

Praca domowa nr 1 (zwrot: piątek 8 marca 2024, godz. 23.00) (3 pkt)

Korzystając z danych rzeczywistych nt. liczby ofiar w wojnach i konfliktach zbrojnych, które miały miejsce w latach 1816 – 2007 (plik wars.txt) przygotuj:

1. Histogram $N(x)$ opisujący liczbę konfliktów o liczbie ofiar w przedziałach o ustalonej szerokości Δx , tj. $(x_0, x_0+\Delta x)$, $(x_0+\Delta x, x_0+2\Delta x)$, $(x_0+2\Delta x, x_0+3\Delta x)$ itd., gdzie x_0 jest najmniejszą liczbą ofiar wśród analizowanych konfliktów. Szerokość przedziału Δx ustal samodzielnie. Możesz poprobować, jak wartość Δx wpływa na histogram. Wykres $N(x)$ przedstaw w równych skalach. Wypróbuj skale: liniowo-liniową (lin-lin), liniowo-logarytmiczną (lin-log) i podwójnie logarytmiczną (log-log).
2. Na podstawie histogramu przygotuj wykres prawdopodobieństwa $P(x)$. Pamiętaj, że $P(x)=N(x)/(N\Delta x)$, gdzie N -liczba wszystkich konfliktów. Zastanów się, dlaczego we wzorze na $P(x)$ występuje dzielenie przez szerokość przedziału. Zauważ, że ponieważ we wzorze na $P(x)$ dzielimy $N(x)$ przez stałą wartość, dlatego wykresy $P(x)$ i $N(x)$ są identyczne. Jedyną rzeczą która je odróżnia jest „przeskalowana” pionowa oś.
3. Przygotuj wykres $P(x)$ korzystając z metody „binowania logarytmicznego”, w której liczbę ofiar $N(x)$ zlicza się w przedziałach o wykładniczo rosnącej szerokości, tj. $(x_0, x_0 a)$, $(x_0 a, x_0 a^2)$, $(x_0 a^2, x_0 a^3)$ itd. Gdy z $N(x)$ będziesz obliczać $P(x)$ nie zapomnij o podzieleniu tej liczby przez długość przedziału, która w tej sytuacji nie jest stała i wynosi odpowiednio: dla pierwszego przedziału: $x_0 a - x_0 = x_0(a-1)$, dla drugiego: $x_0 a^2 - x_0 a$, dla trzeciego: $x_0 a^3 - x_0 a^2$. Przedstaw ten wykres w różnych skalach: liniowo-liniowej, liniowo-logarytmicznej i podwójnie logarytmicznej.
4. Przygotuj wykres przedstawiający tzw. skumulowany rozkład $P^c(x)$ – we wszystkich ww. skalach. Rozkład skumulowany – patrz prezentacja do wykładu.