

I w o n a J a ż d ż e w s k a

STATYSTYKA

Podręcznik dla studentów turystyki i rekreacji



STATYSTYKA

Podręcznik dla studentów turystyki i rekreacji



WYDAWNICTWO
UNIWERSYTETU
ŁÓDZKIEGO

I w o n a J a ż d ż e w s k a

STATYSTYKA

Podręcznik dla studentów turystyki i rekreacji

Iwona Jażdżewska – Uniwersytet Łódzki, Wydział Nauk Geograficznych
Instytut Geografii Miast i Turyzmu, Zakład Geoinformacji, 90-142 Łódź, ul. Kopcińskiego 31

RECENZENT

Jerzy Runge

REDAKTOR INICJUJĄCY

Beata Koźniewska

REDAKTOR WYDAWNICTWA UŁ

Katarzyna Gorzkowska

SKŁAD I ŁAMANIE

Munda – Maciej Torz

OPRACOWANIE TECHNICZNE RYSUNKÓW

Anna Wosiak

PROJEKT OKŁADKI

Katarzyna Turkowska

Zdjęcie wykorzystane na okładce autorstwa Iwa Hencza

© Copyright by Iwona Jażdżewska, Łódź 2019

© Copyright for this edition by Uniwersytet Łódzki, Łódź 2019

Wydane przez Wydawnictwo Uniwersytetu Łódzkiego

Wydanie I. W.07409.16.0.S

Ark. wyd. 11,0; ark. druk. 15,75

ISBN 978-83-8142-535-3

e-ISBN 978-83-8142-536-0

Wydawnictwo Uniwersytetu Łódzkiego

90-131 Łódź, ul. Lindleya 8

www.wydawnictwo.uni.lodz.pl

e-mail: ksiegarnia@uni.lodz.pl

tel. (42) 665 58 63

SPIS TREŚCI

Wprowadzenie	7
Rozdział 1. Zagadnienia wstępne	9
1.1. Podstawowe pojęcia statystyczne	9
1.2. Skale pomiaru	16
1.3. Metoda reprezentacyjna	21
1.4. Techniki zbierania informacji	25
1.5. Zadania	29
1.6. Odpowiedzi do wybranych zadań	35
Rozdział 2. Prezentacja danych statystycznych	39
2.1. Szeregi statystyczne	40
2.2. Tablice statystyczne	50
2.3. Graficzna prezentacja danych statystycznych	53
2.4. Zadania	68
2.5. Odpowiedzi do wybranych zadań	80
Rozdział 3. Rozkłady zmiennych losowych i ich własności	83
Rozdział 4. Analiza jednej zmiennej	89
4.1. Miary średnie	89
4.2. Miary rozproszenia	105
4.3. Miary asymetrii i koncentracji	114

4.4. Zadania	128
4.5. Odpowiedzi do wybranych zadań	137
Rozdział 5. Analiza korelacji i regresji	143
5.1. Analiza korelacji	144
5.2. Analiza regresji	170
5.3. Zadania	178
5.4. Odpowiedzi do wybranych zadań	185
Rozdział 6. Analiza dynamiki	191
6.1. Wykresy dynamiki	191
6.2. Wskaźniki dynamiki	196
6.3. Wyznaczanie tendencji rozwojowych	204
6.4. Zadania	211
6.5. Odpowiedzi do wybranych zadań	221
Podsumowanie – propozycja etapów badania statystycznego	227
Literatura	233
Załączniki	237
Indeks terminów	245

WPROWADZENIE

W badaniach obejmujących szeroko rozumianą turystykę występuje wiele informacji, które zazwyczaj, choć nie zawsze, przyjmują formę liczb. Niekiedy jest ich kilka i można je natychmiast poddać analizie, jednak z czasem ich przybywa – np. liczba informacji wzrasta do kilkuset tysięcy i wtedy, aby je zinterpretować czy wykryć pewne prawidłowości, trzeba postąpić się procedurami badawczymi oferowanymi przez statystykę.

Celem autorki niniejszego podręcznika jest przedstawienie podstawowej wiedzy na temat metod statystycznych oraz umiejętności ich wykorzystania do analizy i oceny wyników prowadzonych badań statystycznych przez studentów różnych kierunków studiów związanych z turystyką. Wiedza ta może być przydatna podczas przygotowywania pracy licencjackiej lub magisterskiej.

Zanim rozpocznie się własne badania, bardzo ważne jest postawienie wielu pytań. Na początek należy określić, po co je wykonujemy. Czy chcemy opisać jedno zjawisko, czy kilka? Jaki problem chcemy rozwiązać? Czy badania będą statyczne (jednokrotne), czy dynamiczne (prowadzone w ciągu kilku lat)? Czy chcemy wskazać związki między badanymi zmiennymi, czy różnice między nimi? Jakimi danymi dysponujemy? Czy przeprowadzimy badania całej populacji, czy określonej próby (badania częstkowe)? Jakie metody

statystyczne możemy wykorzystać? Jakie narzędzia komputerowe są w zasięgu naszych zdolności i możliwości wykorzystania? Jeśli analizujemy wyniki innych autorów (gdy ktoś już podobne badania wykonał), powinniśmy mieć możliwość oceny ich rzetelności, prawidłowo sformułowanych hipotez, zebranych danych źródłowych oraz metod statystycznych, które pozwoliły autorowi na uogólnione wnioski.

Przetwarzanie danych statystycznych wymaga znajomości metod, a także odpowiedniego oprogramowania komputerowego, które eliminuje czasochłonne obliczenia. Można sobie jednak poradzić w sytuacjach, gdy nie mamy go pod ręką i trzeba będzie wykonać obliczenia proste. Ich wykonanie z użyciem kalkulatora lub bez niego pozwala na lepsze zrozumienie stosowanej procedury. Metody tych obliczeń będą prezentowane na przykładach.

W tym miejscu chciałabym podziękować Studentkom i Studentom kierunku turystyka i rekreacja, kształcącym się na Wydziale Nauk Geograficznych Uniwersytetu Łódzkiego, z którymi miałam przez wiele lat przyjemność prowadzić zajęcia ze statystyki i którzy byli pierwszymi czytelnikami tekstu, a także jako pierwsi rozwiązywali zadania oraz stosowali wybrane metody w pracach magisterskich.

Autorka

ROZDZIAŁ 1

ZAGADNIENIA WSTĘPNE

Do rozwiązywania zadań ze statystyki potrzebna jest umiejętność posługiwania się podstawowymi operacjami, symbolami i oznaczeniami matematycznymi, takimi jak np.: +, -, ×, /, $\sqrt{\quad}$, !, %, ‰, <, >, ≤, ≥, ≠, ∈, ∉, ∑ (suma), ∏ (iloczyn), pojęciami: macierz, zbiór liczb naturalnych, całkowitych, rzeczywistych, a także znajomość kilku liter greckich wykorzystywanych w matematyce: α , β , γ , ν , μ , δ , Σ , σ , π , Π , ε , ξ , χ , ϕ (warto sprawdzić, jak się je czyta).

1.1. PODSTAWOWE POJĘCIA STATYSTYCZNE

Jednym z początkowych etapów badania statystycznego powinno być precyzyjne określenie **podmiotu badań** pod względem **rzeczowym** (co lub kogo badamy?), **czasowym** (w jakim okresie lub kiedy odbywają się badania?) oraz **przestrzennym** (gdzie one się odbywają?). Po spełnieniu tych założeń można przystąpić do określenia **zbiorowości statystycznej** (populacji statystycznej¹), którą jest **ogół elementów poddanych badaniu statystycznemu**. Z kolei **jednostkami statystycznymi** są elementy zbiorowości statystycznej

1 Obydwa określenia są równorzędne, lecz niekiedy przyjmuje się, że **populacja statystyczna** (od łac. *populatio* – ludność) dotyczy zbioru ludności.

powiązane ze sobą logicznie, tak aby można je było przyporządkować danej populacji. Każdą z jednostek charakteryzują określone cechy i ich wartości (przykład 1.1.1).

Przykład 1.1.1

Można mówić np. o zbiorowości turystów w Ciechocinku w maju 2013 r. lub hoteli w Londynie w grudniu 2015 r. W każdej z tych zbiorowości występują jednostki statystyczne (turysta, hotel), które wyróżniają się określonymi cechami o określonych wartościach (matematycznych).

Jeśli jednostką statystyczną jest turysta, to cechą wspólną łączącą ją z innymi osobami przebywającymi w Ciechocinku jest fakt, że przyjechał tu w celach turystycznych, a cechami różniącymi są np. takie, jak: wiek, płeć, miejsce zamieszkania, liczba dzieci, z którymi przyjechał, wydatki poniesione podczas pobytu w uzdrowisku, stopień zadowolenia z usług, z jakich korzystał.

Jeśli jednostką statystyczną jest hotel, to cechą wspólną łączącą ją z innymi obiektami noclegowymi jest funkcja, jaką ma do spełnienia, czyli udzielenie noclegu, a cechami różniącymi są m.in.: liczba miejsc noclegowych, liczba zatrudnionych, kategoria (liczba gwiazdek), położenie geograficzne, koszty noclegów, wyposażenie w dodatkowe usługi, np. bar, fryzjer, Wi-Fi.

Z jakimi zbiorowościami możemy spotkać się w badaniach turystycznych? Wbrew pozorom odpowiedź na to pytanie nie jest prosta, gdyż turystyka to zjawisko złożone i może być badana w wielu aspektach.

Nie byłoby turystyki, gdyby nie było osób chętnych do jej uprawiania, dlatego najczęściej wymienianą i badaną populacją są właśnie turyści. Wraz z innymi **osobami związanymi z turystyką** stanowią oni dużą grupę, którą można podzielić m.in. na następujące kategorie:

- a) turyści,
- b) pracownicy zatrudnieni w przedsiębiorstwach turystycznych,
- c) urzędnicy – wydziały promocji i turystyki w administracji publicznej,
- d) pracownicy organizacji turystycznych (np. Polskiej Organizacji Turystycznej – POT, Polskiego Towarzystwa Turystyczno-Krajoznawczego – PTTK),
- e) duchowni zajmujący się pielgrzymami,

- f) mieszkańcy obszarów, na które przyjeżdżają turyści,
- g) pracownicy innych instytucji związanych pośrednio lub bezpośrednio z turystyką,
- h) inne osoby.

Działalnością turystyczną zajmuje się wiele podmiotów – **przedsiębiorstw**, które mogą być również podmiotami zainteresowania i badania statystycznego. Część z nich związana jest bezpośrednio lub pośrednio z turystyką. Można je podzielić na następujące branże²:

- hotelarstwo (cała baza noclegowa),
- gastronomia (wszystkie typy),
- transport kolejowy, lądowy, wodny, lotniczy (obsługa pasażerów i bagażu),
- organizatorzy turystyki, agenci i pośrednicy turystyczni,
- pośrednictwo finansowe (ubezpieczenie podróżnych, wymiana walut),
- administracja publiczna (informacja i administracja turystyczna, wyspecjalizowane oddziały policji, straży granicznej, ochrona lotnisk, wydawanie wiz, pozwoleń itd.),
- organizacje i instytucje turystyczne,
- handel (sprzedaż pamiątek, literatury i map turystycznych),
- kultura (obiekty muzealne, galerie, teatry, parki rozrywki),
- sport i rekreacja (obiekty sportowo-rekreacyjne, ogrody botaniczne i zoologiczne),
- zdrowie i uroda (gabinety odnowy, kosmetyczne, SPA),
- edukacja (studia w zakresie turystyki, hotelarstwa, rekreacji oraz kursy pilotażu, żeglarskie i kursy dla turystów, np. jazdy na nartach),
- inne.

Wyjazdy turystyczne mogą być związane z różnymi **wydarzeniami**, które są elementami zbiorowości statystycznej. Można badać wydarzenia:

- 1) kulturalne – muzealne, folklorystyczne, muzyczne, teatralne, filmowe itd.,

2 Światowa Organizacja Turystyki (UNWTO) w zakres gospodarki turystycznej włącza 10 sektorów gospodarki, które są zgodne z Międzynarodową Klasyfikacją Turystyczną (SICTA – Standard International Classification of Tourism Activities).

- 2) sportowe – jednej z dyscyplin (np. tenisa) lub kilku (np. lekkoatletyczne, olimpijskie),
- 3) religijne – zależą one od religii; obejmują święta religijne cykliczne, jubileusze, święta patronów (odpusty) i inne,
- 4) ekonomiczne – kongresy, szkolenia, targi itd.,
- 5) inne.

Turysta opuszcza czasowo swoje miejsce zamieszkania i udaje się do innej miejscowości, regionu, kraju. Przedmiotami badań w geografii turystyki są zagadnienia związane z kilkoma aspektami, do których wykorzystuje się odpowiednie metody badań. Są wśród nich badania dotyczące walorów turystycznych (przyrodniczych i antropogenicznych), zagospodarowania turystycznego, ruchu turystycznego.

Przestrzeń geograficzna regionów turystycznych może być przedmiotem badań geograficznych, a badaniem mogą być objęte struktury i procesy obserwowane w jednostkach przestrzennych. Ich położenie da się określić w różny sposób, np.:

- a) administracyjny, uwzględniający m.in. podział na:
 - jednostki administracyjne (w Polsce: województwa, powiaty gminy),
 - jednostki osadnicze (miasto, wieś),
 - obszary chronione (np. parki narodowe, rezerваты);
 - państwa (europejskie, należące do UE, z innych kontynentów),
- b) przyrodniczy, z punktu widzenia np.:
 - strefy klimatycznej (międzyzwrotnikowej, umiarkowanej itd.),
 - ukształtowania terenu (górski, wyżynny, nizinny),
 - położenia względem akwenów i sieci rzecznej,
 - typów krajobrazów przyrodniczych;
- c) z uwzględnieniem pełnionych funkcji:
 - regiony turystyczne,
 - regiony rolnicze,
 - regiony ekonomiczne.

Przedstawione przykłady zbiorowości statystycznych, które mogą być objęte badaniami z zakresu przestrzeni geograficznej, wskazują na szerokie spektrum dziedzin naukowych i związanych z nimi metod badawczych. Podstawowe metody statystyczne są identyczne w każdej z dyscyplin, jednak warto sięgnąć po podręczniki dedykowane naukom społecznym (Frankfort-Nachmias, Nachmias 2001; Babbie 2004, 2013; Francuz, Mackiewicz 2005; Bedyńska, Brzezicka-Rotkiewicz, red. 2007), ekonomicznym (Luszniewicz, Słaby 1996; Pocięcha 2002), geograficznym (Jażdżewska 2013; Runge 2007), politologicznym (Mider, Marcinkowska 2013).

Kolejnym etapem w części wstępnej badań jest określenie **obszaru badań**, na którym się one odbywają. Najczęściej bywa on wskazywany jako jednostka administracyjna, np. województwo lubelskie, miasto Kraków, państwo Słowacja. Nie jest to jedyna możliwość zdefiniowania miejsca, w którym przeprowadzono badania statystyczne. W zależności od potrzeb, skali badań i umiejętności można posługiwać się różnymi procedurami **określenia położenia geograficznego**, np. takimi, jak:

- nominalną (podaje się nazwę kraju, miasta lub obiektu – np. Polska, Kraków, Wawel),
- współrzędnych lokalnych (określa się obiekt i odległość – np. w promieniu 500 m od szlaku turystycznego, 50 m od rzeki Warty),
- porządkową (numeracja domów – np. wzdłuż ulicy Piotrkowskiej w Łodzi, od numeru 1 do 63),
- współrzędnych geograficznych (długość i szerokość geograficzna),
- topologiczną (czyli sąsiedztwa obiektów – np. wszystkie państwa graniczące z Morzem Bałtyckim).

Kiedy już zostanie określone, co lub kto będą poddane badaniom i gdzie będą się one odbywały, należy wskazać kiedy się one odbędą (przykład 1.2). W statystyce można wykorzystać dwa sposoby **określenia czasu badań**. Pierwszy polega na podaniu okresu badań, np. tydzień, miesiąc, kwartał, cały rok. Jest to badanie ciągłe zjawiska, m.in. przychodów w hotelu w ciągu roku. Drugi dotyczy badania wykonywanego najczęściej w konkretnym dniu, np. 31 grudnia, lub w dniu występowania wydarzenia turystycznego, np. mecz finałowy w Mistrzostwach Świata w siatkówce kobiet.

Przykład 1.1.2

Jeśli zadaniem badawczym jest wskazanie wielkości ruchu turystycznego w Sopocie w ciągu całego roku, to mamy do czynienia z pierwszym sposobem. Jeśli zadaniem jest podanie liczby pielgrzymów w dniu święta maryjnego (15 sierpnia) w jednym z sanktuariów maryjnych – wykorzystuje się drugi sposób.

W niektórych przypadkach badanie całej zbiorowości statystycznej – nazywane **badaniem pełnym lub wyczerpującym** – jest bardzo kosztowne (np. Narodowy Spis Powszechny – NSP³) lub niemożliwe do zrealizowania (np. badanie pielgrzymów podczas mszy, kontrola jakości produktów spożywczych). Wówczas wykonuje się **badania częściowe lub cząstkowe** na próbie losowej lub nielosowej (szerzej w podrozdziale 1.3).

W zależności od tego, jaki charakter miały badania, wyróżnia się dwa podejścia w statystyce: **statystykę opisową**, gdy mamy do czynienia z całą zbiorowością, lub **wnioskowanie statystyczne**, gdy wnioskujemy na temat całej populacji na podstawie pomiarów z próby⁴.

W statystyce występują liczby **bezwzględne (absolutne)** i **względne**. Liczby bezwzględne są to wielkości, które otrzymujemy w wyniku mierzenia lub sumowania jednostek zbiorowości. Każda z nich, wyrażająca rozmiar badanego zjawiska, jest mianowana, np. koszty noclegu w złotych, dolarach amerykańskich, rublach lub euro. Liczby względne powstają przez porównanie ze sobą dwóch liczb. Odgrywają one ważną rolę przy porównywaniu zjawisk. Są to wielkości procentowe lub wskaźniki natężenia opisujące relacje między różnymi cechami, np. gęstość zaludnienia, przeciętne wydatki turysty na noclegi.

Przykład 1.1.3

Dla jednostek statystycznych z przykładu 1.1.1 wiek turysty będzie wartością bezwzględną, a średnie wydatki codzienne poniesione podczas pobytu

-
- 3 Nie każdy NSP jest badaniem pełnym, niekiedy były to również badania częściowe, np. NSP w 2011 r. w Polsce.
 - 4 W podręczniku zaprezentowana będzie głównie statystyka opisowa, z elementami wnioskowania statystycznego.

w uzdrowisku – wartością względną. Liczba miejsc noclegowych w hotelu będzie wartością bezwzględną, a procent wykorzystania miejsc noclegowych w ciągu roku – wartością względną.

Własności, jakimi charakteryzują się jednostki statystyczne nazywamy **cechami statystycznymi**. Przystępując do badań statystycznych, określamy, ile cech będziemy analizować. Zbiorowość może być jednocechowa (jednowymiarowa) i wielocechowa. Zbiór cech dzielimy na **mierzalne (ilościowe)** i **niemierzalne (jakościowe)**.

Cechy mierzalne można przedstawić za pomocą liczb i jednostek miary, np. koszt podróży (zł), długość podróży (km), czas podróży (dni), waga bagażu (kg), liczba osób towarzyszących (osoby), wiek turysty (lata). Dzieli się je na **ciągłe** i **skokowe**. Zmienna skokowa może przyjmować wartości z określonego przedziału liczb zbioru liczb naturalnych lub całkowitych nieujemnych, np. liczba osób towarzyszących, liczba dzieci, liczba bagażu. Zmienna ciągła może przyjmować wszystkie wartości z określonego przedziału liczb zbioru liczb rzeczywistych, np. koszt biletu, odległość od miejsca zamieszkania.

Przykład 1.1.4

Dla jednostek statystycznych z przykładu 1.1.1

Wydatki dzienne turysty poniesione podczas pobytu w uzdrowisku są cechą mierzalną ciągłą. Liczba dzieci, z którymi przyjechał, są cechą mierzalną skokową. Płeć jest cechą niemierzalną. Stopień zadowolenia z usług, z jakich korzystał, jest **cechą niemierzalną stopniowalną**.

Liczba miejsc noclegowych w hotelu jest cechą mierzalną skokową, a dochody miesięczne są cechą mierzalną ciągłą. Położenie geograficzne, np. ulica, jest cechą niemierzalną, a kategoria hotelu (liczba gwiazdek) jest cechą niemierzalną stopniowalną, wyposażenie w dodatkowe usługi (np. fryzjer, bar, Wi-Fi) jest **cechą niemierzalną dwudzielną**.

1.2. SKALE POMIARU

W statystyce operuje się pojęciem **skali pomiarowej**. Jej zrozumienie jest niezwykle ważne, gdyż część metod statystycznych można wykonywać jedynie dla danych w określonej skali. Wykorzystanie – często bezmyślne – programów komputerowych dla danych w złej skali pomiarowej prowadzi do błędnych wniosków i nierzetelnych analiz statystycznych. Dlatego niezbędna jest znajomość skali, w jakiej prezentowane są dane. Wyróżnia się następujące skale pomiarowe: nominalną, porządkową, interwałową i ilorazową. Cechy jakościowe są mierzone w skalach nominalnej i porządkowej, które nazywane są skalami słabymi. Cechy ilościowe są mierzone w skalach interwałowej i ilorazowej, które nazywane są skalami mocnymi. Im silniejsza jest skala pomiaru, tym więcej metod statystycznych ma do dyspozycji badacz.

Skalę nominalną wykorzystuje się w badaniach jakościowych; określa ona np.:

- a) narodowość: Francuz, Niemiec, Polak, Portugalczyk, Rosjanin, Węgier,
- b) wyznanie: chrześcijanie, muzułmanie, żydzi,
- c) użytkowanie ziemi: lasy, tereny zabudowane, tereny komunikacyjne, wody,
- d) własność: państwowa, prywatna, kościelna, gminna,
- e) branże w gospodarce: przemysł, rolnictwo, turystyka.

Należy pamiętać, że każda jednostka statystyczna badanej zbiorowości może należeć tylko do jednej klasy, czyli podział na nie jest rozłączny, oraz wszystkie jednostki statystyczne są skategoryzowane, zatem klasyfikacja na nie jest zupełna.

Dane przedstawione w skali nominalnej możemy zliczać według kategorii i podać liczebność każdej z nich lub wskazać kategorię najliczniejszą albo najmniej liczną. Nie można ich porównywać, odejmować, dzielić, porządkować (np. rosnąco), ustalać ich rangi („rangować”), a jedyną relacją, jaką można im przyporządkować, jest relacja równości. Są im dedykowane specjalne metody statystyczne, takie jak np. dominanta.

Skalę porządkową również wykorzystuje się w badaniach jakościowych, ale ma ona inne podstawy, gdyż określa cechy, które można uporządkować, takie jak: