

SEMINARIUM UKŁADY DYNAMICZNE

Tytuł: **Prawo Benforda**
Referent: **Magdalena Nowak**
Data: **4 I 2013**

Dane statystyczne mają rozkład Benforda jeśli prawdopodobieństwo wystąpienia cyfry n na pierwszej pozycji w liczbie przedstawia się następująco:

$$P_n = \log_{10} \left(1 + \frac{1}{n} \right)$$

Prawem Benforda nazywamy zjawisko częstego występowania powyższego rozkładu w danych statystycznych. Zbiory danych spełniają to prawo jeśli posiadają następujące własności:

- Dane muszą być zgromadzone w tej samej przestrzeni fizycznej, np. dotyczące jedynie odległości w km.
- Zbiór danych nie może posiadać górnej ani dolnej granicy przyjmowanych wartości.
- Zbiór danych nie może posiadać żadnych przypisanych numerów, np. numerów pracowników, kont bankowych, telefonów.
- Zestaw danych powinien zawierać liczby różnych rzędów. Efektem naturalnego procesu rozwoju jest występowanie większej ilości liczb o niższej wartości, a mniejszej ilości liczb o wyższej, np. faktury. Niemniej jednak zestaw danych nie musi być bardzo rozproszony.

Przykłady dancy spełniających prawo Benforda:

- (1) Ciągi geometryczne (np. 2^n), ciąg Fibonacciego.
- (2) Dane geograficzne - powierzchnie krajów lub jezior, liczba ludności, długości granic lub rzek, wysokości szczytów.
- (3) Dane fizyczne - odległości galaktyczne, masy atomowe pierwiastków, stałe fizyczne.
- (4) Stałe matematyczne.
- (5) Dane ekonomiczne - wysokości PKB, faktury, podatki.

Prawo Benforda zostało zauważone już w roku 1881 przez Simona Newcomba lecz nazwę swą wzięło dopiero od Franka Benforda, który w 1938 roku zauważył dokładnie to samo co Newcomb: tablice logarytmiczne w bibliotece były najbardziej brudne na początkowych stronach a czystsze na kolejnych, co oznaczało, że znacznie częściej korzysta się z tych tablic dla liczb rozpoczynających się od 1. Dowód prawa Benforda został podany dopiero w 1995 roku przez Theodora P. Hilla. Podczas seminarium zostały omówione fragmenty tego rozumowania.

Zestawy danych, które spełniają prawo Benforda, tzw. zbiory Benforda, posiadają następujące, ciekawe własności:

- (1) Niezmienniczość skali – mnożąc wszystkie wartości w zbiorze przez stałą otrzymujemy również zbiór spełniający prawo Benforda. Oznacza to, że przeliczanie danych na inne miary lub waluty nie zmienia rozkładu cyfr pierwszych w tych danych.
- (2) Niezmienniczość bazy – prawo Benforda dotyczy nie tylko zbiorów opartych na dziesiętnym systemie liczbowym, ale również wszystkich innych systemach oprócz unarnego.
- (3) Niezmienniczość w operacjach matematycznych – potęgowanie, mnożenie, dzielenie, dodawanie oraz odejmowanie zbiorów Benforda ze sobą powoduje powstanie nowego zbioru Benforda. Naturalne modele matematyczne ewolucji danych nie powinny zatem "psuć" rozkładu Benforda.

Prawo Benforda stosowane jest między innymi do sprawdzania poprawności obliczeń, prawdziwości danych statystycznych, do wykrywania oszustw w zeznaniach podatkowych i rozliczeniach finansowych oraz do sprawdzania dokładności działania modeli matematycznych opisujących ewolucję danych z różnych dziedzin. Dzięki niemu możemy bez obaw symulować naturalne zjawiska fizyczne, biologiczne i ekonomiczne stosując np. ciągi geometryczne czy ciąg Fibonacciego.