

AUTOR

mgr Edyta Kulawiecka

e.kulawiecka@poczta.fm

Państwowa Wyższa Szkoła Zawodowa im. Prezydenta

Stanisława Wojciechowskiego w Kaliszu, Wydział Zarządzania

RACHUNEK KORELACJI W NAUKACH O BEZPIECZEŃSTWIE Z WYKORZYSTANIEM PROGRAMU STATISTICA

Słowa kluczowe: bezpieczeństwo, zmienne ilościowe, zmienne jakościowe, rachunek korelacji, Statistica

Wstęp

We współczesnym świecie zaobserwować można zjawisko powstawania coraz większej ilości danych w różnych obszarach ludzkiej aktywności, w tym również dotyczących bezpieczeństwa. Dlatego też niezbędne są narzędzia pozwalające na ich analizę m.in. statystyczną. Analizę taką można przeprowadzić z wykorzystaniem na przykład programu *Statistica*.

Celem niniejszego opracowania jest przedstawienie istoty rachunku korelacji, zwłaszcza w naukach o bezpieczeństwie, z wykorzystaniem wymienionego programu. Na przykładzie zostaną pokazane zarówno funkcjonalności, jak i wysoka przydatność programu *Statistica* w analizie statystycznej.

Na podstawie danych zawartych w raporcie *Stan bezpieczeństwa ruchu drogowego oraz działania realizowane w tym zakresie w 2014 roku*¹ opracowanym przez Krajową Radę Bezpieczeństwa Ruchu Drogowego pokazane zostanie wykorzystanie programu *Statistica* do prowadzenia analiz, które umożliwią uzyskanie określonych informacji.

Bezpieczeństwo jako dyscyplina naukowa

Nauki o bezpieczeństwie to jedna z dyscyplin naukowych znajdująca się w dziedzinie nauk społecznych, która z kolei mieści się w obszarze wiedzy nauki społeczne². Pojęcie bezpieczeństwo – *securitas* pochodzące z języka łacińskiego składa się z dwóch członów: *sine* (bez) i *cura* (zmartwienie, strach, obawa). W literaturze przedmiotu jest ono interpretowane na wiele sposobów.

W *Słowniku terminów z zakresu bezpieczeństwa narodowego*³ bezpieczeństwo określone jest jako *stan, który daje poczucie pewności, i gwarancje jego zachowania oraz szansę na doskonalenie. Jedną z podstawowych potrzeb człowieka, sytuacja odznaczająca się brakiem ryzyka utraty czegoś, co człowiek szczególnie ceni, na przykład: zdrowia, pracy, szacunku, uczuć, dóbr materialnych*. Z podanej definicji należy wnioskować, że bezpieczeństwo jest najwyższą wartością w życiu człowieka oraz społeczności. O zaliczeniu bezpieczeństwa do grupy potrzeb podstawowych zdecydował już w latach 70. A. Maslow. Ten psycholog kliniczny uporządkował potrzeby ludzkie w sposób hierarchiczny, zaczynając od potrzeb fizjologicznych czyli jedzenia, picia i schronienia. Następnie wskazał cztery poziomy potrzeb psychologicznych. Należy

1 <http://www.krbrd.gov.pl> [dostęp: 15.05.2016].

2 W miejsce nauk o wojskowości decyzją Ministra Nauki i Szkolnictwa Wyższego z dnia 8 sierpnia 2011 w sprawie obszarów wiedzy, dziedzin nauki i sztuki oraz dyscyplin naukowych i artystycznych wprowadzone zostały dwie nowe dyscypliny: nauki o bezpieczeństwie i nauki o obronności, Dz. U. z 2011 r., nr 179, poz. 1065.

3 J. Kaczmarek, W. Łepkowski, B. Zdrodowski (red.), *Słownik terminów z zakresu bezpieczeństwa narodowego*, WSO, AON, Warszawa 2008, s. 14.

zauważyć, że bezpieczeństwo wg A. Masłowa znajduje się na poziomie niższym w stosunku do potrzeby przynależności, szacunku oraz samorealizacji, co sytuuje je w podstawie piramidy⁴.

W *Słowniku Języka Polskiego* bezpieczeństwo przedstawiono jako *stan niezagrożenia, spokoju*⁵. Natomiast wg *Podręcznego Słownika Języka Polskiego* bezpieczeństwo to *niepodleganie obawie, spokój, spokojność, pewność, że się nic złego nie stanie*⁶. Z przytoczonych definicji wynika, że bezpieczeństwo to tylko pewien stan. Bezpieczeństwa nie należy jednak pojmować statycznie, ale przede wszystkim dynamicznie. Wobec powyższego jest ono nie tylko stanem a również procesem *odnoszącym się do podmiotu działającego w określonej przestrzeni strukturalnej (przedmiotowej)*⁷. Brunon Hołyst o bezpieczeństwie pisze, że to *nie tylko indywidualne czy zbiorowe zachowania ludzkie, to także kwestia organizacji życia społecznego i jego warunków, by nie powstawały sytuacje zagrożeniowe, czy też prowokujące zagrożenia*⁸ czyli postrzega bezpieczeństwo jako proces.

Bezpieczeństwo może być też rozumiane negatywnie bądź pozytywnie. W rozumieniu negatywnym zwraca się przede wszystkim uwagę na brak zagrożeń, w pozytywnym zaś traktuje się je jako stan pewności, przetrwania, swobód rozwojowych⁹.

Bezpieczeństwo należy postrzegać w dwóch aspektach: wewnętrznym i zewnętrznym oraz w czterech wymiarach: jednostkowym, grupowym, narodowym/państwowym, międzynarodowym. W ujęciu ogólnym bezpieczeństwo wewnętrzne oznacza stabilność i harmonijność danego organizmu bądź systemu, a bezpieczeństwo zewnętrzne oznacza brak zagrożenia ze strony innych podmiotów¹⁰.

Bezpieczeństwo jest kategorią złożoną, w ramach której wyodrębnić można wiele jego rodzajów według różnych kryteriów. Podstawowa klasyfikacja obejmuje podział tej nowej dyscypliny naukowej według kryteriów podmiotowego i przedmiotowego, a w każdym z nich wskazuje kilka jego typów (rys.1). Bezpieczeństwo według pierwszego kryterium dzieli się na narodowe i międzynarodowe. W przypadku kryterium przedmiotowego należy zauważyć, że nie jest to jednak lista zamknięta, jest i będzie uzupełniana o jego nowe rodzaje w związku z powstawaniem coraz to nowych zagrożeń.

Bolesław Balcerowicz klasyfikuje bezpieczeństwo szerzej, oprócz podmiotowego i przedmiotowego kryterium, uwzględnia kryteria przestrzenne, czasu oraz sposobu organizowania. Wyróżnia zatem według kryterium:

- przestrzennego: bezpieczeństwo lokalne, subregionalne, regionalne, ponadregionalne i globalne,
- czasu: stan bezpieczeństwa i procesy bezpieczeństwa,
- sposobu organizowania: unilateralne (izolacjonizm, hegemonizm, niezaangażowanie, neutralność) i multilateralne (sojusze, bezpieczeństwo zbiorowe, bezpieczeństwo kooperacyjne itd.)¹¹.

Jest to najbardziej ogólny, funkcjonujący w literaturze przedmiotu podział bezpieczeństwa.

4 A. Masłow, *Motywacja i osobowość*, Wydawnictwo Naukowe PWN, Warszawa 2006, s. 62-71.

5 L. Drabik, A. Kubiak-Sokół, E. Sobol, *Słownik języka polskiego*, Wydawnictwo Naukowe PWN, Warszawa 2008, s. 49.

6 Z. de Bondy-Lempicka, S. Arct, *Podręczny Słownik Języka Polskiego*, Wydawnictwo Wiedza Powszechna, Warszawa 1957, s. 12.

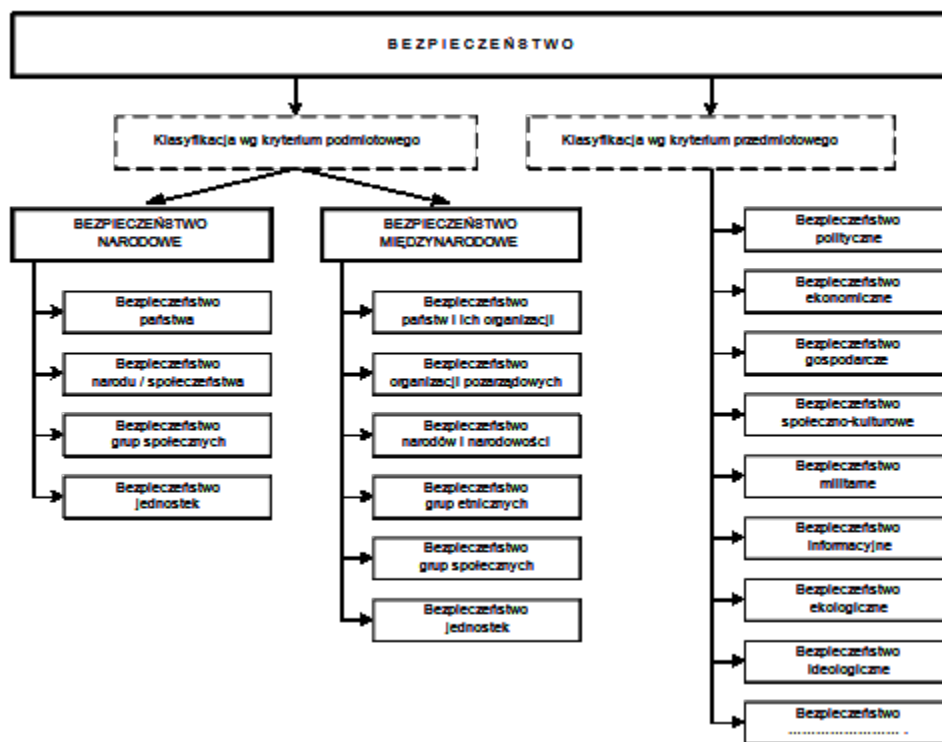
7 K. Drabik, *Paradygmat podmiotowości w naukach o bezpieczeństwie*, [w:] A. Filipek (red.), *Elementy teorii i praktyki transdyscyplinarnych badań problemów bezpieczeństwa*, tom IV, Uniwersytet Przyrodniczo-Humanistyczny w Siedlcach, Siedlce 2015, s. 30.

8 B. Hołyst, *Kryminologia*, Wydawnictwo LexisNexis, Warszawa 2007, s. 1473.

9 M. Huzarski, *Zmienne podstawy bezpieczeństwa i obronności państwa*, AON, Warszawa 2009, s. 12.

10 Tamże, s. 12.

11 B. Balcerowicz, *Bezpieczeństwo polityczne Rzeczypospolitej Polskiej*, ZUMS BN, WSO, AON, Warszawa 2004, s. 7.



Źródło: opracowanie własne na podstawie: J. Gryz, *Zarys podstaw teorii bezpieczeństwa*, AON, Warszawa 2010, s. 74.

Rys. 1. Klasyfikacja bezpieczeństwa

Bezpieczeństwo tradycyjnie łączone jest z kwestiami porządku publicznego, obronności czy zwalczania zagrożeń. O bezpieczeństwie mówi się również w innych aspektach, na przykład ruchu drogowego, energetycznym, pracy. Zatem jest ono domeną wielu nauk: społecznych, prawnych, humanistycznych i ekonomicznych.

Należy pamiętać, że nauki o bezpieczeństwie to bardzo młoda dyscyplina, której obiekt i przedmioty badań istniały jednak już znacznie wcześniej. W związku z powyższym obiektem badań w naukach o bezpieczeństwie są: zbiorowości i zbory społeczne, instytucje społeczne, procesy i zjawiska społeczne¹². Za przedmiot badań nauk o bezpieczeństwie Bogdan Szulc uznaje *współczesne potrzeby, stany i systemy we wszelkich wymiarach oraz ich funkcjonowanie w aspekcie efektywności jednostki oraz na różnych poziomach organizacyjnych*¹³. Współczesne systemy bezpieczeństwa obejmują zatem działania organizacji o charakterze państwowym, rządowym, samorządowym, gospodarczym oraz społecznym w wymiarze militarny i niemilitarny. Według Stanisława Zajasa nauki o bezpieczeństwie powinny jednak obejmować przede wszystkim całość problematyki bezpieczeństwa w wymiarze pozamilitarnym. Przedmiotem badań tej nowej dyscypliny nauk społecznych powinny być zatem współczesne systemy bezpieczeństwa oraz ich działanie na różnych poziomach organizacyjnych, w których biorą udział różne instytucje administracji publicznej¹⁴. Bardzo obszernie przedmiot badań tej nowej dyscypliny nauk społecznych sformułował Marian Cieślarczyk, który przyjął, że jest nim *fenomen bezpieczeństwa oraz składające się na niego fakty (zdarzenia), procesy, i bardziej szczegółowe zjawiska w sferze bezpieczeństwa, w jego różnych wymiarach przedmiotowych i między nimi, rozpatrywane w odniesieniu do konkretnych podmiotów, z punktu widzenia których analizujemy bezpieczeństwo, przy uwzględnieniu także wpływu warunków środowiskowych*. Ponadto, zdaniem autora, *oczywiste*

12 B. Wiśniewski (red.), *Bezpieczeństwo w teorii i badaniach naukowych*, Wydawnictwo Wyższej Szkoły Policji, Szczytno 2011, s. 118.

13 M. Huzarski, J. Wołejko (red.), *Leksykon obronności, Polska i Europa*, Wydawnictwo Bellona, Warszawa 2014, s. 14.

14 S. Zajas, *Nauki wojskowe a nauki o bezpieczeństwie i obronności*, [w:] P. Sienkiewicz, *Metodologia badań bezpieczeństwa narodowego*, materiały z konferencji naukowej, Warszawa 2010, s. 105.

jest, że przedmiotem badań będą także, a może przede wszystkim, relacje między podmiotem a jego otoczeniem (środowiskiem), charakter tego otoczenia, a także cechy danego podmiotu¹⁵.

Analiza podanych definicji i kryteriów podziału bezpieczeństwa oraz jego podmiotu i przedmiotu badań daje podstawę do wyciągnięcia wniosku o bardzo szerokim znaczeniu tego terminu. Nie jest zadaniem łatwym sformułowanie uniwersalnej, zwartej i precyzyjnej definicji bezpieczeństwa. Termin ten ma bowiem wiele aspektów, a jego charakter sprawia, że kryteria pozwalające zdefiniować pojęcie bezpieczeństwa dotyczą między innymi wojskowości, ekonomii, socjologii, prawa, psychologii. Bezpieczeństwo jest więc pojęciem interdyscyplinarnym.

Podstawy rachunku korelacji

Rachunek korelacji jako jedna z najbardziej popularnych miar statystycznych znalazł szerokie zastosowanie w wielu dziedzinach nauki. Z powodzeniem miarę tę można wykorzystać w naukach o bezpieczeństwie.

Korelacja (łac. *correlatio*) ogólnie jest to współzależność, wzajemne powiązanie; w rachunku prawdopodobieństwa i statystyce matematycznej związek między zmiennymi przypadkowymi¹⁶. Najprostszym przykładem korelacji jest związek dwóch zjawisk. Jego siłę można wyrazić liczbowo za pomocą wielu mierników. Wybór miernika zaś uzależniony jest między innymi od rodzaju cech (zmiennych), między którymi zależność jest badana (mieralne, niemieralne bądź mieszane), liczby obserwacji, kształtu zależności¹⁷. Należy pamiętać, że bardzo często na wyróżnione zjawisko oddziałuje wiele innych zjawisk. W takich przypadkach mówi się o korelacji wielorakiej (wielu zmiennych).

Badanie współzależności cech¹⁸ polega na wykrywaniu i ustalaniu ilościowych zależności między pewną badaną cechą zaobserwowaną w zbiorowości statystycznej zwaną zmienną zależną (Y), a innymi cechami tej zbiorowości zwanymi zmiennymi niezależnymi (X)¹⁹. Celem badań współzależności jest zatem stwierdzenie, czy między badanymi zmiennymi zachodzi zależność, jeżeli tak, to jaka jest jej siła, rodzaj i kierunek oraz w jakiej postaci matematycznej można ją zapisać²⁰.

Wyróżnia się dwa rodzaje zależności, jakie można zaobserwować między cechami opisującymi daną zbiorowość: funkcyjną i stochastyczną.

Zależność funkcyjna polega na tym, że zmiana wartości jednej zmiennej niezależnej powoduje ściśle określone zmiany wartości drugiej zmiennej zależnej.

Zależność stochastyczna występuje wtedy, gdy zmiany wartości zmiennej niezależnej powodują zmianę rozkładu prawdopodobieństwa zmiennej zależnej, co oznacza, że przyporządkowanie to ma charakter losowy²¹. Szczególnym przypadkiem zależności

15 M. Cieślarczyk, *Teoretyczne i metodologiczne podstawy badania problemów bezpieczeństwa i obronności państwa*, Siedlce 2009, s. 212.

16 J. Apanowicz, *Zarys metodologii prac dyplomowych z organizacji i zarządzania*, Wydawnictwo Wyższego Seminarium Duchownego Bernardinum, Gdynia 1997, s. 24.

17 M. Sobczyk, *Statystyka*, Wydawnictwo Naukowe PWN, Warszawa 2000, s. 218.

18 Dział statystyki zajmującej się badaniem związków zachodzących między zmiennymi nosi nazwę teorii współzależności.

19 A. Bielecka, *Statystyka w biznesie i ekonomii, teoria i praktyka*, WSPiZ im. Leona Koźmińskiego, Warszawa 2005, s. 254.

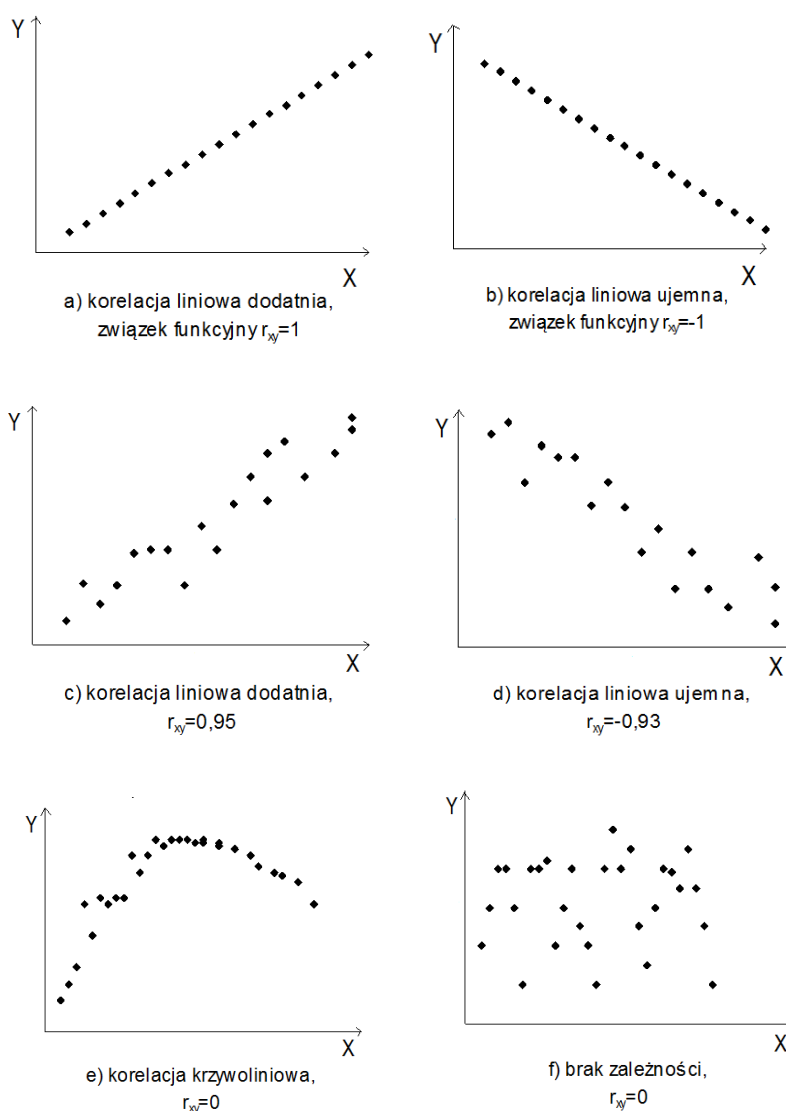
20 Badaniem siły, kierunku i rodzaju zależności między zmiennymi zajmuje się rachunek korelacji, zaś ustalaniem kształtu zależności tzw. analiza regresji.

21 A. Bielecka, *Statystyka...*, s. 255.

stochastycznej jest zależność korelacyjna. Polega ona na tym, że określonym wartościom jednej zmiennej odpowiadają ściśle określone średnie wartości drugiej zmiennej²².

Badanie współzależności cech (zmiennych) przeprowadzić można tylko, gdy pomiędzy cechami zachodzi związek logiczny, czyli istnieje więź przyczynowo-skutkowa. Dlatego też w pierwszej kolejności należy przeprowadzić analizę jakościową, czyli ustalić, czy istnieje wspomniany logiczny związek między badanymi cechami. Dopiero wówczas gdy prawdopodobnie takie relacje występują, przechodzi się do analizy ilościowej tych współzależności.

Analizę zależności rozpocząć również należy od sporządzenia korelacyjnego diagramu rozrzutu, który jest zbiorem punktów na płaszczyźnie, odpowiadających zbiorowi par liczb (x_i, y_i) , gdzie x_i jest i -tą obserwacją zmiennej niezależnej X, natomiast y_i i -tą obserwacją zmiennej zależnej Y. Wzrokowa ocena korelacyjnego diagramu rozrzutu umożliwi wstępne określenie siły i rodzaju zależności. Na rys. 2 zaprezentowano różne sytuacje związane z układem danych wraz z odpowiadającymi im wartościami współczynnika korelacji liniowej Pearsona.



Źródło: opracowanie własne na podstawie: U. Siedlecka, *Analiza korelacji i regresji*, [w:] S. Ostasiewicz, Z. Rusnak, U. Siedlecka, *Statystyka. Elementy teorii i zadania*, Wyd. AE im. Oskara Langego we Wrocławiu, Wrocław 2001, s. 310.

Rys. 2. Przykłady korelacyjnych diagramów rozrzutu

Korelacyjny diagram rozrzutu jest często wykorzystywanym w praktyce sposobem stwierdzenia czy istnieje korelacja między dwiema zmiennymi mierzalnymi. W przypadku gdy

22 U. Siedlecka, *Analiza korelacji i regresji*, [w:] S. Ostasiewicz, Z. Rusnak, U. Siedlecka, *Statystyka, Elementy teorii i zadania*, AE im. Oskara Langego we Wrocławiu, Wrocław 2001, s. 309.

między dwiema badanymi zmiennymi ilościowymi zachodzi związek liniowy²³, wówczas miarą siły i kierunku korelacji między nimi jest współczynnik korelacji liniowej Pearsona r_{xy} wyznaczany w drodze standaryzacji kowariancji. Kowariancja jest średnią arytmetyczną iloczynu odchyleń poszczególnych zmiennych od ich średnich arytmetycznych, czyli²⁴:

$$\text{cov}(x, y) = \text{cov}(y, x) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \quad [1]$$

Kowariancja przekazuje następujące informacje o związku korelacyjnym:

$\text{cov}(x, y) = 0$ - brak zależności korelacyjnej,

$\text{cov}(x, y) < 0$ - ujemna zależność korelacyjna,

$\text{cov}(x, y) > 0$ - dodatnia zależność korelacyjna.

Standaryzując kowariancję, otrzyma się unormowany miernik siły i kierunku współzależności liniowej dwóch cech mierzalnych. Ostatecznie współczynnik korelacji Pearsona przyjmuje postać²⁵:

$$r_{xy} = r_{yx} = \frac{\text{cov}(x, y)}{S(x) \cdot S(y)} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n \cdot S(x) \cdot S(y)} \quad [2]$$

gdzie: $S(x), S(y)$ to odchylenia standardowe zmiennych: niezależnej X i zależnej Y .

W analizie korelacji badacz jednakowo traktuje obie zmienne: zależną i niezależną. Korelacja między zmiennymi jest pewną miarą liniowości związku między nimi. Pokazuje ona, na ile obie zmienne zmieniają się równocześnie w sposób liniowy. Korelacja między X i Y jest więc taka sama jak między Y i X ²⁶. Dlatego też wartość obliczonego współczynnika korelacji liniowej Pearsona spełnia równość $r_{xy} = r_{yx}$.

Współczynnik korelacji liniowej Pearsona informuje o kierunku zależności, określa to znak współczynnika. Bezwzględna jego wartość mówi zaś o sile związku. Dodatni znak współczynnika wskazuje na istnienie współzależności pozytywnej (dodatniej), ujemny oznacza współzależność negatywną (ujemną). Przyjmuje on wartości z przedziału od -1 do $+1$ ²⁷. Jeżeli współczynnik r_{xy} osiągnie wartość:

$r_{xy} = 1$, oznacza ona korelację dodatnią, związek funkcyjny,

$r_{xy} = -1$, oznacza ona korelację ujemną, związek funkcyjny,

$r_{xy} = 0$, oznacza brak korelacji liniowej, czyli nie ma liniowego związku między dwiema zmiennymi;

$-1 < r_{xy} < 0$, oznacza korelację ujemną niedoskonałą;

$0 < r_{xy} < 1$, oznacza korelację dodatnią niedoskonałą.

W analizach statystycznych zwykle przyjmuje się też, że jeżeli współczynnik korelacji wynosi²⁸:

- mniej niż $|0,2|$ to praktycznie nie ma związku liniowego między badanymi cechami;
- $|0,2-0,4|$ – zależność liniowa wyraźna, lecz niska;
- $|0,4-0,7|$ – zależność umiarkowana;
- $|0,7-0,9|$ – zależność znacząca;
- powyżej $|0,9|$ – zależność bardzo silna.

23 Związek liniowy wystąpi wtedy, gdy jednostkowym przyrostom zmiennej niezależnej towarzyszyć będzie, średnio, stały przyrost zmiennej zależnej.

24 M. Sobczyk, *Statystyka, podstawy teoretyczne, przykłady zadania*, Uniwersytetu Marii Curie-Skłodowskiej, Lublin 2000, s. 240.

25 Tamże, s. 241.

26 A. Aczel, *Statystyka w zarządzaniu*, Wydawnictwo Naukowe PWN, Warszawa 2000, s. 479.

27 M. Sobczyk, *Statystyka...*, s. 241.

28 U. Siedlecka, *Analiza...*, s. 311.

Należy zwrócić uwagę, że w literaturze przedmiotu przedziały stanowiące podstawę ustalania siły korelacji nieznacznie różnią się²⁹. Przy interpretacji współczynnika korelacji liniowej Pearsona należy pamiętać o trzech rzeczach:

- najważniejsza jest istotność korelacji, ponieważ badaczowi nie potrzebna jest korelacja nawet bardzo wysoka, jeżeli nie jest istotna;
- wartość współczynnika bliska 0 nie zawsze oznacza brak zależności, a jedynie brak zależności liniowej;

wielkość współczynnika podlega wpływom wartości skrajnych i odstających³⁰. Istotność korelacji związana jest z tak zwanym poziomem istotności p, który określa prawdopodobieństwo popełnienia błędu przy przyjęciu hipotezy, że pomiędzy badanymi zmiennymi występuje korelacja. W naukach humanistycznych przyjmuje się, że współczynnik korelacji jest istotny, jeśli $p \leq 0,05$ lub przy bardziej rygorystycznych wymaganiach $p \leq 0,01$ ³¹. W badaniach naukowych badacz ma nie tylko do czynienia ze zmiennymi ilościowymi, lecz także ze zmiennymi o charakterze jakościowym. W związku z powyższym analiza zależności cech statystycznych może dotyczyć zarówno cech ilościowych, jak i jakościowych. Chcąc dokonać oceny związku pomiędzy zmiennymi jakościowymi, należy przedstawić zebrane dane indywidualne w postaci tablicy wielodzzielczej inaczej kontyngencji³². W tablicy takiej o określonej liczbie wierszy i kolumn zamieszcza się liczebności poszczególnych wariantów cech (a nie ich wartości). Liczba kolumn i wierszy, a także specyfikacja kategorii cech, zależą od badacza. Najprostszą tablicą wielodzzielczą jest tablica asocjacji 2 x 2, tj. o dwóch wierszach i dwóch kolumnach³³. Tablice wielodzzielcze stanowią podstawę do obliczania miar określających siłę związku. Miary korelacji zmiennych jakościowych nie pozwalają bowiem na określenie kierunku zależności, wskazują tylko jej siłę. W literaturze przedmiotu spotkać można wiele miar służących pomiarowi zależności między cechami niemierzalnymi.

Do najczęściej stosowanych należą³⁴:

- współczynnik F Yule'a,
$$\Phi = \sqrt{\frac{\chi^2}{N}} \in \langle 0, 1 \rangle$$
 ; [3]
- współczynnik zbieżności V–Cramera,
$$V = \sqrt{\frac{\chi^2}{N \times \min(w-1; k-1)}} \in \langle 0, 1 \rangle$$
 ; [4]
- współczynnik kontyngencji Pearsona
$$C = \sqrt{\frac{\chi^2}{\chi^2 + N}} \in \langle 0, \frac{\sqrt{\frac{w-1}{w}} + \sqrt{\frac{k-1}{k}}}{2} \rangle$$
 ;

29 Por., $0 < r_{xy} < 0,1$ korelacja nikła; $0,1 \leq r_{xy} < 0,3$ korelacja słaba; $0,3 \leq r_{xy} < 0,5$ korelacja przeciętna; $0,5 \leq r_{xy} < 0,7$ korelacja wysoka; $0,7 \leq r_{xy} < 0,9$ korelacja bardzo wysoka; $0,9 \leq r_{xy} < 1$ korelacja prawie pełna, [w:] A. Stanisławski, *Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny*, StatSoft, tom 1, Kraków 2006, s. 293; $r_{xy} \leq 0,3$ korelacja niewyraźna; $0,3 < r_{xy} \leq 0,5$ korelacja średnia; $r_{xy} > 0,5$ korelacja niewyraźna, [w:] M. Sobczyk, *Statystyka...*, s. 225;

[5]

gdzie w - liczba wierszy, k - liczba kolumn w tablicy wielodzielczej.

Interpretacja wszystkich wskaźników jest taka sama:

- jeżeli ma on wartość zero, to cechy X i Y są niezależne;
im bliższa jedynki jest wartość tych współczynników, tym silniejsze powiązanie pomiędzy analizowanymi cechami X i Y ³⁵. Wszystkie te miary bazują na wartości pewnej statystyki χ^2 (chi-kwadrat) obliczonej zgodnie z wzorem:

$$\chi^2 = \sum_{i=1}^w \sum_{j=1}^k \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}, \quad [6]$$

gdzie:

n_{ij} - liczebności empiryczne,

\hat{n}_{ij} - liczebności teoretyczne, będące liczbą elementów, które powinny wystąpić w i -tym wierszu oraz j -tej kolumnie, gdyby między zmiennymi X i Y nie występowała żadna korelacja.

Wartość \hat{n}_{ij} oblicza się z wzoru:

$$\hat{n}_{ij} = \frac{\sum_{i=1}^w n_{ij} \cdot \sum_{j=1}^k n_{ij}}{n} \quad [7]$$

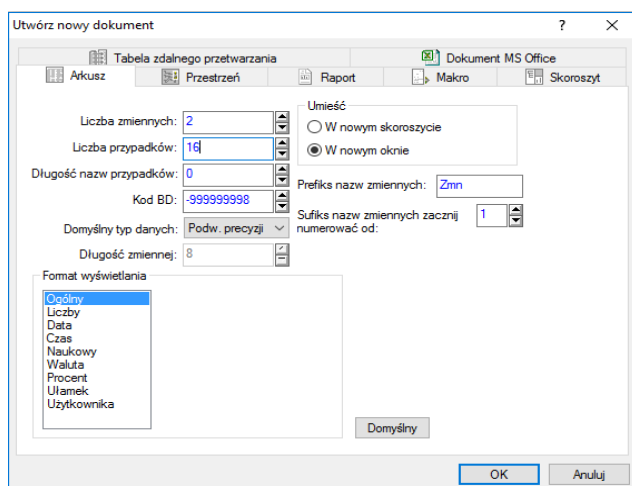
Statystyka chi-kwadrat przyjmuje wartości z przedziału

$(0, n \cdot \sqrt{(w-1)(k-1)})$ ³⁶.

Podsumowując, podstawowym celem każdego badania naukowego winno być poszukiwanie zależności między zmiennymi. Wg autora nie ma innego sposobu wyrażania relacji między zmiennymi ilościowymi bądź jakościowymi.

Przykład zastosowania pakietu Statistica w rachunku korelacji

Statistica jest uniwersalnym, zintegrowanym systemem służącym do statystycznej analizy danych, tworzenia wykresów, operowania na bazach danych, wykonywania transformacji danych i tworzenia aplikacji. W skład systemu wchodzi wszechstronny zestaw zaawansowanych procedur analitycznych i data mining. Statistica zawiera nie tylko procedury statystyczne i graficzne ogólnego przeznaczenia i związane z nimi narzędzia zarządzania danymi, ale także specjalistyczne techniki analityczne wykorzystywane m.in. w marketingu, finansach, zarządzaniu produkcją, badaniach naukowych i wielu innych³⁷. Oprogramowanie to pozwala wyznaczyć parametry klasyczne oraz pozycyjne, przeprowadzić rachunek korelacji i regresji – prostej i wielorakiej, wykonać analizy wielowymiarowe: kanoniczną, dyskryminacyjną, czynnikową, log-liniową i wiele innych. Przeprowadzenie analizy przy użyciu pakietu Statistica wymaga przygotowania arkusza o odpowiedniej liczbie zmiennych i przypadków (rys. 3). Dane w arkuszu można edytować analogicznie jak w arkuszach kalkulacyjnych. Za pomocą programu Statistica można zatem przeprowadzić analizę korelacji zarówno pomiędzy zmiennymi ilościowymi jak i jakościowymi. Narzędzie to z powodzeniem może być wykorzystywane w naukach o bezpieczeństwie.



Źródło: opracowanie własne z wykorzystaniem programu Statistica.

Rys. 3. Okno wejściowe programu STATISTICA (ver. 12)

Jednym z elementów bezpieczeństwa w ogólnej działalności człowieka jest zagadnienie bezpieczeństwa ruchu drogowego. Zaś ważnym jego aspektem ustalenie, jakie czynniki wpływają na liczbę wypadków drogowych³⁸ na danym obszarze oraz czy pomiędzy rodzajem zderzenia a obszarem, na którym wydarzył się wypadek drogowy, istnieje skorelowanie. Przeprowadzona analiza dała odpowiedź na pytanie, która z określonych zmiennych objaśniających jest najsilniej skorelowana ze zmienną objaśnianą, czyli liczbą wypadków drogowych zaobserwowaną w poszczególnych województwach w Polsce w 2014 roku. W pracy posłużono się danymi opublikowanymi przez Krajową Radę Bezpieczeństwa Ruchu Drogowego w dokumencie *Stan bezpieczeństwa ruchu drogowego oraz działania realizowane w tym zakresie w 2014 r.* Raport ten zawiera między innymi informacje na temat:

- bezpieczeństwa ruchu drogowego w Polsce;
- bezpieczeństwa ruchu drogowego w województwach;
- kosztów wypadków drogowych;
- zachowań uczestników ruchu drogowego.

Przeprowadzając rachunek korelacji dwóch cech ilościowych, na początku należy zdefiniować zmienną niezależną, zmienną zależną oraz ustalić, czy zachodzi pomiędzy z nimi związek

przyczynowo-skutkowy. Informacje zawarte w raporcie pozwoliły przyjąć na potrzeby analizy za zmienne niezależne:

- liczbę ludności w danym województwie (X1);
- długość dróg publicznych w danym województwie w km (X2);
- liczbę zarejestrowanych samochodów w danym województwie w tys. (X3);
- gęstość zaludnienia w danym województwie mierzona liczbą osób/km² (X4);
- wskaźnik I mierzony liczbą zarejestrowanych samochodów na kilometr drogi w danym województwie (X5);
- wskaźnik II mierzony liczbą zarejestrowanych samochodów na mieszkańca w danym województwie (X6).

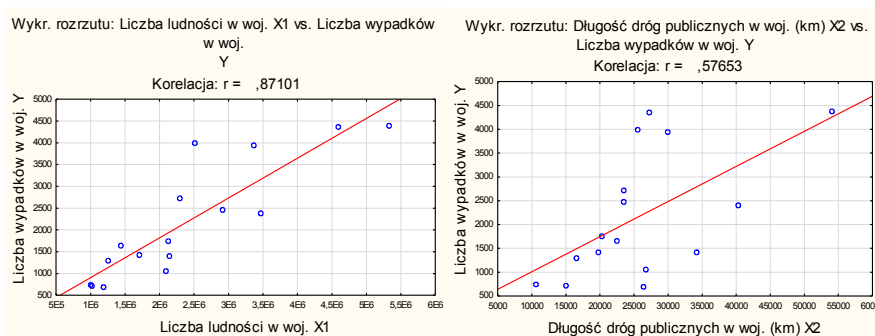
Rys. 4. zawiera dane wejściowe niezbędne do przeprowadzenia rachunku korelacji cech ilościowych.

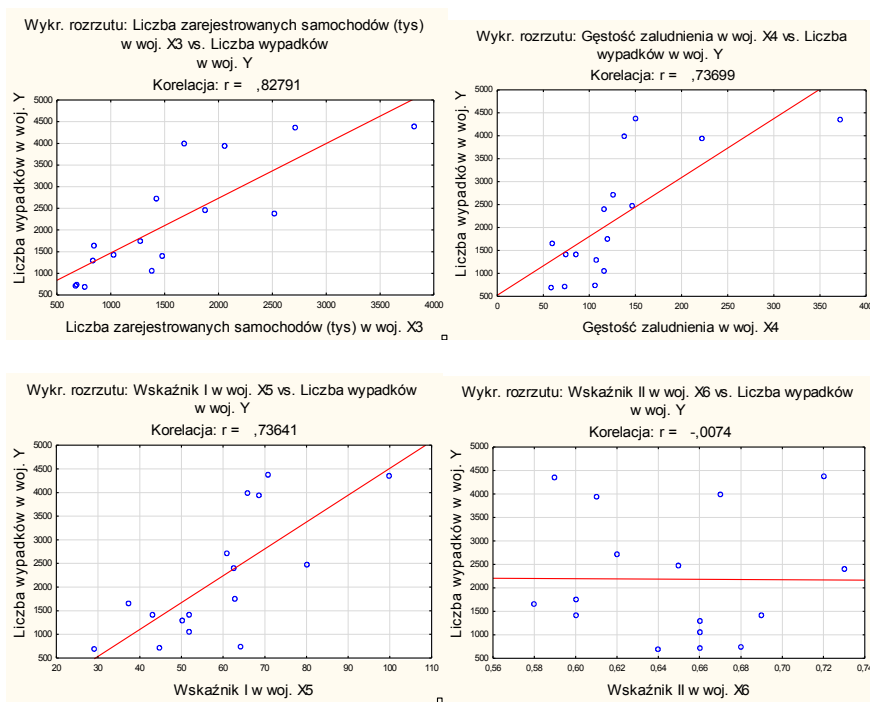
	1 Liczba ludności w woj. X1	2 Długość dróg publicznych w woj. (km) X2	3 Liczba zarejestrowanych samochodów (tys.) w woj. X3	4 Gęstość zaludnienia w woj. X4	5 Wskaźnik I w woj. X5	6 Wskaźnik II w woj. X6	7 Liczba wypadków w woj. Y
1	2 504 136	25 633	1 686,80	137,45	65,81	0,67	39
2	5 334 511	54 045	3 819,50	150,02	70,67	0,72	43
3	3 368 336	29 960	2 056,30	221,85	68,63	0,61	39
4	4 585 924	27 189	2 716,00	371,84	99,89	0,59	43
5	2 147 746	34 262	1 475,30	85,49	43,06	0,69	14
6	2 129 187	20 222	1 272,20	119,31	62,91	0,60	17
7	1 191 918	26 328	762,50	59,04	28,96	0,64	6
8	1 263 176	16 583	834,30	107,86	50,31	0,66	13
9	1 020 307	14 975	668,60	72,94	44,65	0,66	7
10	3 472 579	40 288	2 523,40	116,43	62,63	0,73	23
11	1 715 431	19 699	1 023,30	74,94	51,95	0,60	14
12	2 908 457	23 430	1 879,30	145,81	80,21	0,65	24
13	1 000 858	10 613	682,10	106,34	64,27	0,68	7
14	2 089 992	26 752	1 385,20	116,29	51,78	0,66	10
15	2 302 077	23 439	1 427,30	125,73	60,89	0,62	27
16	1 443 967	22 553	842,00	59,73	37,33	0,58	16

Źródło: opracowanie własne na podstawie danych zawartych w dokumencie *Stan bezpieczeństwa ruchu drogowego oraz działania realizowane w tym zakresie w 2014 r.* opracowanym przez Krajową Radę Bezpieczeństwa Ruchu Drogowego.

Rys. 4. Dane wejściowe w arkuszu Statistica

Analizę zależności rozpoczęto od wykreślenia korelacyjnych diagramów rozrzutu (rys. 5). Celem ich sporządzenia była ocena, czy zachodzące zależności mogą mieć charakter liniowy.



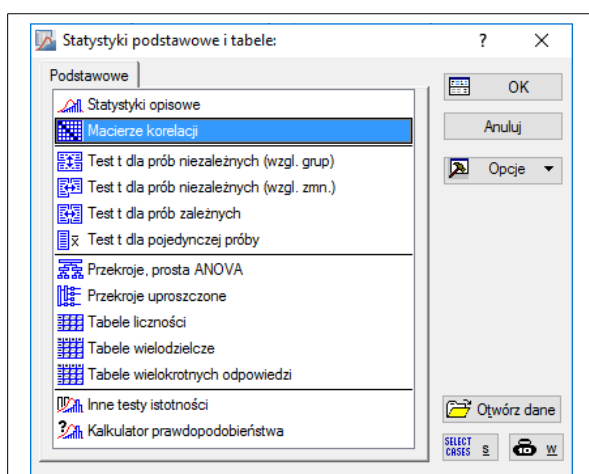


Źródło: opracowanie własne na podstawie danych zawartych w dokumencie *Stan bezpieczeństwa ruchu drogowego oraz działania realizowane w tym zakresie w 2014 r.* opracowanym przez Krajową Radę Bezpieczeństwa Ruchu Drogowego.

Rys. 5. Korelacyjne diagramy rozrzutu

Do ustalenia siły i kierunku zależności wykorzystano powszechnie stosowany współczynnik korelacji liniowej Pearsona r_{xy} . W programie *Statistica* narzędzie to dostępne jest w opcji Macierze korelacji (rys. 6). Szczegółowa ścieżka obejmuje następujące kroki:

Statystyka Statystyki podstawowe i tabele Macierze korelacji OK. dwie listy zmiennych wprowadzamy zmienne OK. Podsumowanie



Źródło: opracowanie własne z wykorzystaniem programu Statistica.

Rys. 6. Macierze korelacji w programie Statistica

Wyniki obliczeń (rys. 7) wskazują na istnienie znaczącej, dodatniej zależności liniowej pomiędzy liczbą wypadków drogowych w województwach w 2014 roku a liczbą ludności zamieszkałą na ich terenie. Wskazuje na to wartość współczynnika korelacji równa 0,87108.

		Korelacje Oznaczone wsp. korelacji są istotne z $p < ,05000$ N=16
Zmienna		Liczba wypadków w woj. zmienna Y
X1 - Liczba ludności w woj.		0,871008
X2 - Długość dróg publicznych w woj. (km)		0,576527
X3 - Liczba zarejestrowanych samochodów w woj. (tys.)		0,827912
X4 - Gęstość zaludnienia w woj.		0,736993
X5 - Wskaźnik I w woj.		0,736414
X6 - Wskaźnik II w woj.		0,001119

Źródło: opracowanie własne z wykorzystaniem programu Statistica.

Rys. 7. Wyniki korelacji zmiennych ilościowych

Wyniki przedstawione na rysunku 8 pozwalają rozpoznać zmienne objaśniające istotnie wpływające na liczbę wypadków drogowych. Program zmienne istotne statystycznie zaznacza bowiem kolorem czerwonym.

Równie mocno na badaną zmienną zależną wpływa:

- liczba zarejestrowanych samochodów w danym województwie w tys. ($r_{xy}=0,827912$);
- gęstość zaludnienia w danym województwie mierzona liczbą osób/km² ($r_{xy}=0,736993$);
- wskaźnik I mierzony liczbą zarejestrowanych samochodów na kilometr drogi w danym województwie ($r_{xy}=0,736414$).

Umiarkowana, dodatnia korelacja występuje pomiędzy badaną zmienną zależną a długością dróg publicznych w danym województwie w km ($r_{xy}=0,576527$). Jedyna zmienna objaśniająca, która nie jest skorelowana ze zmienną objaśnianą, to wskaźnik II mierzony liczbą zarejestrowanych samochodów na mieszkańca w danym województwie; zmienna ta nie jest statystycznie istotnie ($p>0,05$). Przy interpretacji obliczonego współczynnika korelacji liniowej Pearsona należy jednak pamiętać, że wartość zbliżająca się do zera nie musi oznaczać braku zależności, a jedynie brak zależności liniowej.

		Korelacje Oznaczone wsp. korelacji są istotne z $p < ,05000$ N=16
Zmienna		Liczba wypadków w woj. zmienna Y
X1 - Liczba ludności w woj.		,8710 p=,000
X2 - Długość dróg publicznych w woj. (km)		,5765 p=,019
X3 - Liczba zarejestrowanych samochodów w woj. (tys.)		,8279 p=,000
X4 - Gęstość zaludnienia w woj.		,7370 p=,001
X5 - Wskaźnik I w woj.		,7364 p=,001
X6 - Wskaźnik II w woj.		,0011 p=,997

Źródło: opracowanie własne z wykorzystaniem programu Statistica.

Rys. 8. Wyniki istotności korelacji zmiennych ilościowych

Rachunek korelacji dwóch zmiennych jakościowych również przeprowadzono z wykorzystaniem danych pochodzących z wspomnianego już w artykule raportu *Stan bezpieczeństwa ruchu drogowego oraz działania realizowane w tym zakresie w 2014 r.* Analizą objęto 34970 wypadków drogowych w Polsce w badanym roku. Dane wprowadzono do arkusza roboczego o liczbie zmiennych dwie i liczbie przypadków 34970. Rysunek 9 przedstawia tylko wycinek tego arkusza. Następnie przeprowadzono analizę, która ma dać odpowiedź na pytanie:

czy pomiędzy rodzajem zderzenia a obszarem, na którym wydarzył się wypadek drogowy, istnieje skorelowanie statystycznie istotne, jeśli tak, to jaka jest jego siła.

Punktem wyjścia do przeprowadzenia każdej analizy korelacji zmiennych jakościowych jest zestawienie danych dotyczących badanych cech w formie tablicy wielodzieldzej oraz wyznaczenie statystyki c^2 (chi-kwadrat).

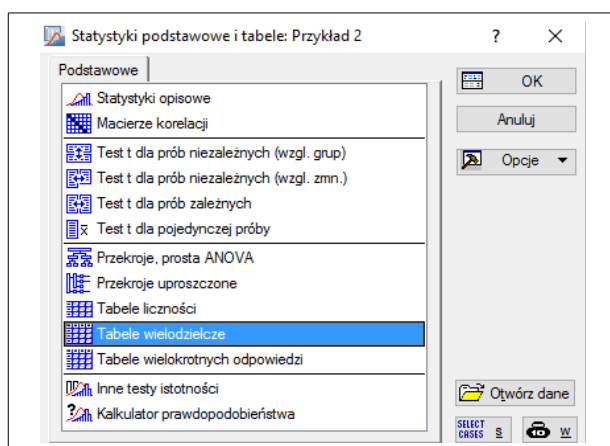
	1 Obszar	2 Rodzaj zdarzenia
1	ZABUDOWANY	ZDERZENIE POJAZDÓW CZOŁOWE
2	ZABUDOWANY	ZDERZENIE POJAZDÓW CZOŁOWE
3	ZABUDOWANY	ZDERZENIE POJAZDÓW CZOŁOWE
4	ZABUDOWANY	ZDERZENIE POJAZDÓW CZOŁOWE
5	ZABUDOWANY	ZDERZENIE POJAZDÓW CZOŁOWE
6	ZABUDOWANY	ZDERZENIE POJAZDÓW CZOŁOWE
7	ZABUDOWANY	ZDERZENIE POJAZDÓW CZOŁOWE
8	ZABUDOWANY	ZDERZENIE POJAZDÓW CZOŁOWE
9	ZABUDOWANY	ZDERZENIE POJAZDÓW CZOŁOWE
10	ZABUDOWANY	ZDERZENIE POJAZDÓW CZOŁOWE
11	ZABUDOWANY	ZDERZENIE POJAZDÓW CZOŁOWE
12	ZABUDOWANY	ZDERZENIE POJAZDÓW CZOŁOWE
13	ZABUDOWANY	ZDERZENIE POJAZDÓW CZOŁOWE
14	ZABUDOWANY	ZDERZENIE POJAZDÓW CZOŁOWE
15	ZABUDOWANY	ZDERZENIE POJAZDÓW CZOŁOWE
16	ZABUDOWANY	ZDERZENIE POJAZDÓW CZOŁOWE
17	ZABUDOWANY	ZDERZENIE POJAZDÓW CZOŁOWE
18	ZABUDOWANY	ZDERZENIE POJAZDÓW CZOŁOWE
19	ZABUDOWANY	ZDERZENIE POJAZDÓW CZOŁOWE
20	ZABUDOWANY	ZDERZENIE POJAZDÓW CZOŁOWE
21	ZABUDOWANY	ZDERZENIE POJAZDÓW CZOŁOWE

Źródło: opracowanie własne z wykorzystaniem programu Statistica.

Rys. 9. Dane wejściowe w arkuszu Statistica

W tym celu skorzystać można z narzędzia dostępnego w opcji Tabele wielodzieldzej programu Statistica (rys. 10). Szczegółowa ścieżka obejmuje następujące kroki:

Statystyka → Statystyki podstawowe i tabele → Tabele wielodzieldzej → OK → Zbiorcze → Określ tabele → wprowadzamy zmienne → OK → OK → Opcje → zaznaczamy Liczności oczekiwane i chi-kwadrat Pearsona → Podsumowanie



Źródło: opracowanie własne z wykorzystaniem programu Statistica.

Rys. 10. Tabele wielodzieldzej w programie Statistica

Rysunki 11 i 12 prezentują statystyki podsumowujące i obejmują: liczności empiryczne, liczności teoretyczne, wartość statystyki c^2 oraz p-wartość.

Tabela liczności (Przykład 2)								
Tabela: Obszar(2) x Rodzaj zdarzenia(7)								
Obszar	Rodzaj zdarzenia ZDERZENIE POJAZDÓW CZOŁOWE	Rodzaj zdarzenia ZDERZENIE POJAZDÓW BOCZNE	Rodzaj zdarzenia ZDERZENIE POJAZDÓW TYLNE	Rodzaj zdarzenia NAJECHANIE NA PIESZEGO	Rodzaj zdarzenia NAJECHANIE NA DRZEWO	Rodzaj zdarzenia WYWRÓCENIE SIĘ POJAZDU	Rodzaj zdarzenia POZOSTAŁE	Wiersz Razem
ZABUDOWANY	1891	8337	2871	8012	536	1283	2438	25368
NIEZABUDOWANY	1340	2174	1347	921	1289	1377	1154	9602
Łącznie	3231	10511	4218	8933	1825	2660	3592	34970

Źródło: opracowanie własne z wykorzystaniem programu Statistica.

Rys. 11. Liczebności empiryczne

Podsumowanie: Wyliczenie licznosci (Przyklad 2)								
Chi ² Pearsona: 4458,19, df=6, p=0,00000								
Obszar	Rodzaj zdarzenia ZDERZENIE POJAZDÓW CZOŁOWE	Rodzaj zdarzenia ZDERZENIE POJAZDÓW BOCZNE	Rodzaj zdarzenia ZDERZENIE POJAZDÓW TYLNE	Rodzaj zdarzenia NAJECHANIE NA PIESZEGO	Rodzaj zdarzenia NAJECHANIE NA DRZEWO	Rodzaj zdarzenia WYWRÓCENIE SIĘ POJAZDU	Rodzaj zdarzenia POZOSTAŁE	Wiersz Razem
ZABUDOWANY	2343,838	7624,91	3059,829	6480,193	1323,895	1929,622	2605,715	25368,00
NIEZABUDOWANY	887,162	2886,09	1158,171	2452,807	501,105	730,378	986,285	9602,00
Ogół	3231,000	10511,00	4218,000	8933,000	1825,000	2660,000	3592,000	34970,00

Źródło: Opracowanie własne z wykorzystaniem programu Statistica.

Rys. 12. Liczebności teoretyczne

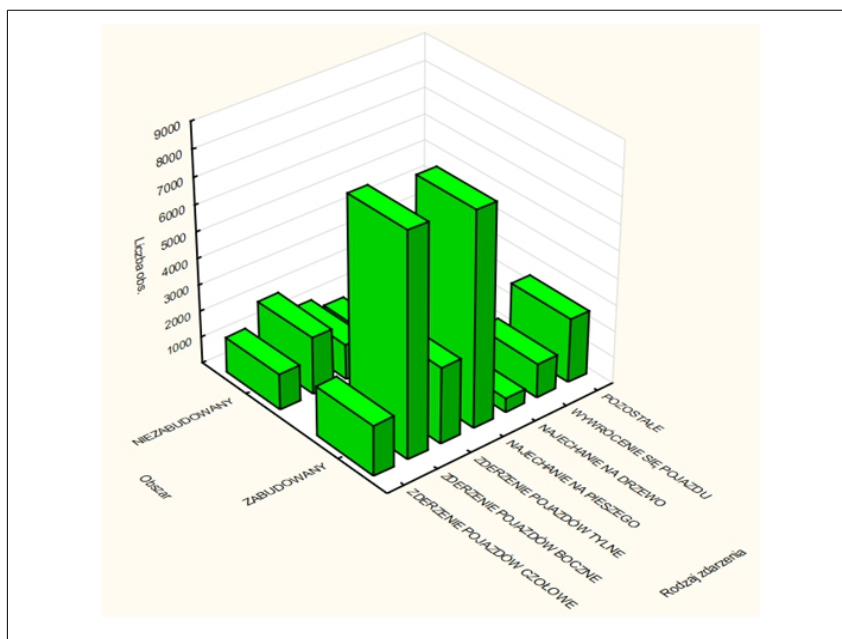
Otrzymane wyniki pozwalają na wyciągnięcie następujących wniosków: przy dowolnie małym (mniejszym niż 0,0000001) poziomie istotności istnieje istotna korelacja pomiędzy rodzajem zderzenia a obszarem, na którym wydarzył się wypadek drogowy. Wskazuje na to wielkość współczynnika zbieżności V – Cramera równa 0,3570519 (rys. 13).

Statystyka: Obszar(2) x Rodzaj zdarzenia(7)	
statystyka	Chi-kwadr.
Chi² Pearsona	4458,188
V Craméra	,3570519

Źródło: opracowanie własne z wykorzystaniem programu Statistica.

Rys. 13. Miary korelacji zmiennych jakościowych

Analizę uzupełniono interpretacją graficzną zależności zmiennych: rodzaj zderzenia a obszar, na którym wydarzył się wypadek drogowy (rys. 14).



Źródło: opracowanie własne z wykorzystaniem programu Statistica.

Rys. 14. Rozkład dwumianowy

Zakończenie

W naukach o bezpieczeństwie można zastosować oprócz podejścia jakościowego również podejście ilościowe. Z powodzeniem wykorzystać można w tej dyscyplinie naukowej rachunek korelacji jako jedną z najbardziej popularnych miar statystycznych

Na podstawie danych wtórnych zawartych w raporcie *Stan bezpieczeństwa ruchu drogowego oraz działania realizowane w tym zakresie w 2014 r.* opracowanym przez Krajową Radę Bezpieczeństwa Ruchu Drogowego ustalono, które zmienne objaśniające są skorelowane z zaobserwowaną w 2014 roku liczbą wypadków w poszczególnych województwach w Polsce. W tym celu wykorzystano program Statistica, który w sposób sprawny umożliwił wyznaczenie współczynników korelacji liniowej Pearsona. Uzyskane wyniki pozwoliły na sformułowanie następujących wniosków:

- pomiędzy liczbą wypadków drogowych w województwach w 2014 roku a liczbą ludności zamieszkałą na ich terenie istnieje znacząca, dodatnia zależność liniowa;
- równie mocno z badaną zmienną zależną skorelowana jest dodatnio: liczba zarejestrowanych samochodów w danym województwie w tys., gęstość zaludnienia w danym województwie mierzona liczbą osób/km², wskaźnik I mierzony liczbą zarejestrowanych samochodów na kilometr drogi w danym województwie.

Przeprowadzono również rachunek korelacji zmiennych jakościowych. Z wykorzystaniem pakietu Statistica ustalono, że pomiędzy rodzajem zderzenia a obszarem, na którym wydarzył się wypadek drogowy, istnieje skorelowanie statystycznie istotne. Jego siłę określono w oparciu o miarę zwaną współczynnikiem zbieżności V – Cramera.

Należy zwrócić uwagę na fakt, że wykorzystanie programu Statistica w rachunku korelacji nie jest jedyną jego funkcją. Oprogramowanie to pozwala przeprowadzić rachunek regresji, którego celem jest ustalanie kształtu zależności oraz opis związku za pomocą funkcji matematycznej. Rachunek regresji może również służyć do przewidywania wartości jednej zmiennej przy określonym poziomie drugiej zmiennej, czyli pozwala modelować przyszłość na podstawie przeszłości. W związku z powyższym należałoby przeprowadzoną w artykule analizę korelacji uzupełnić rachunkiem regresji.

Podsumowując, program Statistica jest to bardzo rozbudowane narzędzie, w skład którego wchodzi wszechstronny zestaw zaawansowanych procedur analitycznych przydatnych w pracy każdego naukowca.

Bibliografia

1. Aczel Amir, *Statystyka w zarządzaniu*, Wydawnictwo Naukowe PWN, Warszawa 2000.
2. Apanowicz Jerzy, *Zarys metodologii prac dyplomowych z organizacji i zarządzania*, Wydawnictwo Wyższego Seminarium Duchownego Bernardinum, Gdynia 1997.
3. Balcerowicz Bolesław, *Bezpieczeństwo polityczne Rzeczypospolitej Polskiej*, ZUMS BN, WSO, AON, Warszawa 2004.
4. Bielecka Anna, *Statystyka w biznesie i ekonomii, teoria i praktyka* Wyd. WSPiZ im. Leona Koźmińskiego, Warszawa 2005.
5. Bielecka Anna, *Statystyka dla menedżerów, teoria i praktyka*, Wolters Kluwer Polska Sp. z o.o., Warszawa 2011.
6. Cieślarczyk Marian (red.), *Metody, techniki i narzędzia badawcze oraz elementy statystyki stosowane w pracach magisterskich i doktorskich*, AON, Warszawa 2003.
7. Cieślarczyk Marian, *Teoretyczne i metodologiczne podstawy badania problemów bezpieczeństwa i obronności państwa*, Siedlce 2009.

8. de Bondy-Łempicka Zofia, Arct Stanisław, *Podręczny Słownik Języka Polskiego*, Wydawnictwo Wiedza Powszechna, Warszawa 1957.
9. Drabik Lidia, Kubiak-Sokół Aleksandra, Sobol Elżbieta, *Słownik języka polskiego*, Wydawnictwo Naukowe PWN, Warszawa 2008.
10. Filipek Agnieszka (red.), *Elementy teorii i praktyki transdyscyplinarnych badań problemów bezpieczeństwa*, tom IV, Uniwersytet Przyrodniczo-Humanistyczny w Siedlcach, Siedlce 2015.
11. Gryz Jarosław, *Zarys podstaw teorii bezpieczeństwa*, AON, Warszawa 2010.
12. Hołyst Brunon, *Kryminologia*, Wydawnictwo LexisNexis, Warszawa 2007.
13. Huzarski Michał, *Zmienne podstawy bezpieczeństwa i obronności państwa*, AON, Warszawa 2009.
14. Huzarski Michał, Wołęjszo Jarosław (red.), *Leksykon obronności, Polska i Europa*, Wydawnictwo Bellona, Warszawa 2014.
15. Kaczmarek Julian, Łepkowski Wojciech, Zdrodowski Bogdan (red.), *Słownik terminów z zakresu bezpieczeństwa narodowego*, WSO, AON, Warszawa 2008.
16. Malinowski Andrzej, *Statystyka opisowa dla studentów administracji i prawa*, Liber Sp. z o.o., Warszawa 2004.
17. Maslow Abraham, *Motywacja i osobowość*, Wydawnictwo Naukowe PWN, Warszawa 2006.
18. Ostasiewicz Stanisława, Rusnak Zofia, Siedlecka Urszula, *Statystyka, Elementy teorii i zadania*, Wyd. AE im. Oskara Langego we Wrocławiu, Wrocław 2001.
19. Piłatowska Mariola, *Repetitorium ze statystyki*, Wydawnictwo PWN, Warszawa 2009.
20. Równiak Michał, *Statistica w badaniach naukowych i nauczaniu statystyki*, StatSoft Polska Sp. z o.o., Kraków 2010.
21. Sobczyk Mieczysław, *Statystyka, podstawy teoretyczne, przykłady zadania*, Wyd. Uniwersytetu Marii Curie-Skłodowskiej, Lublin, 2000.
22. Sobczyk Mieczysław, *Statystyka*, Wydawnictwo Naukowe PWN, Warszawa 2000.
23. Stanisław Andrzej, *Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny*, StatSoft tom 1, Kraków 2006.
24. Starzyńska Wacława, *Statystyka praktyczna*, Wydawnictwo Naukowe PWN, Warszawa 2000.
25. Steczkowski Jan, *Opis Statystyczny, pozyskiwanie, przetwarzanie i analizowanie informacji*, Wydawnictwo Wyższej Szkoły Informatyki i Zarządzania, Rzeszów 2005.
26. Wiśniewski Bernard (red. nauk.), *Bezpieczeństwo w teorii i badaniach naukowych*, Wydawnictwo Wyższej Szkoły Policji, Szczytno 2011.

CORRELATION CALCULATION IN SECURITY SCIENCES USING THE STATISTICA APPLICATION

The aim of this paper is to present the possibility of using the Statistica application to analyse correlation in the science of security.

On the basis of the secondary data contained in the report *The Security of Traffic Movement and the Related Actions Taken in 2014*, which was compiled by the National Roads and Traffic Safety Council, an attempt was made to determine which independent variables are correlated with the number of accidents observed in Polish voivodeships in 2014. A correlation calculation has also been conducted of qualitative variables. With the help of the Statistica application, it was possible to determine whether there existed a correlation that was statistically relevant between a collision type and the area where the road accident occurred. The analysis was illustrated with graphs and screenshots from the running Statistica application.