Zadania ze statystyki dla I roku socjologii

1. Dane poniżej prezentują pomiary dochodu. Uporządkuj poniższe dane w szeregu strukturalnym prostym, od wartości najmniejszych do największych. Następnie utwórz szereg przedziałowy z powyższych danych, przyjmując 5 klas o równych szerokościach przedziałów. Oblicz procentowy udział w poszczególnych klasach i szereg skumulowany.

|  |  |  |
| --- | --- | --- |
| 1234 | 2890 | 974 |
| 2345 | 356 | 2085 |
| 5467 | 3790 | 5207 |
| 3679 | 3890 | 3419 |
| 4800 | 19001 | 4540 |
| 1910 | 7904 | 1650 |
| 987 | 5100 | 727 |
| 3020 | 4280 | 2760 |
| 4100 | 5790 | 3840 |
| 2101 | 8010 | 1841 |

1. Oblicz średnią wartość dochodu w badanej zbiorowości z szeregu prostego oraz szeregu przedziałowego.
2. Tabele poniżej prezentują rozkłady częstości wielkości gospodarstw domowych (xi1 i xi2) społeczności dwóch miasteczek oraz ich odpowiednie częstości (n1 i n2). Należy obliczyć średnią wielkość gospodarstwa domowego oraz medianę wielkości gospodarstwa domowego w obu tych miasteczkach.

Należy również porównać ze sobą te wartości, wskazując, która zbiorowość charakteryzuje się przeciętnie większym gospodarstwem domowym.

|  |  |  |  |
| --- | --- | --- | --- |
| **Xi1** | **ni1** | **Xi2** | **ni2** |
|  1 | 14 | 1 | 12 |
|  2 | 22 | 2 | 30 |
| 3 | 26 | 3 | 25 |
| 4 | 30 | 4 | 28 |
| 5 | 21 | 5 | 21 |
| 6 | 12 | 6 | 12 |
| 7 | 9 | 7 | 9 |
| 8 | 7 | 8 | 7 |
| **∑** | **141** |   | 144 |

1. Uporządkowano dochód mieszkańców jednego z tych miasteczek w szeregu przedziałowym. Należy obliczyć przeciętny dochód oraz medianę dochodów mieszkańców tego miasteczka.

|  |  |
| --- | --- |
| dochód |  liczebność |
| 0 - 999 | 20 |
| 1000 - 1999 | 21 |
| 2000 - 2999 | 15 |
| 3000 - 3999 | 11 |
| 4000 - 4999 | 7 |
| **∑** | **74** |

1. Dochód mieszkańców drugiego z miasteczek został uporządkowany nieco inaczej, co prezentuje tabela poniżej. Proszę obliczyć średni dochód mieszkańców tego miasteczka, wiedząc, że osoba o najwyższym dochodzie zarabia 8200.

|  |  |
| --- | --- |
| 0-2000 | 40 |
| 2001 - 4000 | 31 |
| 4001 - 6000 | 16 |
| powyżej 6000 | 4 |

6. Ustal dla danych zawartych w tabeli poniżej, prezentujących rozkład liczebności (ni) różnej wielkości gospodarstw domowych w dwóch populacjach, następujące miary:

- odchylenie ćwiartkowe

- odchylenie średnie

-odchylenie standardowe

Wiedząc, że w innej miejscowości odchylenie przeciętne tej cechy wynosi 1,8, a odchylenie ćwiartkowe 0,5, określ, która z nich składa się z bardziej zróżnicowanych jednostek.

|  |  |
| --- | --- |
| Xi A | ni A |
| 1 | 82 |
| 2 | 98 |
| 3 | 95 |
| 4 | 84 |
| 5 | 65 |
| 6 | 38 |
| 7 | 28 |
| 8 | 12 |
| 9 | 6 |
| ∑ | 508 |

1. W szeregu przedziałowym zawarte są dane o rozkładach liczebności dzieci ze względu na wiek. Należy ustalić wartość mediany oraz pierwszego i trzeciego kwartyla (Q), odchylenie ćwiartkowe oraz odchylenie standardowe.

|  |  |
| --- | --- |
| Xi | ni |
| 1 - 3 | 35 |
| 4 - 6 | 46 |
| 7 - 9 | 31 |
| 10 - 12 | 21 |
| 13 - 15 | 11 |
|   | 144 |

1. Podany poniżej szereg należy przekształcić w szereg przedziałowy o 3 klasach, równej szerokości. Następnie dla szeregu przedziałowego należy obliczyć medianę oraz odchylenie standardowe, odchylenie ćwiartkowe i współczynnik asymetrii (skośności).

|  |  |
| --- | --- |
| Xi A | ni A |
| 1 | 80 |
| 2 | 91 |
| 3 | 70 |
| 4 | 84 |
| 5 | 65 |
| 6 | 38 |
| 7 | 28 |
| 8 | 15 |
| 9 | 12 |
| ∑ | 483 |

1. Wiemy, że mediana rozkładu pewnej cechy, która ma symetryczny oraz równomierny rozkład, wynosi 5. Ustal, ile wynosić będzie odchylenie ćwiartkowe w tej zbiorowości. Ustal, ile wynosić będzie odchylenie średnie w tej zbiorowości, wiedząc, że natężenie cechy zmienia się co 1 (o jeden interwał) i że rozstęp wynosi 8.

PODSUMOWANIE WYKŁADU

Cechy statystyczne dzielimy na mierzalne i niemierzalne:

* + **Mierzalne** – dające się wyrazić liczbowo (np. wzrost, wiek, dochód)
	+ **Niemierzalne (jakościowe)** – nie dające się wyrazić liczbowo (np. pochodzenie narodowe, miejsce zamieszkania, płeć)

Poziomy pomiaru:

1. Poziom nominalny
2. Poziom porządkowy
3. Poziom interwałowy
4. Poziom ilorazowy
5. Poziom absolutny

Poziom nominalny:

* Najniższy poziom pomiaru
* Charakteryzuje się przypisaniem jednostki analizy do określonej kategorii zmiennej (cechy)
* Kategorie opisują właściwości danej zmiennej
* Podział na kategorie musi być rozłączny i wyczerpujący
* Przykład cechy: płeć (kobiety i mężczyźni), narodowości (np. Polacy, Niemcy, Rosjanie)

Poziom porządkowy:

* Wyższy od nominalnego poziom pomiaru.
* Pomiar na tym poziomie polega na przypisaniu jednostek analizy do kategorii zmiennych, które pozostają względem siebie w pewnym porządku (kolejności).
* Kategorie zmiennej tworzą układ hierarchiczny, wyznaczający miejsce danej kategorii w tym układzie
* Miejsce kategorii w układzie oznacza jakiś potencjalny stopień nasycenia cechy (kontinuum natężenia cechy)
* Przykład cechy: wykształcenie (podstawowe, średnie, wyższe)

Poziom interwałowy:

* Wyższy od porządkowego poziom pomiaru
* Każdej jednostce analizy przypisywana jest wartość liczbowa, wyrażająca natężenie danej cechy statystycznej
* Cecha ma stopniowalny charakter, gdzie każdy stopień (interwał) wyraża w sposób liczbowy jej stopień
* Różnice między natężeniem danej cechy można wyrazić liczbowo, jako odległość między nimi i tylko w taki sposób (o ile dane jednostki różnią się od siebie)
* Punkt początkowy (0) wyznacza się arbitralnie
* Przykład cech: Skala temperatury Celsjusza

Poziom ilorazowy i absolutny

* Najwyższe poziomy pomiaru
* Każdej jednostce analizy przypisywana jest wartość liczbowa, wyrażająca natężenie danej cechy statystycznej
* Cecha ma stopniowalny charakter, lecz nie sposób w inny niż arbitralny sposób wyznaczyć stopnia natężenia cechy
* Punkt początkowy ma naturalny charakter (0 jako brak natężenia cechy)
* Różnice mogą być wyrażone proporcjami (jednostka różni się od innej o określoną proporcję razy – ileś tam razy)

Mierniki tendencji centralnej:

* Klasyczne: średnia arytmetyczna, średnia arytmetyczna ważona, średnia ważona dla szeregów przedziałowych
* Pozycyjne: mediana, dominanta

Średnia arytmetyczna

Dla szeregu punktowego:

$$\overbar{x}= \frac{\sum\_{}^{}x\_{i}}{N}$$

Średnia arytmetyczna ważona

Dla szeregu rozdzielczego:

$$\overbar{x}= \frac{\sum\_{}^{}x\_{i}n\_{i}}{N}$$

Średnia arytmetyczna

Dla szeregu przedziałowego:

$$\overbar{x}= \frac{\sum\_{}^{}°\dot{x\_{i }n\_{i}}}{N}$$

Mediana:

1. Pozycja mediany:

$$poz.Me = \frac{N+1}{2}$$

1. Mediana dla N nieparzystego:

$$Me=x\_{n}$$

1. Mediana dla N parzystego:

$$Me= \frac{\left(x\_{n-1}+x\_{n+1}\right)}{2}$$

1. Mediana dla szeregu przedziałowego:

$$Me= X\_{0}+ \left(\frac{N}{2}-\sum\_{}^{}nk\right)\frac{C}{n\_{0}}$$

I kwartyl:

1. Pozycja kwartyla:

$$poz.Q\_{1}= \frac{N+1}{4}$$

1. Dla szeregu rozdzielczego:

$$Q\_{1}= x\_{n}$$

1. dla szeregu przedziałowego:

$$Q\_{1}=X\_{0}+ \left(\frac{N}{4}-\sum\_{}^{}nk\right)\frac{C}{n\_{0}}$$

III kwartyl

1. Pozycja kwartyla:

$$poz.Q\_{3}= \frac{3N+1}{4}$$

1. Dla szeregu rozdzielczego:

$$Q\_{3}= x\_{n}$$

1. dla szeregu przedziałowego:

$$Q\_{3}=X\_{0}+ \left(\frac{3N}{4}-\sum\_{}^{}nk\right)\frac{C}{n\_{0}}$$

Rozstęp:

$$R=x\_{max}-x\_{min}$$

Odchylenie ćwiartkowe:

$$Q=\frac{Q\_{3}-Q\_{1}}{2}$$

Odchylenie ćwiartkowe – przykład z wykładu

Odchylenie przeciętne(średnie):

1. dla szeregu punktowego

$$D= \frac{\sum\_{}^{}\left|x\_{i}-\overbar{x}\right|}{N}$$

1. dla szeregu rozdzielczego

$$D= \frac{\sum\_{}^{}\left|x\_{i}-\overbar{x}\right|n\_{i}}{N}$$

1. dla szeregu przedziałowego

$$D= \frac{\sum\_{}^{}\left|°x\_{i}-\overbar{x}\right|n\_{i}}{N}$$

Odchylenie przeciętne(średnie) – przykład z wykładu

Odchylenie standardowe:

1. dla szeregów punktowych

$$s= \sqrt{\frac{\sum\_{}^{}\left(x\_{i}-\overbar{x}\right)^{2}}{N}}$$

1. dla szeregów rozdzielczych

$$s= \sqrt{\frac{\sum\_{}^{}\left(x\_{i}-\overbar{x}\right)^{2}n\_{i}}{N}}$$

1. dla szeregów przedziałowych

$$s= \sqrt{\frac{\sum\_{}^{}\left(°x\_{i}-\overbar{x}\right)^{2}n\_{i}}{N}}$$

Przykład z wykładu dla szeregu zwykłego

Przykład dla szeregu przedziałowego

Wariancja:

1. dla szeregów punktowych 
2. dla szeregów rozdzielczych 
3. dla szeregów przedziałowych: 

Klasyczny współczynnik zmienności:

$$V=\frac{s}{\overbar{x}}$$

Współczynnik zmienności dla miar położenia:

$$V\_{Q}=\frac{Q}{Me}$$

Skośność:

$$A=\frac{3\left(\overbar{x}-Me\right)}{s}$$