

z CÉŁÓÓ

1. Przedmiot zainteresowania statystyki

Przedmiotem zainteresowania statystyki są ilościowe metody badania zjawisk masowych.

2. Cechy zmienne a cechy stałe

2.1. Cechy stałe

Cechy stałe określają jednostki pod względem rzeczowym (co?), czasowym (kiedy?) oraz przestrzennym (gdzie?). Cechy stałe są wspólne wszystkim jednostkom badanej zbiorowości. Nie podlegają badaniu, jedynie decydują o zaliczeniu jednostek do określonej zbiorowości.

2.2. Cechy zmienne to właściwości, którymi różnią się poszczególne jednostki statystyczne. Dzielą się na:

- Ilościowe (niemierzalne) - nie da się ich zmierzyć, jedynie określić słownie (np. pochodzenie społeczne, płeć, kolor włosów)
- Ilościowe - można je wyrazić za pomocą liczb o różnych mianach (np. wzrost w cm, wiek w latach, zarobki w zł). Dzielą się one dalej na:
 - Zmienne skokowe - mogą się wyrażać jedynie określonymi liczbami, bez wartości pośrednich (np. liczba studentów w grupie, liczba pokoi w mieszkaniu). Cechy skokowe przyjmujące przyjmujące bardzo dużo wartości to tzw. cechy quasi ciągłe (np. zarobki w groszach)
 - Zmienne ciągłe - mogą przyjmować każdą wartość z określonego przedziału liczbowego

3. Pomiar wielkości prostej i złożonej oraz jego błąd (niepewność)

3.1. Pomiary proste (bezpośrednie)

3.2. Pomiary złożone (pośrednie)

Polegają na wyznaczaniu wartości wielkości złożonej na podstawie znanych zależności między różnymi wielkościami mierzonymi bezpośrednio, np. pomiar oporności elektrycznej metodą techniczną na podstawie zmierzonych bezpośrednio wartości natężenia i napięcia.

Niepewność pomiaru

jest miarą rozrzutu wyników powtarzanych pomiarów danej wielkości fizycznej. Zapisując wynik pomiaru fizycznego x należy wyraźnie zaznaczyć jednostkę podanej wartości i opatrzyć przedziałem niepewności Δx :

$$x \pm \Delta x$$

np. zmierzona mikrometrem średnica drutu d wynosi $d = (2,53 \pm 0,01) \text{ mm}$.

Z powodu występowania przypadkowych niepewności (błędów) pomiarowych, powstają pomiary wielkości fizycznej dają różne wyniki. Otrzymane wartości (wyniki) rozkładają się wokół wartości rzeczywistej a ich rozrzut zależy od dokładności prowadzonych pomiarów. Niepewności przypadkowe można zmniejszyć stosując dokładniejsze przyrządy i dbając o zapewnienie niezmiennych warunków doświadczenia - nie można ich jednak całkowicie uniknąć.

Błędy grube powstają wskutek fałszywego odczytu przyrządu lub ewidentnej pomyłki eksperymentatora, np. zapisanie wyniku pomiaru długości w centymetrach zamiast w milimetrach. Pomiar obarczony błędem grubym różni się zasadniczo od pozostałych wyników i można go łatwo zauważyć. Powtarzanie pomiarów pozwala zatem dostrzec i wyeliminować wyniki obciążone błędem grubym.

Błędy systematyczne wynikają z wadliwego działania przyrządu pomiarowego (np. amperomierz ze skrzywioną wskazówką, spieszący się stoper itp.) lub ze złe zaprojektowanego doświadczenia (np. waga jest ustawiona blisko grzejnika i jedno ramie jej belki jest dłuższe od drugiego). W takich przypadkach występuje stała różnica między wartościami zmierzonymi i wartością rzeczywistą. Błędy systematyczne można eliminować przez wprowadzenie poprawek lub takie projektowanie układów pomiarowych aby błędy te nie występowały.

Ocena niepewności przypadkowych

Wielokrotne niezależne powtarzanie tego samego pomiaru fizycznego pozwala otrzymać serie wyników $x_1, x_2, x_3, \dots, x_n$, gdzie n - liczba wykonanych pomiarów. Rzeczywistej wartości wielkości fizycznej nie znamy ale można wykazać, że najbardziej zbliżona do niej jest średnia arytmetyczna otrzymanych wyników \bar{x}

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

4. Skale pomiarowe wraz z omowieniem

Wyróżnia się cztery skale pomiarowe:

4.1. Skala nominalna (niemetryczna)

Najmniej precyzyjny sposób pomiaru. Liczby pełnią jedynie rolę umownych symboli służących do identyfikacji jednostek statystycznych i ich klasyfikacji do wyróżnionych kategorii. Jedyną dopuszczalną operacją matematyczną to zliczenie jednostek należących do określonej kategorii danej cechy i obliczanie np. proporcji, odsetek. O poszczególnych wariantach danej cechy można powiedzieć tylko, że są różne lub równe (np. gatunki sera, numery telefonów, grupy krwi).

4.2. Skala porządkowa (rangowa - niemetryczna)

Posiada wszystkie cechy skali nominalnej, dodatkowo pozwala na porządkowanie jednostek statystycznych w ramach wyróżnionych kategorii pod względem nateżenia badanej cechy. Można powiedzieć nie tylko, czy dane warianty są równe lub różne, ale również określić, czy jeden jest większy czy mniejszy od drugiego. W skali nominalnej liczby, zwane rangami, wyznaczają kolejność występowania jednostek, ale nie określają odległości między nimi (np. sok A jest słodszy od soku B, ale nie można określić o ile słodszy). Porządkowanie w tej skali może być słabe (np. marka A jest tak samo dobra lub mniej dobra od marki B - relacja \leq). Przykładami skali nominalnej są skala Richtera, stopnie wojskowe, wysztalcenie, itp. Skala porządkowa jest bardziej precyzyjna niż nominalna.

4.3. Skala przedziałowa (interwalowa - metryczna)

Posiada wszystkie własności skali porządkowej, a oprócz tego umożliwia określenie odległości (dys-tansu) między jednostkami. W skali przedziałowej punkt zerowy jest ustalony arbitralnie (brak zera absolutnego, np. skala Celsjusza - między 10 stopniami a 20 jest 10 stopni różnicy, ale to nie znaczy, że 20 stopni jest 2 razy cieplejsze niż 10 stopni)

4.4. Skala stosunkowa (ilorazowa)

Posiada własności trzech poprzednich skal, ale posiada naturalny punkt zerowy. Punkt zerowy oznacza brak danej cechy. Można wykonywać na nich wszystkie operacje matematyczne łącznie z dzieleniem. Przykładem może być temperatura w stopniach Kelvina, długość, ciężar, cena towaru.

5. Rodzaje badan statystycznych

- **Badania pelne (calkowite, wyczerpujace)**, obejmujace wszystkie jednostki danej zbiorowosci statystycznej
- **Badania niepelne (czesciowe)**, obejmujace niektore jednostki zbiorowosci statystycznej
Decyzja o przeprowadzeniu badania czesciowego zamiast pelnego moze byc podjeta z nastepujacych przyczyn:
 - zbiorowosc statystyczna jest tak liczna, ze badanie pelne byloby zbyt kosztowne badz wymagaloby zbyt dlugiego czasu
 - badanie ma charakter niszczaczy (np. badanie jakosci konserw)
 - chodzi jedynie o wyniki orientacyjne

Zarowno badanie pelne jak i czesciowe moga byc ciagle, okresowe lub dorazne.

- **Badania ciagle**- przeprowadzane i analizowane sukcesywnie oraz nieprzerwanie
- **Badania okresowe**- podejmowane w pewnych, zazwyczaj scisle okreslonych, odstepach czasu (np. coroczne spisy rolne czy przeprowadzane co 10 lat powszechne spisy ludnosci).
- **Badania dorazne**- przeprowadzane w pewnych szczegolnych okolicznosciach spowodowanych z reguly nieprzewidzianymi przyczynami (np. badanie dotyczace strat czasu pracy wywolanych epidemia grypy, badanie strat materialnych spowodowanych powodzią lub pozarem).
- **Szacunki interpolacyjne i ekstrapolacyjne**
 - **Interpolacja** polega na szacowaniu nieznanych wartosci cechy na podstawie znanych wartosci sasiednich (pozniejszych i wczesniejszych)
 - **Ekstrapolacja** polega na szacowaniu wartosci wykraczajacych poza przedzial wartosci znanych.

Zarowno ekstrapolacja jak i interpolacja moga miec charakter liniowy lub nieliniowy. Interpolacja i ekstrapolacja liniowa opieraja sie na zalozeniu proporcjonalnego rozkladu wartosci cechy pomiedzy liczebnosci lub jednostki czasu.

6.