Rozwiązania zadań 15 z listy nr. 5 i 5 z listy nr.6 z przedmiotu "Rachunek prawdopodobieństwa dla informatyków"

Mateusz Małowiecki

24 stycznia 2022

Zadanie 15

Treść

Przed głosowaniem nad impeachmentem prezydenta pewnego państwa, przeprowadzono sondaż. W ramach sondażu spytano n osób o to czy popierają usunięcie prezydenta. Zakładamy, że każda z pytanych osób została losowo i niezależnie od pozostałych wybrana ze zbioru osób uprawnionych do głosowania (w szczególności nie wykluczamy powtórzeń) oraz odpowiedziała TAK /NIE. Interesuje nas oszacowanie poparcia wniosku. Oszacuj jak duże powinno być n, żeby prawdopodobieństwo popełnienia błędu względnego większego niż 1% było mniejsze niż 5%.

Rozwiązanie

Niech:

- m - liczba mieszkańców państwa
- k - liczba osób, które zamierzają odpowiedzieć "TAK"
- X_n - zmienna losowa, mówiąca ilu mieszkańców odpowiedziało "TAK" spośród n wybranych

Oczywiście:

$$X_n \sim B\Big(n, \frac{k}{m}\Big) \tag{1}$$

więc:

$$E(X_n) = \frac{n * k}{m} \tag{2}$$

Zauważmy teraz że błąd względny dany jest wzorem:

$$\delta_x = \frac{|x - x_0|}{|x_0|} \tag{3}$$

gdzie

- x - wartość zmierzona

- x_0 - dokładna wartość

W naszym przypadku $x = \frac{X_n}{n}$ i $x_0 = \frac{k}{m}$. Zatem

$$\delta_x = \frac{|\frac{X_n}{n} - \frac{k}{m}|}{|\frac{k}{m}|} = \frac{|m|}{|n * k|} * \left|X_n - \frac{k * n}{m}\right| \tag{4}$$

Korzystając z (2), możemy zauważyć, że

$$\delta_x = \frac{1}{E(X_n)} * |X_n - E(X_n)| \tag{5}$$

Zatem

$$P(\delta_x \geqslant 0.01) = P(\frac{1}{E(X_n)} * |X_n - E(X_n)| \geqslant 0.01) = P(|X_n - E(X_n)| \geqslant 0.01 * E(X_n)) \tag{6}$$

Skorzystamy teraz z nierówności Chernoffa:

$$P(|Y - E(Y)| \geqslant \delta * E(Y)) \leqslant 2 * e^{-\frac{\delta^2 * E(Y)}{3}} \tag{7}$$

A zatem:

$$P(|X_n - E(X_n)| \geqslant 0.01 * E(X_n)) \leqslant 2 * e^{-\frac{0.01^2 * E(X_n)}{3}} \tag{8}$$

Zatem wystarczy znaleźć n dla którego

$$2 * e^{-\frac{0.01^2 * E(X_n)}{3}} < 0.05 \tag{9}$$

Po prostych przekształceniach otrzymamy:

$$n > 30000 * ln(4) * \frac{m}{k} \approx 41588 * \frac{m}{k} \tag{10}$$

Zadanie 5

Treść

Wynik pomiaru opóźnienia w transmisji na pojedynczym odcinku sieci komputerowej obarczony jest błędem systematycznych 0.5 i błędem losowym będącym zmienną losową o średniej 0 i wariancji 1. Dokonano pomiaru sieci składającej się ze 100 odcinków (pomiary były niezależne od siebie). Niech Y będzie błędem całkowitym popełnionych przy badaniu całej sieci. Oszacować prawdopodobieństwo tego, że:
a) $Y < 75$
b) Wynik pomiaru nie przekracza rzeczywistej wartości mierzonego opóźnienia na trasie 100 odcinków.

Rozwiązanie

Niech Y_i - opóźnienie na i-tym odcinku. Oczywiście:

$$Y = \sum_{i=1}^{n} Y_i \tag{11}$$

Zauważmy, że

$$Y_i = Z_i + 0.5 \tag{12}$$

dla pewnych zmiennych losowych Z_i takich, że $E(Z_i) = 0$ i $V(Z_i) = 1$. Zatem $E(Y_i) = 0.5$ i $V(Y_i) = 1$. W celu obliczenia prawdopodobieństw skorzystamy z centralnego twierdzenia granicznego. A zatem:

$$U = \frac{Y - 100 * \frac{1}{2}}{\sqrt{100 * 1}} \sim N(0,1) \tag{13}$$

podpunkt a

Chcemy oszacować $P(Y < 75)$, a zatem:

$$P(Y < 75) = P\Big(\frac{Y - 100 * \frac{1}{2}}{\sqrt{100 * 1}} < \frac{75 - 100 * \frac{1}{2}}{\sqrt{100 * 1}}\Big) = P(U < 2,5) \approx \Phi(2,5) = 0,993 \tag{14}$$

podpunkt b

Ponieważ rzeczywista wartość opóźnienia na trasie 100 odcinków wynosi 50, więc:

$$P(Y < 50) = P\Big(\frac{Y - 100 * \frac{1}{2}}{\sqrt{50 * 1}} < \frac{75 - 100 * \frac{1}{2}}{\sqrt{100 * 1}}\Big) = P(U < 0) \approx \Phi(0) = 0,5 \tag{15}$$