

Temat:

Statystyczny opis danych

Oznaczenia

żółty – nowe pojęcie

*cz*erwony – uwaga

* - materiał nadobowiązkowy

Zagadnienia

- 1. Wprowadzenie**
- 2. Pojęcia: populacja, próba, cecha, typy cech**
- 3. Elementy opisu statystycznego:**
 - a. empiryczny rozkład wartości**
 - b. opis parametryczny**

Termin *statystyka*

Termin *statystyka*

Statystyka (łac. *status*) – stan rzeczy

Znaczenie historyczne (do poł. XIX w.)

Statystyka – uporządkowany, czyli podany w formie tabeli, zbiór danych liczbowych dotyczących stanu państwa.

Termin *statystyka*

Znaczenie współczesne

Statystyka – zgromadzone różnorakie dane liczbowe dotyczące wybranego zjawiska, np.:

- statystyka urodzeń,
- statystyka wyjazdów turystycznych,
- statystyka zachorowań na grype,
- statystyka produkcji rolnej.

Termin *statystyka*

Statystyka – dyscyplina naukowa zajmująca się tworzeniem metod służących:

- pozyskiwaniu danych,
- ich prezentacji,
- oraz analizie i interpretacji.

Celem tych metod jest otrzymanie użytecznych, uogólnionych informacji na temat badanego zjawiska.

Elementy badań statystycznych (1)

Pozyskiwanie danych to element badania statystycznego, w ramach którego dokonuje się obserwacji danego zjawiska, a następnie pomiaru lub zliczenia.

Elementy badań statystycznych (2)

Prezentacja danych polega na przedstawieniu nawet bardzo licznych zbiorów danych w takiej postaci tabelarycznej lub graficznej, która ułatwi ich ocenę i analizę.

Elementy badań statystycznych (3)

Analizę danych można przedstawić:

- metodami **statystyki opisowej** (wykonać opis statystyczny),
- metodami **statystyki matematycznej** (przeprowadzić wnioskowanie statystyczne).

UWAGA. Statystyka matematyczna bazuje na metodach rachunku prawdopodobieństwa.

Przedmiot badań statystycznych

Badanie statystyczne zawsze dotyczy pewnej zbiorowości (statystycznej), której elementami są obiekty materialne lub zjawiska, np.:

- studenci kierunku Rolnictwo I SGGW,
- drzewa owocowe w ustalonym sadzie,
- krowy wybranej rasy,
- wypadkowość w gospodarstwach rolnych danej gminy.

Przedmiot badań statystycznych

Zbiorowość statystyczną określa się mianem **populacji (generalnej)**, a jej elementy to **jednostki statystyczne**.

Populacja skończona - ma skończenie wiele jednostek statystycznych.

Populacja nieskończona - ma nieskończenie wiele jednostek statystycznych.

*

Komentarz

Przypadek populacji nieskończonej na ogół jest wynikiem przyjętego modelu zjawiska, np. populacja pomiarów wysokości pędu rośliny ustalonej odmiany, populacja rzutów monetą. Czasem populacja, choć skończona, jest tak liczna, że można ją traktować jak nieskończoną.

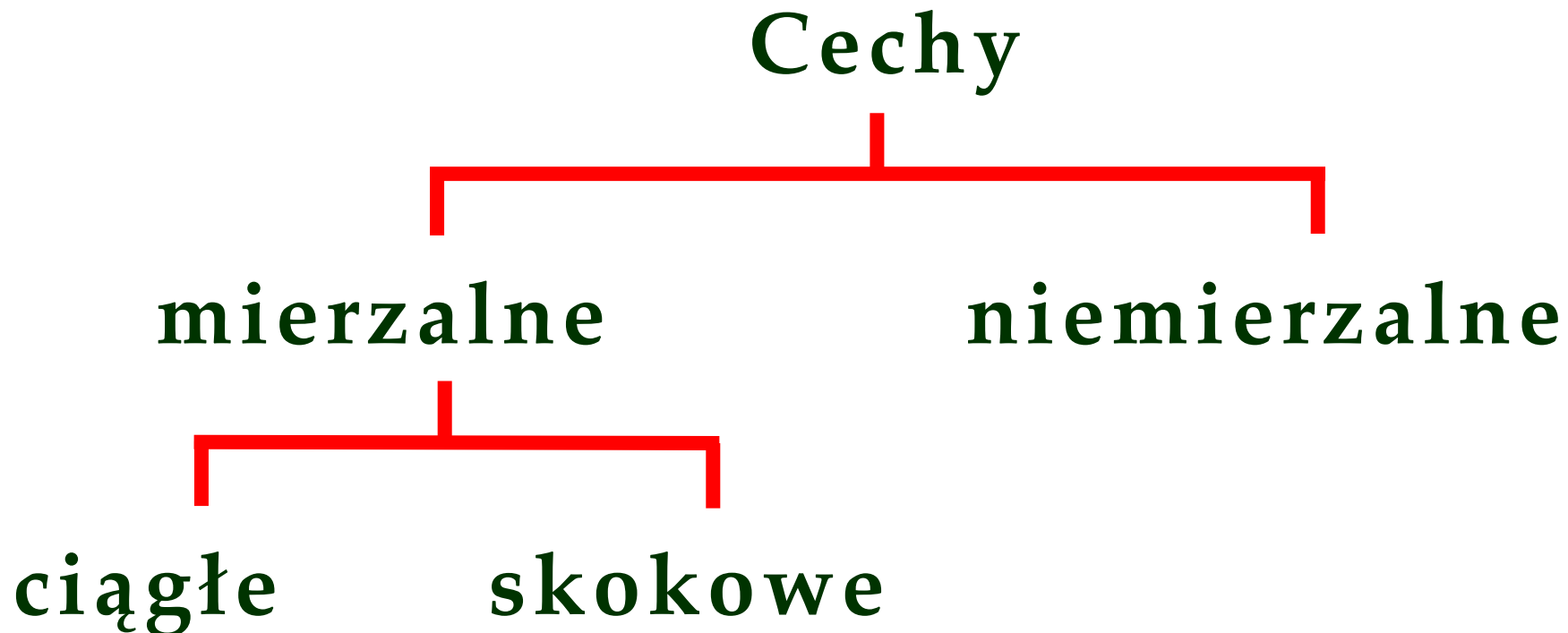
Cechy statystyczne - przykłady

Populacja - studenci kierunku
Rolnictwo I SGGW.

Jednostki statystyczne tej populacji mają właściwości, które nazywamy cechami statystycznymi, np.: płeć, wzrost, wagę, wiek, kolor oczu, liczba niezdanych egzaminów z semestru 1.

Komentarz o typach cech.

Klasyfikacja cech



Uwaga. Cechy miarzelne nazywa się także ilościowymi, a niemierzalne – jakościowymi.

Cecha mierzalna

Cechę mierzalną można otrzymać przez:

1. pomiar bezpośredni,
2. pomiar pośredni wielkości związanych z badaną cechą i dokonanie przeliczeń w oparciu o znane zależności,
3. zliczanie.

Przykłady cech mierzalnych

- wysokość drzew owocowych w sadzie,
- plon ziarna jęczmienia zebranego z poletka doświadczalnego,
- stężenie roztworu,
- średni dobowy przyrost masy zwierzęcia hodowlanego,
- liczba jabłek zepsutych w skrzynce,
- liczba szkodników znalezionych na poletku po zastosowaniu preparatu,
- liczba krów w gospodarstwie rolnym

Definicja cechy mierzalnej ciągłej

Cecha mierzalna ciągła – może przyjąć każdą wartość z określonego przedziału liczbowego, np.:

- wydajność dzienna mleka od krowy (l),
- zawartość cukru w burakach cukrowych (%).

Definicja cechy mierzalnej skokowej

Cecha mierzalna skokowa (dyskretna)
– może przyjąć tylko niektóre wartości liczbowe, np.:

- liczba prosiąt w jednym miocie,
- liczba rozgałęzień na krzewie róży,
- liczba nasion w jednym strąku grochu.

*

Uwaga

Wiele cech ciągłych „z natury”, np. wysokość rośliny, ze względu na stosowaną technikę pomiaru ma charakter cech skokowych, np. wysokość rośliny określona z dokładnością do centymetra.

Definicja cechy niemierzalnej

- Cecha niemierzalna (jakościowa) -**
każda badana jednostka statystyczna
może być zakwalifikowana do jednej
z wielu kategorii, bez przypisywania
określonej wartości liczbowej, np.:
- barwa oczu - niebieska, szara, piwna,
 - ocena nasion - zdatne lub niezdatne do siewu,
 - barwa kwiatu - czerwona, różowa, biała

Uwaga nt. cechy niemierzalnej

Kategoriom cechy niemierzalnej można przyporządkować liczby, np.:

- liczba 0 - ziarno pszenicy o sile kiełkowania mniejszej niż 90%,
- liczba 1 – pozostałe ziarna,

jednak liczby te nie wyrażają wartości cechy.

Rodzaje badań statystycznych

Badanie pełne (wyczerpujące) -
badaniu poddana jest cała populacja.

Badanie częściowe (niewyczerpujące)
- badaniu poddana jest tylko część
populacji (niektóre jednostki).

Uwaga

Badanie pełne jest niemożliwe do wykonania np. wtedy, gdy:

- populacja jest nieskończona,**
- metoda badawcza ma charakter niszczący,**
- wysoki jest koszt pojedynczego badania.**

Przykłady badań niszczących

- analiza ilości skrobi w bulwie ziemniaka,
- analiza jakości mąki pszennej,
- analiza zawartości składnika koncentratu pomidorowego w puszcze.

Próba

Próba (populacja próbna) - ta część populacji generalnej, która bezpośrednio podlega badaniu.

Próba reprezentatywna stanowi taką część populacji, która zachowuje wszelkie właściwości struktury całej populacji.

Jak otrzymać próbę reprezentatywną?

Przy pobieraniu próby z populacji należy spełnić dwa warunki:

losowość - każda jednostka populacji ma jednakową szansę dostania się do próby,

niezależność - dodawanie następnej jednostki do próby jest niezależne od pobranych wcześniej jednostek.

Badanie próby reprezentatywnej

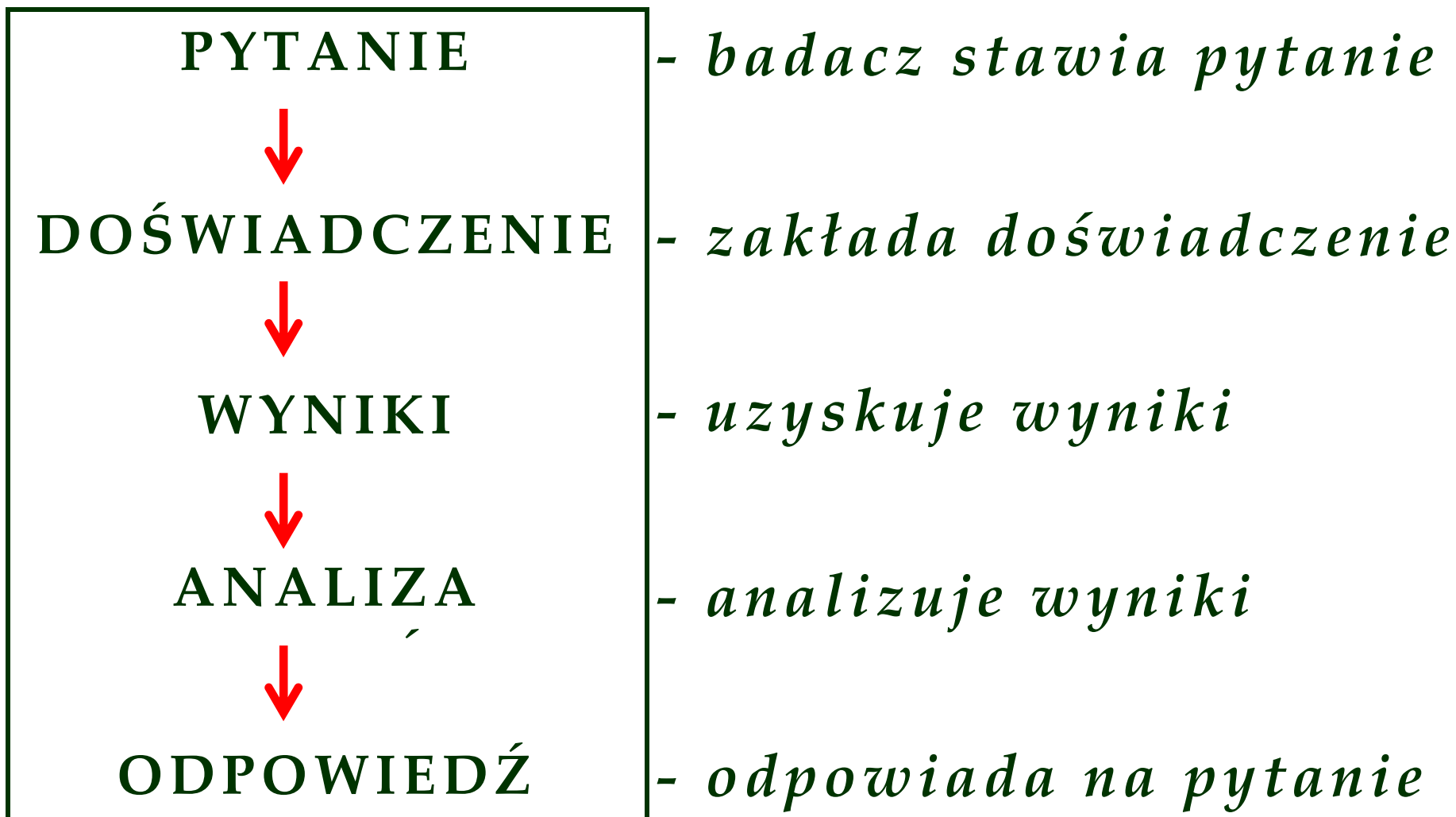
Zalety

**Badanie próby reprezentatywnej
pobranej z populacji pozwala
poprawnie wnioskować
o całej populacji.**

Źródła danych statystycznych w rolnictwie

- sprawozdania z kontroli np. udojów mleka, nieśności kur, zużycia parku maszynowego w pracach polowych
- rejestracje i spisy rolne
- doświadczenie naukowe i obserwacja

Schemat badania statystycznego



Przykład

Badamy porażenie roślin pewnego gatunku.

1. sposób pomiaru cechy

Obliczamy procent powierzchni porażonej pędów dla sześciu roślin (w %):

0,0 20,0 18,9 3,5 1,7 20,8

Badana cecha to powierzchnia porażenia, a wynikami pomiaru są liczby rzeczywiste (**cecha mierzalna ciągła**).

Przykład cd.

2. sposób pomiaru cechy

Liczymy porażone pędy dla sześciu roślin:

0 4 3 2 1 4

Badana cecha to liczba pędów porażonych,
a wynikami pomiaru są liczby naturalne
(**cecha mierzalna skokowa**).

Przykład cd.

3. sposób pomiaru cechy

Obserwujemy, czy roślina została porażona, czy nie. Wyniki dla sześciu roślin:

O P P P P P

Badana cecha to fakt porażenia rośliny, a wyniki pomiaru w ogóle nie są wyrażone liczbowo! P - oznacza roślinę porażoną, O - roślinę odporną (**cecha niemierzalna**).

Przykład cd.

W każdym z przytoczonych przykładów ogólny cel działania był określony jednakowo – badanie porażenia. Ale za każdym razem inaczej określono cechę: powierzchnia porażonej rośliny, liczba porażonych pędów na roślinie, fakt porażenia lub nie rośliny. Za każdym razem inny był typ badanej cechy. Od typu cechy (typu uzyskanych wyników) zależy dobór metody analizy, więc przy późniejszej prezentacji metod zawsze pojawi się ważna informacja, dla jakiego typu cechy dana metoda jest przeznaczona.

Opis statystyczny

Opis statystyczny

Opis statystyczny polega na przedstawieniu:

- struktury zbiorowości
- współzależności cech
- zmian zjawisk w czasie

Formy opisu statystycznego

- tabelaryczna (dane liczbowe zestawione w szeregi i tablice)
- graficzna (dane w postaci wykresów)
- parametryczna (parametry to charakterystyki liczbowe)

Opis statystyczny – zadanie

Dla wskazanych danych utwórz:

- a.** szereg rozdzielczy
- b.** histogram
- c.** wielobok częstości
- d.** dystrybuantę empiryczną

Tworzenie szeregu rozdzielczego

1. ustalamy zakres wartości,
2. ustalamy liczbę klas,
3. zakres wartości dzielimy na klasy,
4. wyznaczamy liczebność (lub częstość) w każdej klasie,
5. przedstawiamy klasy w postaci zestawienia w tabeli (szereg rozdzielczy) lub graficznie (histogram, wielobok częstości).

Liczba klas szeregu rozdzielczego

Ozn.: k - liczba klas, n - liczebność próby

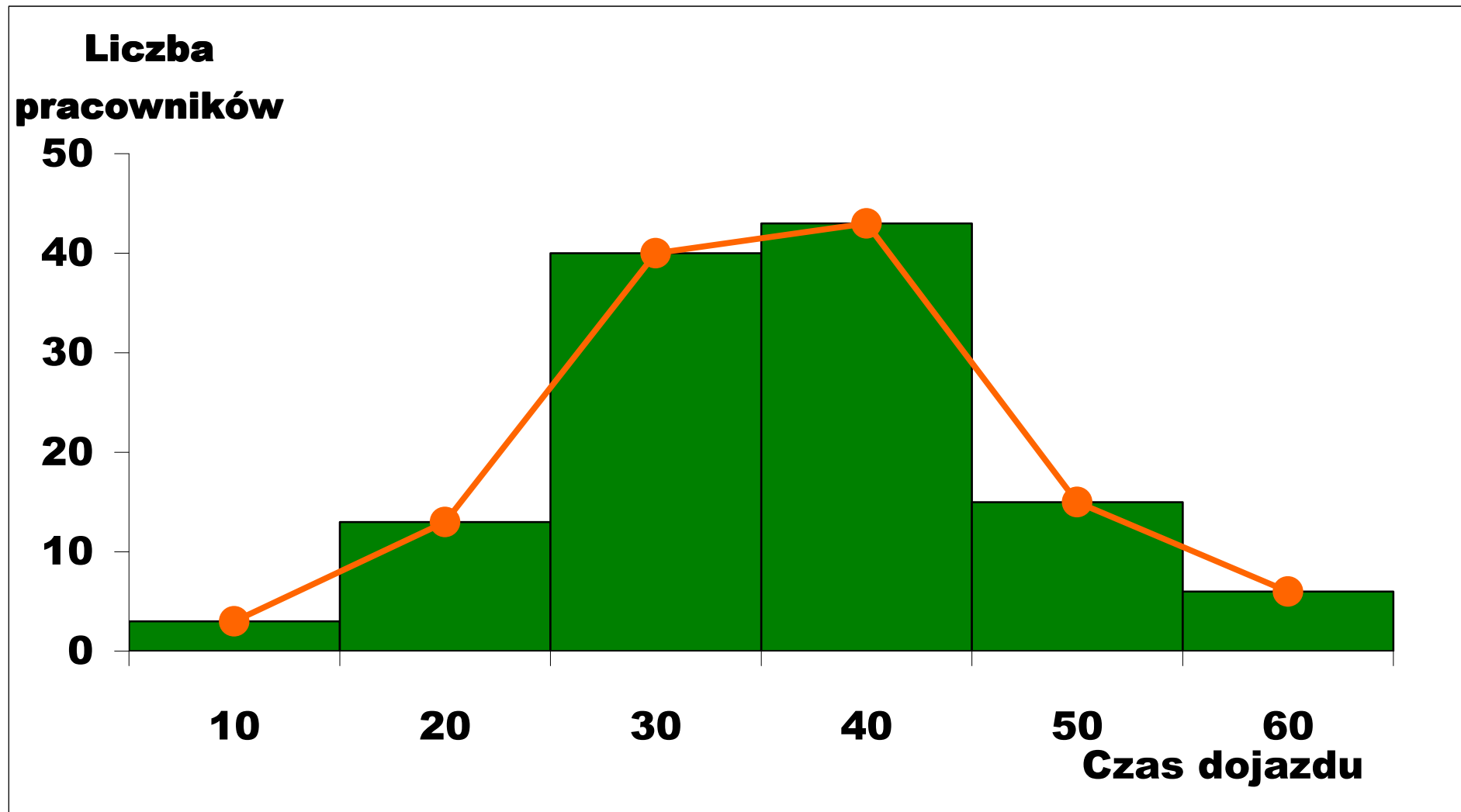
n	k
30 - 60	6 - 8
60 - 100	7 - 10
100 - 200	9 - 12
200 - 500	11 - 17
500 - 1500	16 - 25
przeważnie najwyżej 30	

Szereg rozdzielczy – przykład 1

Określono czasu dojazdu do pracy pracowników pewnego zakładu.

Czas dojazdu w minutach	Liczba pracowników
5 - 15	3
15 - 25	13
25 - 35	40
35 - 45	43
45 - 55	15
55 - 65	6
Ogółem	120

Histogram, wielobok cz. – przykład 1



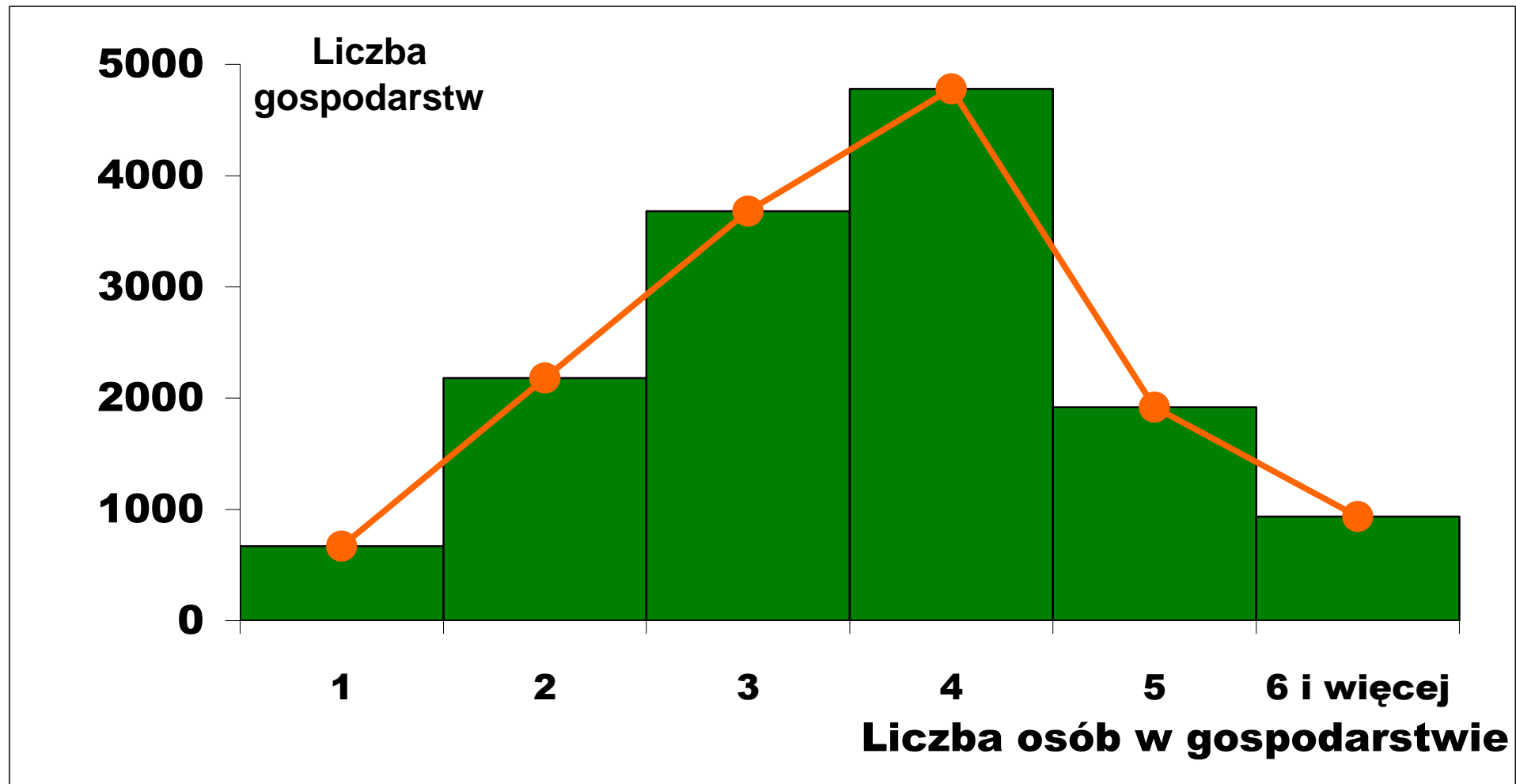
Histogram (zielone słupki)
Wielobok częstości (pomarańczowa łamana)

Szereg rozdzielczy – przykład 2

Gospodarstwa domowe według wielkości

Liczba osób w gosp. domowym	Liczba gospodarstw
1	670
2	2 182
3	3 681
4	4 779
5	1 918
6 i więcej	937
Ogółem	14 167

Histogram, wielobok cz. – przykład 2



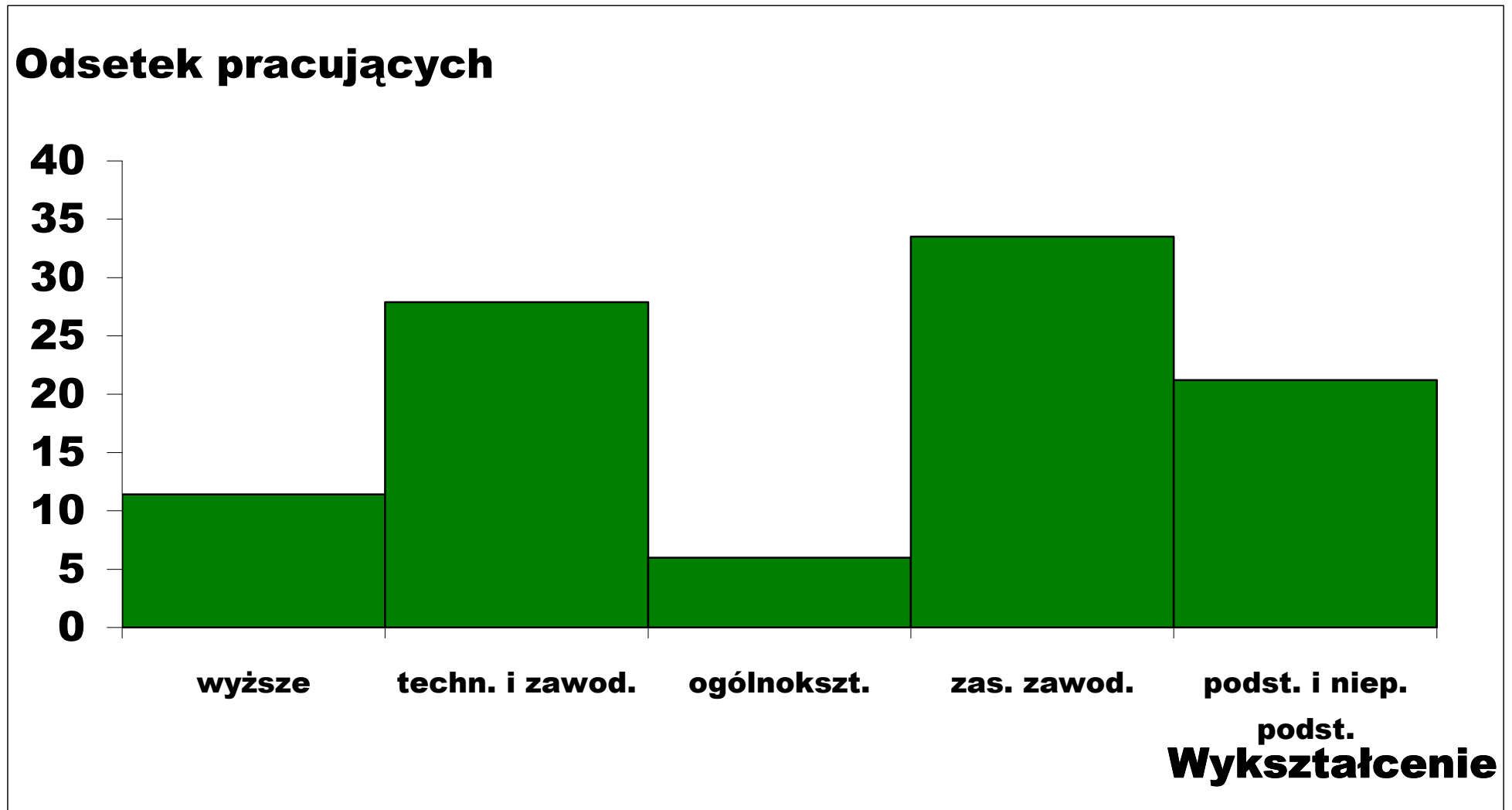
Histogram
Wielobok częstości

Szereg rozdzielczy – przykład 3

Struktura pracujących według poziomu wykształcenia.

Wykształcenie	Odsetek pracujących
Wyższe	11,4
Średnie:	
techniczne i zawodowe	27,9
ogólnokształcące	6,0
zasadnicze zawodowe	33,5
Podstawowe i niepełne podstawowe	21,2
Razem	100,0

Histogram – przykład 3



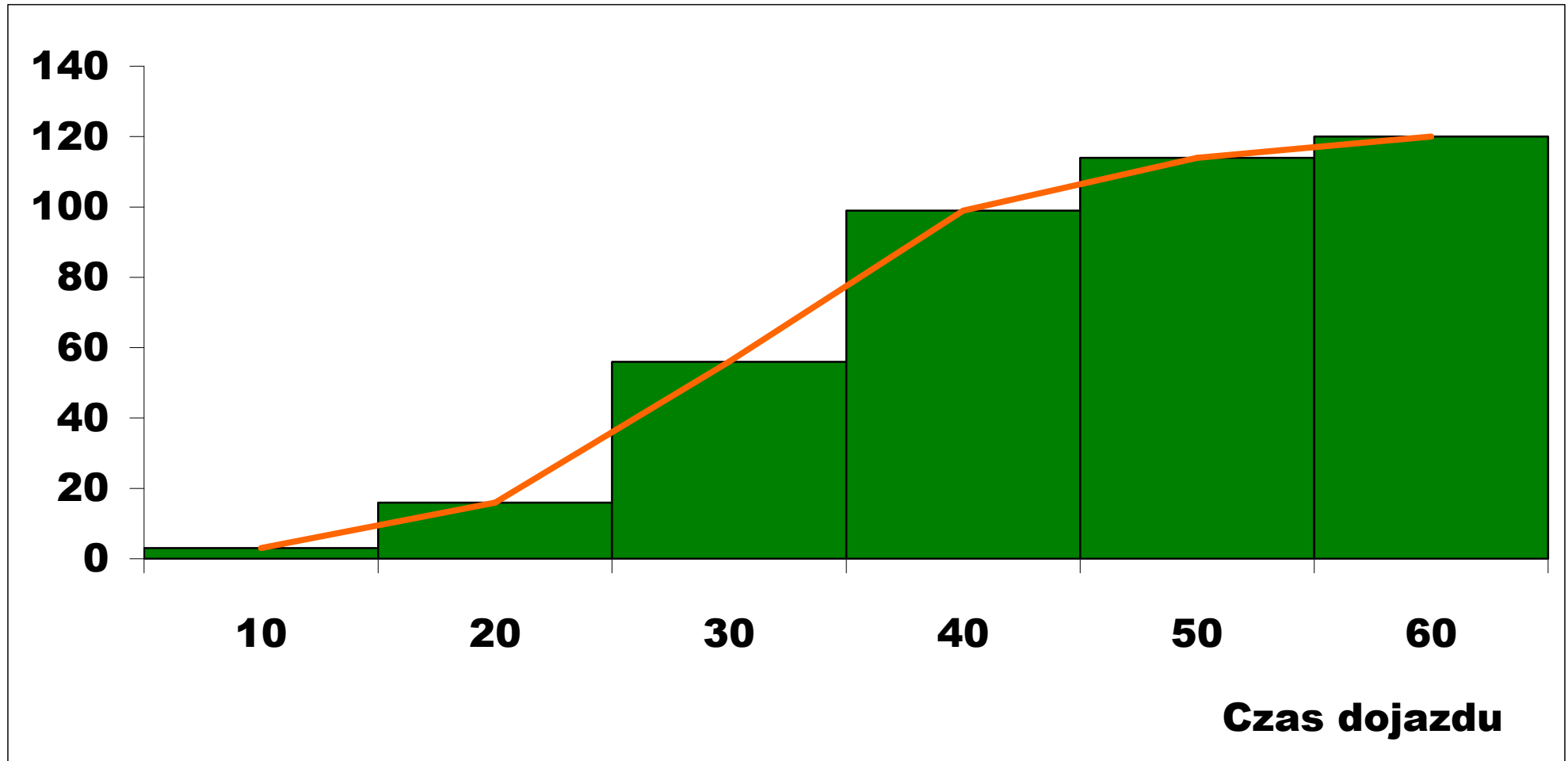
Histogram

Szereg rozdzielczy skumulowany

Przykład. Pracownicy pewnego zakładu według czasu dojazdu do pracy.

Czas dojazdu w minutach	Liczba pracowników	Liczba pracowników narastająco
5 - 15	3	3
15 - 25	13	16
25 - 35	40	56
35 - 45	43	99
45 - 55	15	114
55 - 65	6	120
Ogółem	120	

Histogram skumulowany - przykład



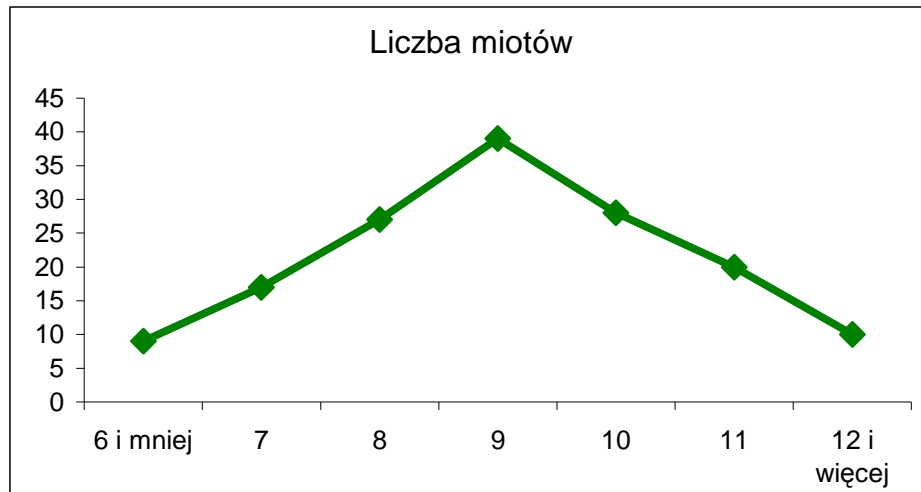
Histogram skumulowany
Dystrybuanta empiryczna (!)

Rozkłady emp. cechy mierzalnej

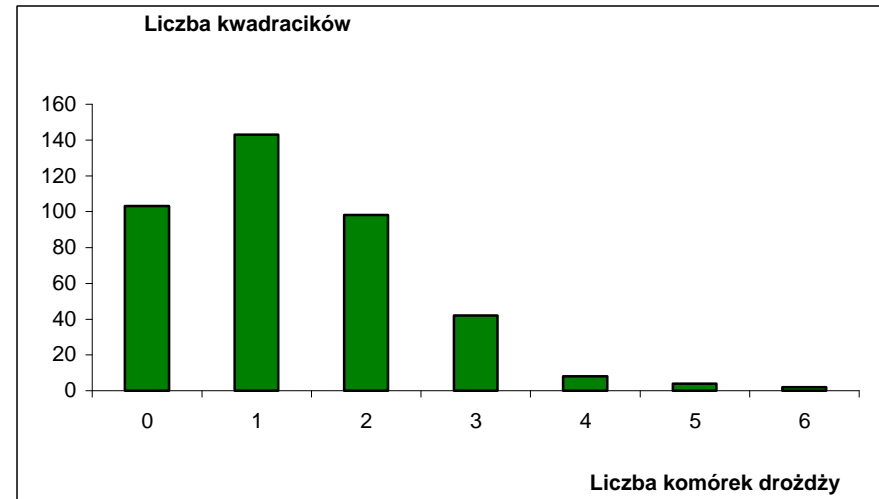
Typy rozkładów

Ze względu na rozmieszczenie liczebności badanych jednostek przy odpowiednich wartościach cechy mierzalnej.

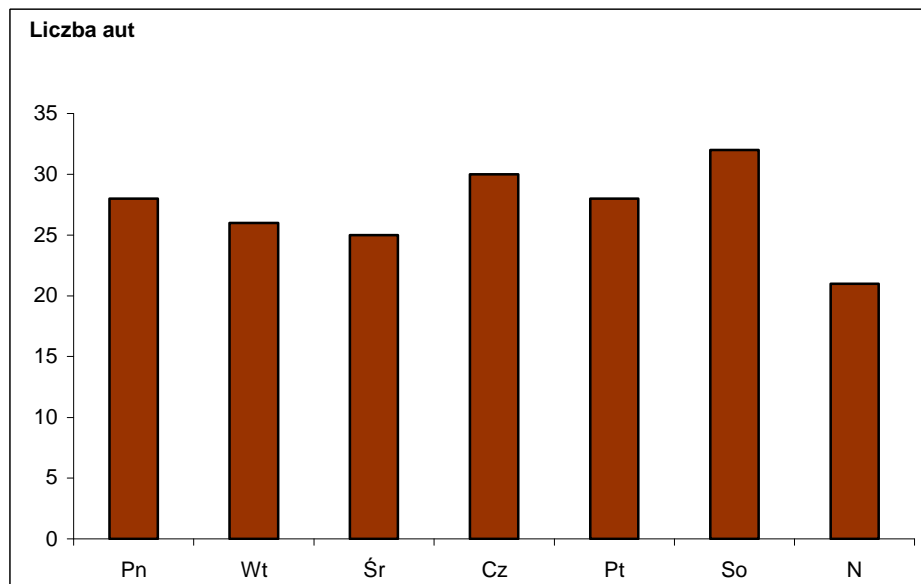
Przykłady typów rozkładów



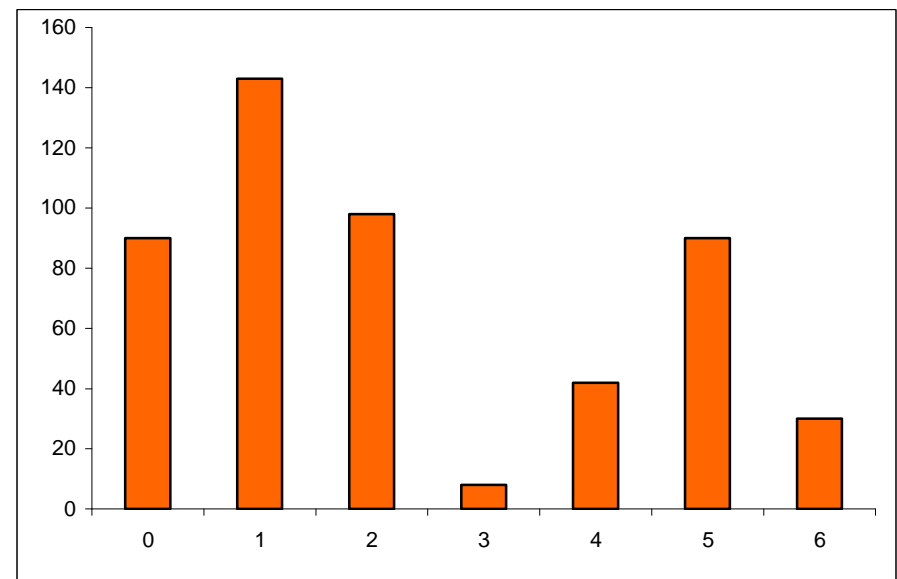
Rys. 1. R. symetryczny



Rys. 2. R. asymetryczny

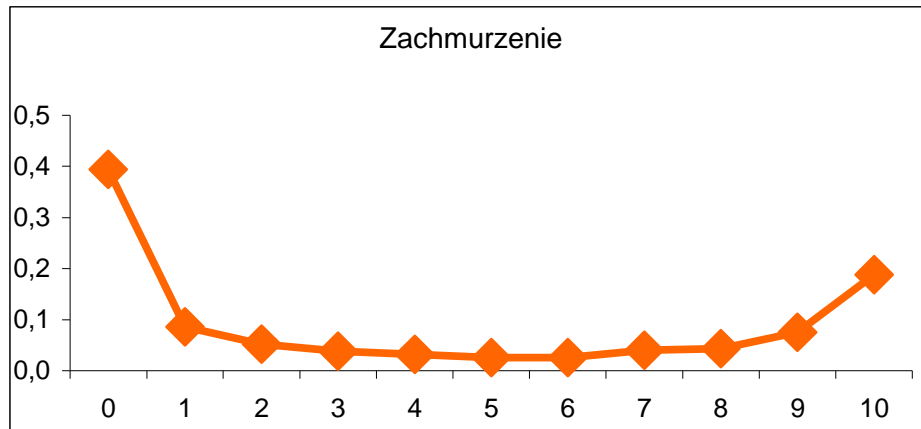


Rys. 3. R. równomierny



Rys. 4. Rozkład bimodalny

Przykłady typów rozkładów, cd.



Rys. 5. Rozkład typu U