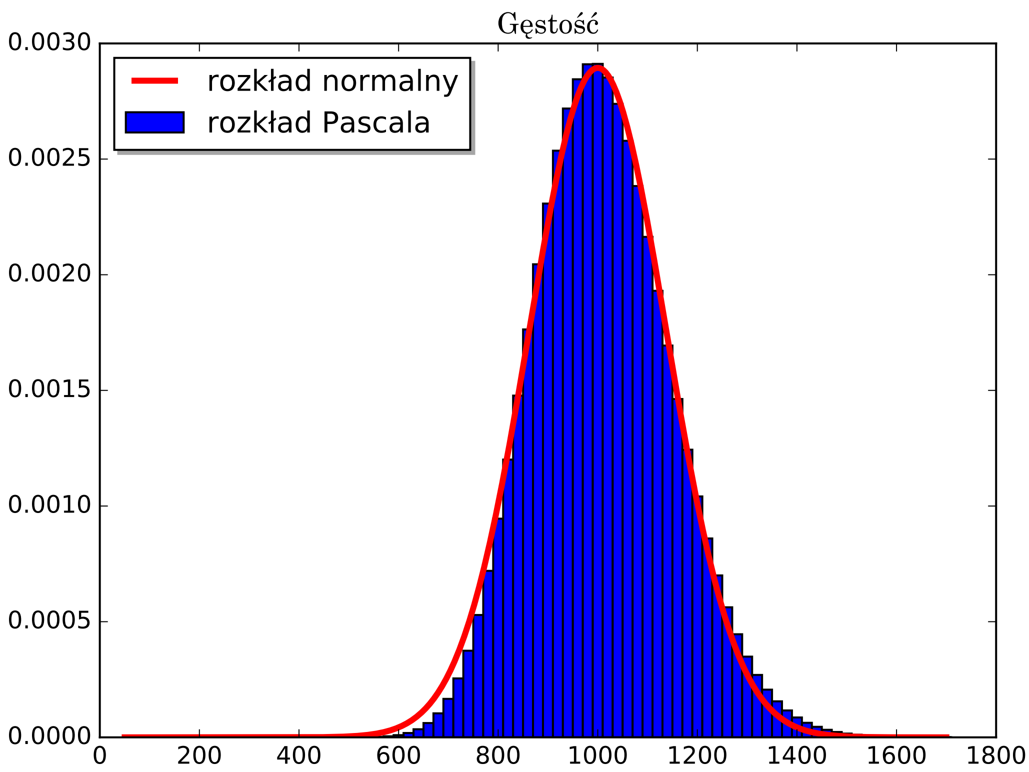
$$P(X \leqslant k) \approx \Phi\left(\frac{k-1000}{\sqrt{19000}}\right).$$

- Podstawiając $k = 1100$ otrzymujemy

$$P(X \leqslant 1100) \approx \Phi\left(\frac{1100-1000}{\sqrt{19000}}\right) \approx \Phi(0.73) \overset{z\ tab.stat.}{=} 0.7673.$$

Otrzymane oszacowanie szukanego prawdopodobieństwa (czyli 0.7673) jest istotnie bliskie wynikowi otrzymanego ze wzorów dokładnych (0.7659).

Porównanie rozkładów dokładnego i asymptotycznego



Rysunek 1: Porównanie dyskretnego rozkładu dokładnego (Pascala) zmiennej losowej X z gęstością normalnego rozkładu asymptotycznego.