

ANALIZA JEDNOZMIENNOWA

podstawowe pojęcia



DZISIAJ

- Krótka retrospekcja – przypomnienie typów zmiennych
- O obliczaniu znanych już wartości:
 - średniej
 - mediany
 - mody
- Ćwiczenia praktyczne, czyli liczenie zadań

PRZYPOMNIJMY

● ZMIENNA CIĄGŁA

„Zmienna X jest ciągła, jeśli może ona przyjmować każdą dowolną wartość z określonego przedziału liczb rzeczywistych”



● ZMIENNA DYSKRETNA

Zmienna jest dyskretna „jeżeli może ona przyjmować skończony lub przeliczalny zbiór wartości”
(Sambor 1990)



ĆWICZENIE

ZMIENNA
CIĄGŁA

ZMIENNA
DYSKRETNA

- płeć
- ilość głosek w wyrazie
- długość dźwięku
- wysokość pierwszego formantu
- ilość zapożyczeń w tekście
- zarobki

ĆWICZENIE

Proszę określić na jakim poziomie mierzone są poniższe zmienne:

- płeć
- ilość głosek w wyrazie
- długość dźwięku
- wysokość pierwszego formantu
- ilość zapożyczeń w tekście
- zarobki
- nacechowanie emocjonalne wyrazu
- wykształcenie
- częstość występowania wyrazu
- kategoria gramatyczna wyrazu
- region pochodzenia
- przynależność tekstu odmiany stylistycznej
- rodzaj użytego chwytu perswazyjnego

ROZWIĄZANIE

- płeć
- ilość głosek w wyrazie
- długość dźwięku
- wysokość pierwszego formantu
- kategoria gramatyczna wyrazu
- region pochodzenia
- przynależność tekstu odmiany stylistycznej
- rodzaj użytego chwytu perswazyjnego
- ilość zapożyczeń w tekście
- zarobki
- nacechowanie emocjonalne wyrazu
- wykształcenie
- częstość występowania wyrazu

NOMINALNA

PORZĄDKOWA

ŚREDNIA

Dobrze znany wzór:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

- poziom: minimum interwałowy
- w przypadku zmiennych pogrupowanych w przedziały klasowe: bierzemy pod uwagę środek każdego przedziału

MEDIANA

- poziom: minimum porządkowy
- Dla zbioru danych (tabeli liczebnościowej) LUB danych dyskretnych:
 1. Szeregujemy obserwacje od najmniejszej do największej (lub odwrotnie)
 2. Numerujemy uszeregowane obserwacje od 1 do N
 3. Wybieramy medianę:

Przy *nieparzystej* ilości obserwacji:

wartość znajdująca się na miejscu $(N+1)/2$

Przy *parzystej* ilości obserwacji:

średnia z wartości znajdujących się na miejscach $N/2$ i $N/2+1$

MEDIANA

Dla danych ciągłych znajdujących się w przedziałach klasowych

1. Obliczamy $N/2$ (w przypadku nieparzystej ilości obserwacji zaokrąglamy w górę)
2. Szukamy przedziału klasowego, w którym znajduje się wartość z tym numerem.
3. Znajdujemy dolną granicę przedziału, w którym znajduje się mediana L
4. Określamy liczebność skumulowaną poniżej przedziału, w którym znajduje się mediana F
5. Określamy liczebność przedziału, w którym znajduje się mediana f
6. Określamy długość przedziału klasowego, w którym znajduje się mediana h
7. Podstawiamy do wzoru:

$$Mn = L + \frac{N/2 - F}{f} * h$$

MODA

- Może być obliczana na każdym poziomie zmiennej
- *Dla zbioru danych (tabeli liczebnościowej) LUB danych dyskretnych: wybieramy tę wartość (te wartości), których częstości są najwyższe.*
- *Dla danych ciągłych znajdujących się w przedziałach klasowych: obliczamy modę wg wzoru:*

$$Mo = X_s + \frac{n_s - n_{s-1}}{(n_s - n_{s-1}) + (n_s - n_{s+1})}$$

X_s – dolna granica przedziału klasowego zawierającego wartość modalną

$n_s ; n_{s-1} ; n_{s+1}$ – liczebność klasy zawierającej modalną i klas sąsiadujących

PROJEKT – cz. 2 obliczenia

do końca tygodnia: na stronie ZSL bada danych wraz z opisem zmiennych (poziomy na jakich znajdują się poszczególne zmienne)

CO NALEŻY ZROBIĆ:

- dobranie wartości dla zmiennych do poszczególnych wartości liczbowych
- prezentacja najważniejszych (z perspektywy problemu badawczego) wyników w postaci wykresów i tabel
- obliczenie odpowiednich miar statystycznych (przynajmniej po dwie każdego typu)
- omówienie i interpretacja tych miar
- obliczenie korelacji dla przynajmniej 3 stawianych hipotez
- interpretacja otrzymanych wyników

PROJEKT – cz. 2 obliczenia

TERMINY:

- Gotowe projekty: mailem do 18.01.2008
- Indywidualne omówienie wraz z wpisem: trzy terminy do wyboru w pierwszym tygodniu sesji (21.01., 23.01., 25.01.)
- wybór terminu: na ostatnich zajęciach