Statystyka matematyczna w Excelu dla szkół. Ćwiczenia praktyczne

Autor: Andrzej Obecny
ISBN: 83-7197-711-5
Format: B5, stron: około 98

Zaawansowane możliwości arkusza kalkulacyjnego Microsoft Excel czynią zeń doskonałe narzędzie do przeprowadzania analiz statystycznych. MS Excel może w wielu przypadkach zastąpić drogie i skomplikowane w obsłudze pakiety statystyczne.

Książka ta jest drugą publikacją z tej serii poświęconą zastosowaniom Excela w statystyce. W pierwszej książce pt.: „Statystyka opisowa w Excelu dla szkół. Ćwiczenia praktyczne" przedstawione zostały możliwości tego programu w zakresie statystycznej analizy struktury. Tym razem zaprezentowane zostało zastosowanie Excela w innych działach statystyki, a mianowicie w statystyce matematycznej oraz w analizie współzależności i dynamiki zjawisk.

„Statystyka matematyczna w Excelu dla szkół. Ćwiczenia praktyczne" to cenna pomoc dla studentów rozmaitych kierunków, którzy stają przed koniecznością wykonania analizy danych, a także dla wszystkich osób, korzystających w swojej pracy z narzędzi statystyki matematycznej.

Książka opisuje:

- Rozkłady empiryczne i teoretyczne
- Estymację parametryczną
- Metody weryfikacji hipotez (testy istotności i zgodności)
- Korelację i regresję
- Badanie trendów i wahań okresowych

Wydawnictwo Helion
ul. Chopina 6
44-100 Gliwice
tel. (32)230-98-63
e-mail: helion@helion.pl

Spis treści

Rozdział 1.

Rozkłady empiryczne i teoretyczne

Wprowadzenie

Pogrupowane w szereg statystyczny wyniki obserwacji przeprowadzone na populacji empirycznej (w wyniku badania pełnego lub częściowego) nazywamy rozkładem empirycznym. Mówi nam on, ile razy dana wartość badanej cechy występuje w zbiorze obserwacji (dla cechy skokowej) lub ile jednostek należy do określonego przedziału wartości cechy (dla cechy ciągłej).

Dysponując postacią rozkładu empirycznego, możemy dokonać opisu statystycznego zbiorowości lub prowadzić wnioskowanie co do charakteru całej zbiorowości.

W parze z rozkładami empirycznymi idą rozkłady teoretyczne, czyli funkcje (modele) matematyczne, służące przeprowadzeniu analizy statystycznej danego zjawiska. Wśród nich dominującą pozycję zajmują: rozkład normalny, rozkład dwumianowy oraz rozkład Poissona.

W rozdziale tym zajmiemy się tworzeniem rozkładów empirycznych w oparciu o wyniki przeprowadzonych obserwacji oraz tworzyć będziemy zbiorowości o strukturach określonych przez modele teoretyczne. Rozkłady te przedstawimy w formie graficznej w postaci histogramów lub diagramów. Ponadto ćwiczenia dotyczące rozkładów teoretycznych spróbujemy powiązać z pewnymi podstawowymi, ważnymi twierdzeniami.

Za pomocą Excela — jak zobaczymy — wykonanie wszystkich tych zadań nie będzie skomplikowane.

Histogramy i diagramy rozkładów

Histogram będący formą prezentacji szeregu statystycznego (ich budowę i rodzaje omówiono wyczerpująco w *Statystyce opisowej*) to wykres zbudowany z przylegających do siebie prostokątów, w którym podstawa każdego prostokąta oznacza wartość badanej zmiennej, a jego wysokość — liczebność lub częstość względną. Powierzchnie prostokątów są proporcjonalne do liczebności klas.

Z technicznego punktu widzenia (wykonania histogramu w Excelu) nie ma znaczenia, czy badana cecha ma charakter ciągły czy skokowy. Pokażemy to w kilku pierwszych ćwiczeniach.

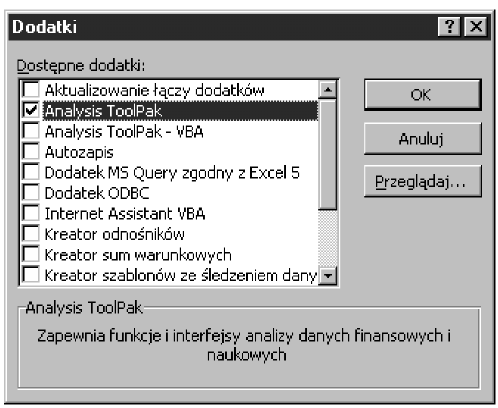
Ćwiczenie 1.1.

W jednej z wyższych uczelni ekonomicznych na Śląsku przeprowadzono ankietę, w której zapytano grupę 192 pracowników naukowych o to, w ilu uczelniach lub szkołach (poza macierzystą) prowadzą jakiekolwiek zajęcia dydaktyczne (cecha y_i) oraz ile osób mają na utrzymaniu (cecha x_i). Wykonaj histogram rozkładu zmiennej X oraz Y.

Zadanie to można rozwiązać podobnie jak to zrobiliśmy to w *Statystyce opisowej*, tj. zastosować funkcję statystyczną *LICZ.JEŻELI()* lub *CZĘSTOŚĆ()* i zbudować szereg rozdzielczy punktowy, po czym na jego podstawie wykonać wykres kolumnowy. Tym razem wykorzystamy specjalnie narzędzie przygotowane do tego w Excelu. Jest nim *Histogram*.

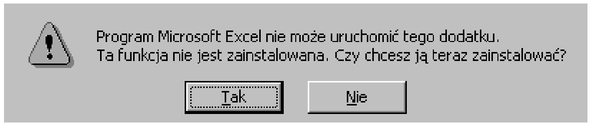
Nim przejdziemy do właściwego wykonania tego ćwiczenia, wyjaśnijmy, czym jest *Histogram* i jak oraz kiedy z niego można skorzystać. Otóż *Histogram* w Excelu to jedno z narzędzi analizy danych statystycznych programu o nazwie *Analysis Toolpak*. Aby możliwe było użycie *Histogramu*, program Analysis Toolpak (jeden z tzw. dodatków Excela) musi być wcześniej zainstalowany i załadowany. W tym celu należy wybrać z paska menu polecenie *Narzędzia/Dodatki*. W oknie, które się wtedy otworzy, należy zaznaczyć wybór programu Analysis Toolpak i zaakceptować ten wybór przyciskiem *OK* (rysunek 1.1).

Rysunek 1.1.
Rysunek pomocniczy do ćwiczenia 1.1



Jeżeli program był wcześniej zainstalowany, to po tej czynności zostanie załadowany do pamięci komputera, czyli stanie się dostępny. Jeżeli jednak po naciśnięciu przycisku *OK* pojawi się okno z komunikatem takim jak na rysunku 1.2, to należy wybrać odpowiedź *TAK* i zainstalować dodatek Analysis Toolpak. Instalacja wymagać będzie jednak dostępu do niezbędnych plików znajdujących się na instalacyjnej płycie CD!

Rysunek 1.2.
Rysunek pomocniczy do ćwiczenia 1.1



Dysponując załadowanym dodatkiem Analysis Toolpak, można przystąpić do wykonywania ćwiczenia.

1. Otwórz skoroszyt *Ćwiczenie1_1.xls*.
2. Określ wartości maksymalne, jakie przyjmują obie badane cechy zmienne.

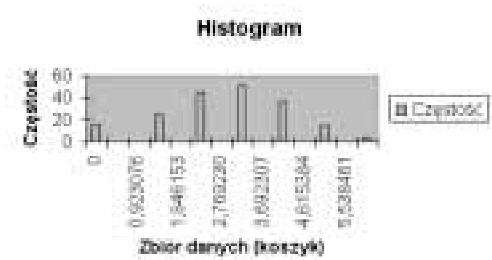
   Do komórek *C2* i *D2* wpisz następujące formuły: `=MAX(A2:A193)`, `=MAX(B2:B193)`.

   Znajomość ich jest konieczna, aby określić przedziały histogramu. Jeżeli końce przedziałów nie zostaną podane, zakres wartości pomiędzy minimum a maksimum zbioru danych zostanie podzielony na przedziały o równej szerokości i w oparciu o nie zostanie zbudowany szereg rozdzielczy oraz wykres (patrz rysunki 1.3 i 1.4). Jak widać, zmienna, która w naszym ćwiczeniu ma charakter skokowy, została „potraktowana" przez program jako cecha ciągła, stąd taki podział na przedziały (klasy). Aby temu zapobiec, należy podać własne zakresy, do których „dostosuje" się program.

Rysunek 1.3.
Rysunek pomocniczy do ćwiczenia 1.1

	A	B
1	Zbiór danych (koszyk)	Częstość
2	0	16
3	0,461538462	0
4	0,923076923	0
5	1,384615385	25
6	1,846153846	0
7	2,307692308	46
8	2,769230769	0
9	3,230769231	51
10	3,692307692	0
11	4,153846154	36
12	4,615384615	0
13	5,076923077	15
14	5,538461538	0
15	Więcej	4

Rysunek 1.4.
Rysunek pomocniczy do ćwiczenia 1.1

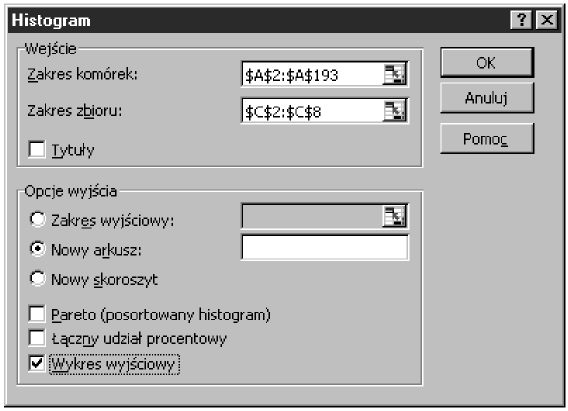


Znasz więc już teraz wszystkie możliwe wartości cechy, jakie mogą przyjmować zmienne *X* i *Y*. Dla zmiennej *X* są to liczby z przedziału między 0 a 6, a dla zmiennej *Y* — od 1 do 7.

3. Przedstaw rozkład zmiennej *X* oraz jej histogram.

Do komórek od *C2* do *C8* wpisz kolejno liczby: 0, 1, 2, 3, 4, 5, 6. Z paska menu wybierz polecenie *Narzędzia/Analiza danych*. Następnie znajdź i wybierz narzędzie analizy o nazwie *Histogram*. Wprowadź wartości odpowiednich pól wg rysunku 1.5.

Rysunek 1.5.
Rysunek pomocniczy do ćwiczenia 1.1

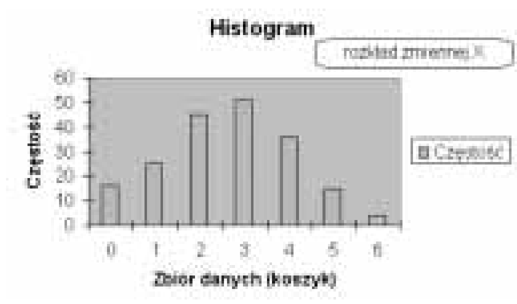


Usuń 9. wiersz powstałego arkusza (z częstością równą 0), aby „poprawić" opis na osi OX wykresu.

Rysunek 1.6.
Fragment arkusza przedstawiający rozwiązanie ćwiczenia 1.1

	A	B
1	*Zbiór danych (koszyk)*	*Częstość*
2	0	16
3	1	25
4	2	45
5	3	51
6	4	36
7	5	15
8	6	4
9		
10	rozkład zmiennej X	

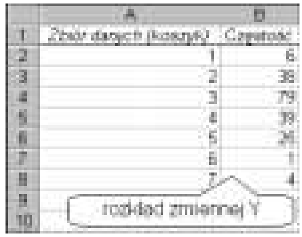
Rysunek 1.7.
Fragment arkusza przedstawiający rozwiązanie ćwiczenia 1.1



4. Przedstaw rozkład zmiennej *Y* oraz jej histogram.

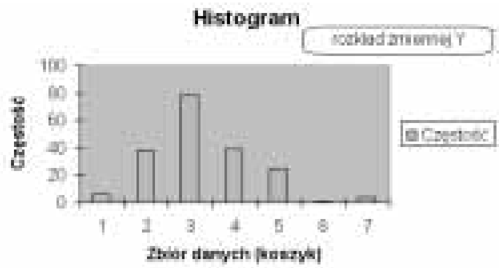
Do komórek od *D2* do *D8* wpisz kolejno liczby: 1, 2, 3, 4, 5, 6, 7. Z paska menu wybierz polecenie *Narzędzia/Analiza danych*, a następnie znajdź i wybierz narzędzie analizy o nazwie *Histogram*. W zakresie danych wejściowych, w polu *Zakres komórek* wpisz `$B$2:$B$193`, zaś w polu *Zakres zbioru* wpisz `$D$2:$D$8`. W opcjach wyjścia zaznacz *Nowy arkusz* i *Wykres wyjściowy*. Na koniec usuń 9. wiersz powstałego arkusza (z częstością równą 0), aby „poprawić" opis na osi OX wykresu.

Rysunek 1.8.
Fragment arkusza przedstawiający rozwiązanie ćwiczenia 1.1



	A	B
1	Zbiór danych (koszyk)	Częstość
2	1	6
3	2	38
4	3	79
5	4	39
6	5	24
7	6	1
8	7	4
9		
10		

Rysunek 1.9.
Fragment arkusza przedstawiający rozwiązanie ćwiczenia 1.1



W poprzednim ćwiczeniu stworzyłeś histogram dla szeregu punktowego, w który zostały pogrupowane wyniki obserwacji. Teraz zbudujesz szereg przedziałowy i również zaprezentujesz go w formie histogramu.

Ćwiczenie 1.2.

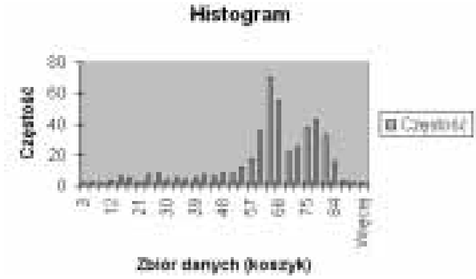
Przeprowadzono badanie statystyczne, w którym ustalono długość życia 456 osób pochowanych na cmentarzu komunalnym w jednym z miast na Pomorzu w ostatnim roku (długość życia zaokrąglano w górę do pełnego roku). Wykonaj histogram rozkładu empirycznego, grupując wyniki w klasy o rozpiętości $C_x = 3$.

1. Otwórz skoroszyt *Ćwiczenie1_2.xls*.
2. Wprowadź następujące granice przedziałów klasowych, na które podzielisz badaną zbiorowość: 1 – 3, 4 – 6, 7 – 9 itd. aż do 88 – 90.

   Do komórek *B2* oraz *B3* wpisz odpowiednio liczby 3 i 6. Następnie zaznacz, przeciągając myszą, obie te komórki i ustaw kursor na uchwycie wypełniania komórki *B3* (uchwyt wypełniania to prawy dolny róg komórki aktywnej). Kursor zmieni wtedy swą postać z grubego, białego krzyżyka na cienki i czarny. Teraz przeciągnij myszą do obszaru *B31*.
3. Wykonaj histogram rozkładu empirycznego.

   Z paska menu wybierz *Narzędzia/Analiza danych*. Następnie znajdź i wybierz *Histogram*. W polu *Zakres komórek* wpisz `$A$2:$A$457`, w polu *Zakres zbioru* wpisz `$B$2:$B$31`, natomiast w opcjach wyjścia zaznacz *Nowy arkusz* oraz *Wykres wyjściowy*.

Rysunek 1.10.
Fragment arkusza przedstawiający rozwiązanie ćwiczenia 1.2



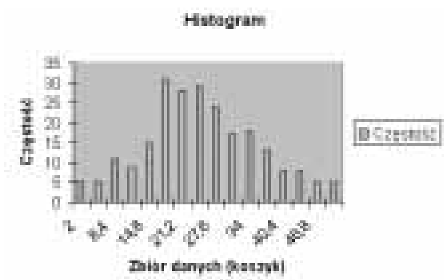
Ćwiczenie 1.3.

Przygotowany został wykaz transakcji dziennych (231 operacji) jednej z kas hipermarketu, który zawiera informacje o liczbie zakupionych przez klienta towarów (zmienna X) i ich wartość (zmienna Y). Wykonaj histogram rozkładu empirycznego zmiennej X oraz Y.

1. Otwórz skoroszyt *Ćwiczenie1_3.xls*.
2. Wykonaj histogram rozkładu zmiennej *X*.

   Z paska menu wybierz *Narzędzia/Analiza danych*. Następnie znajdź i wybierz *Histogram*. W polu *Zakres komórek* wpisz `$A$2:$A$232`. Natomiast w opcjach wyjścia zaznacz *Nowy arkusz* oraz *Wykres wyjściowy*.

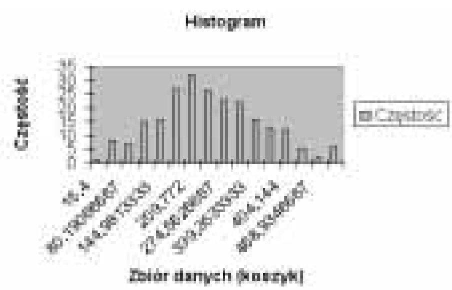
Rysunek 1.11.
Fragment arkusza przedstawiający rozwiązanie ćwiczenia 1.3



3. Wykonaj histogram rozkładu zmiennej *Y*.

   Z paska menu wybierz *Narzędzia/Analiza danych*. Następnie znajdź i wybierz *Histogram*. W polu *Zakres komórek* wpisz `$B$2:$B$232`, natomiast w opcjach wyjścia zaznacz *Nowy arkusz* oraz *Wykres wyjściowy*.

Rysunek 1.12.
Fragment arkusza przedstawiający rozwiązanie ćwiczenia 1.3



Teraz wykonasz ćwiczenie, prezentując rozkład empiryczny w postaci wieloboku liczebności, czyli linii łamanej (diagramu), który uzyskuje się z histogramu przez połączenie odcinkami środków kolejnych górnych boków poszczególnych prostokątów.

Ćwiczenie 1.4.

W pewnej cementowni w Wielkopolsce przeprowadzono kontrolę przestrzegania procedury pakowania cementu do worków. Zważono 100 nominalnie 50-kilogramowych worków cementu z dokładnością do 0,1 kg (zmienna X). Ponadto przyjęto, że pusty worek waży 0,4 kg i uwzględniono to w wynikach pomiarów. Opierając się na uzyskanych danych, zbuduj odpowiedni szereg rozdzielczy i wykonaj diagram uzyskanego rozkładu.

1. Otwórz skoroszyt *Ćwiczenie1_4.xls*.

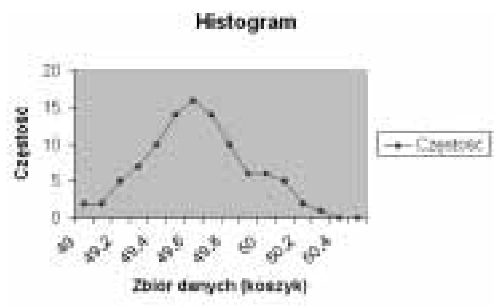
2. Wykonaj histogram rozkładu zmiennej *X*.

Z paska menu wybierz *Narzędzia/Analiza danych*. Następnie znajdź i wybierz *Histogram*. W polu *Zakres komórek* wpisz `$A$2:$A$101`, a w polu *Zakres Zbioru* — `$B$2:$B$16`. W opcjach wyjścia zaznacz *Nowy arkusz* oraz *Wykres wyjściowy*.

3. Zmień histogram na diagram.

W obszarze kreślenia wykresu kliknij prawym przyciskiem myszy. Z rozwiniętego menu podręcznego wybierz polecenie *Typ wykresu*. Na liście typów wykresu wskaż typ *liniowy*.

Rysunek 1.13. *Fragment arkusza przedstawiający rozwiązanie ćwiczenia 1.4*



Rozkład normalny

Najważniejszym rozkładem teoretycznym dla cechy ciągłej w statystyce jest rozkład normalny Gaussa-Laplace'a.

Jego funkcja gęstości prawdopodobieństwa określona dla wszystkich rzeczywistych wartości *x* wyraża się wzorem:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{(x-M)^2}{2\sigma^2}},$$

gdzie:

M — oznacza wartość oczekiwaną zmiennej X,

σ — to odchylenie standardowe zmiennej X.

Jednym z ważnych twierdzeń opartym na założeniu o rozkładzie normalnym zmiennej X jest tzw. reguła trzech sigm. Mówi ona, że realizacje zmiennej losowej nie będą się różniły (in plus ani in minus) od wartości oczekiwanej więcej aniżeli o trzy odchylenia standardowe, co opisuje równanie:

$$P\{M - 3\sigma < X < M + 3\sigma\} = 0{,}9973\,.$$

Ćwiczenie poniższe przybliży tę ważną regułę.

Ćwiczenie 1.5.

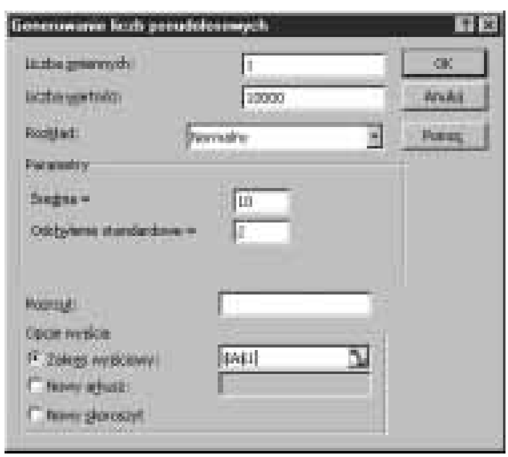
Wygeneruj 10 000 liczb losowych według rozkładu normalnego o średniej M = 10 i odchyleniu standardowym σ = 2. Następnie wykonaj histogram uzyskanego rozkładu empirycznego oraz oblicz liczebności cząstkowe dla następujących przedziałów liczbowych:

- *<4; 16>*,
- *<6; 14>*,
- *<8; 12>*.

Generowanie liczb losowych o zadanych rozkładach możliwe jest w Excelu dzięki narzędziu analizy danych *Generowanie liczb pseudolosowych*. Podobnie jak *Histogram* należy on do pakietu Analysis Toolpak z dodatków Excela.

1. Otwórz skoroszyt *Ćwiczenie1_5.xls*.
2. Wygeneruj 10 000 liczb losowych wg rozkładu normalnego N(10;2).

   Z paska *Narzędzi* wybierz polecenie *Analiza danych*, a następnie wybierz *Generowanie liczb pseudolosowych*. Wprowadź wartości odpowiednich pól wg rysunku 1.14.

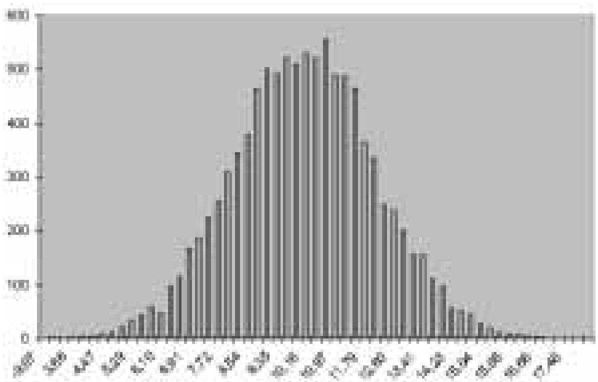


Rysunek 1.14. *Rysunek pomocniczy do ćwiczenia 1.5*

3. Wykonaj histogram rozkładu empirycznego.

Z paska menu wybierz *Narzędzia/Analiza danych*. Następnie znajdź i wybierz *Histogram*. W polu *Zakres komórek* wpisz `$A$1:$A$10000`, natomiast w opcjach wyjścia zaznacz *Nowy arkusz* oraz *Wykres wyjściowy*.

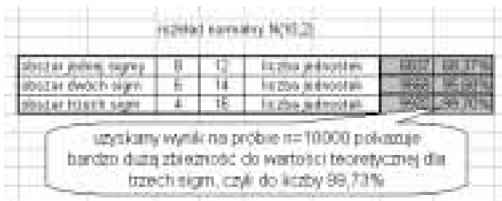
Rysunek 1.15.
Fragment arkusza przedstawiający rozwiązanie ćwiczenia 1.5



4. Oblicz liczebności cząstkowe dla zadanych przedziałów wartości zmiennej *X*.

Do komórek *I8* oraz *J8* wpisz kolejno
`=CZĘSTOŚĆ($A$1:$A$10000;F8)-CZĘSTOŚĆ($A$1:$A$10000;E8), =I8/10000.`
Następnie przekopiuj ich zawartości do obszaru *I9 – J10*.

Rysunek 1.16.
Fragment arkusza przedstawiający przykładowe rozwiązanie ćwiczenia 1.5



Rozkład dwumianowy

Podstawowym rozkładem dla cechy skokowej jest rozkład dwumianowy Bernoulliego.

Prawdopodobieństwo k sukcesów w n próbach określa się w tym rozkładzie wg następującego wzoru:

$$P(X=k)=\binom{n}{k}p^k q^{n-k}, \quad \binom{n}{k}=\frac{n!}{k!\,(n-k)!}, \quad k=0,1,2,\ldots,n,$$

gdzie:

p — oznacza prawdopodobieństwo zajścia zdarzenia A (sukcesu),

q — określa prawdopodobieństwo zajścia zdarzenia przeciwnego do A (porażki).

Generując liczby losowe wg rozkładu dwumianowego, spróbujemy zilustrować inne fundamentalne twierdzenie wykorzystywane w statystyce, a mianowicie prawo wielkich liczb (dokładniej mówiąc, tzw. słabe uogólnienie prawa wielkich liczb).

Mówi ono, że przy wzroście liczebności próby prawdopodobieństwo tego, że bezwzględna różnica między średnią arytmetyczną z próby a wartością oczekiwaną ze zbiorowości generalnej jest mniejsza od dowolnie małej liczby dodatniej dąży do jedności, co zapisuje się wzorem:

$$\lim_{n\to\infty} (P|\bar{x} - M| < \varepsilon) = 1.$$

Dla rozkładu dwumianowego wartość oczekiwana i odchylenie standardowe wynoszą odpowiednio:

$$E(X) = n\,p, \qquad D(X) = \sqrt{n\,p\,q}.$$

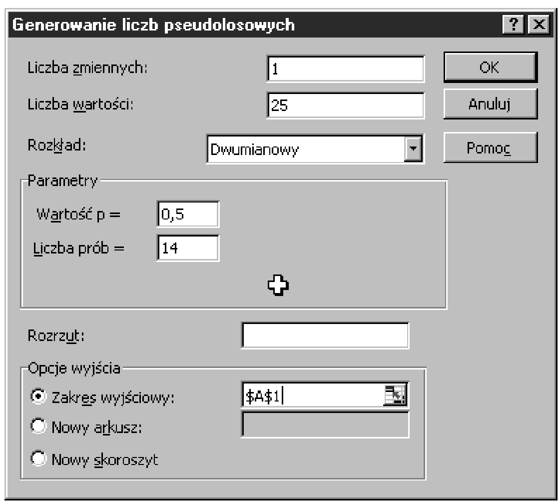
Ćwiczenie 1.6.

Wygeneruj kolejno 25, 100, 400 i 800 liczb losowych według rozkładu dwumianowego o parametrach: p = 0,5 i liczbie prób (doświadczeń) n = 14. Następnie dla każdej wygenerowanej serii liczby oblicz ich średnie arytmetyczne oraz odchylenia standardowe.

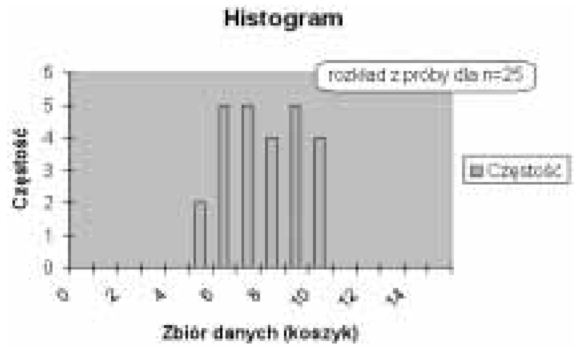
1. Otwórz skoroszyt *Ćwiczenie1_6.xls*.
2. Wygeneruj 4 serie liczb losowych wg rozkładu dwumianowego o zadanych parametrach.

   Z paska *Narzędzi* wybierz polecenie *Analiza danych*, a następnie wybierz *Generowanie liczb pseudolosowych*. Wprowadź wartości odpowiednich pól wg rysunku 1.17.

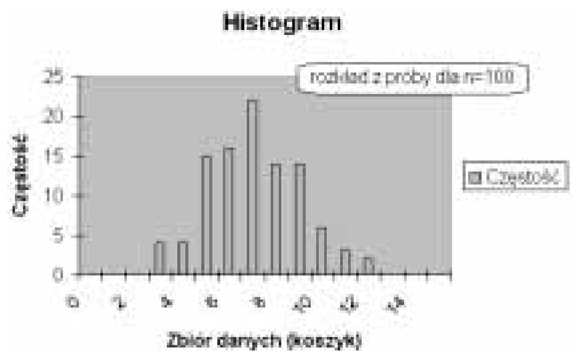
Rysunek 1.17.
Rysunek pomocniczy do ćwiczenia 1.6



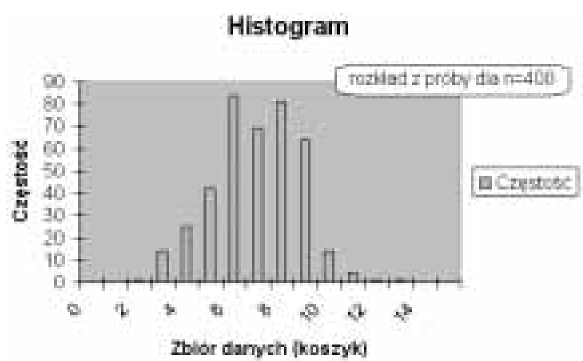
Następnie powtórz procedurę dla 100, 400 i 800 liczb, zmieniając jedynie wartość pola *Zakres wyjściowy* w *Opcjach wyjścia*, podając kolejno następujące adresy komórek: `$B$1`, `$C$1`, `$D$1`.



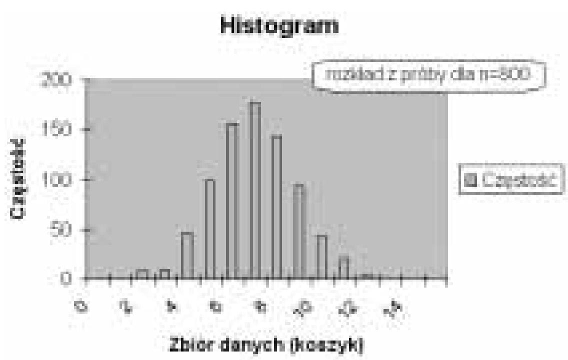
Rysunek 1.18.
Rysunek pomocniczy do ćwiczenia 1.6



Rysunek 1.19.
Rysunek pomocniczy do ćwiczenia 1.6



Rysunek 1.20.
Rysunek pomocniczy do ćwiczenia 1.6



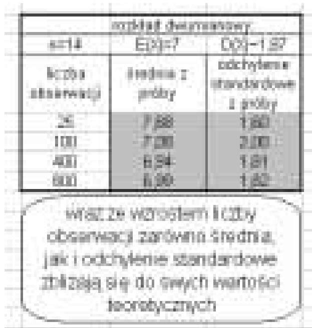
Rysunek 1.21.
Rysunek pomocniczy do ćwiczenia 1.6

3. Oblicz średnie arytmetyczne dla wygenerowanych serii danych.

Do komórek od *G8* do *G11* wpisz kolejno: `=ŚREDNIA(A1:A25)`, `=ŚREDNIA(B1:B100)`, `=ŚREDNIA(C1:C400)`, `=ŚREDNIA(D1:D800)`.

4. Oblicz odchylenia standardowe dla wygenerowanych serii danych.

Do komórek od *H8* do *H11* wpisz kolejno: `=ODCH.STANDARDOWE(A1:A25)`, `=ODCH.STANDARDOWE(B1:B100)`, `=ODCH.STANDARDOWE(C1:C400)`, `=ODCH.STANDARDOWE(D7:D800)`.



Rysunek 1.22. *Fragment arkusza przedstawiający przykładowe rozwiązanie ćwiczenia 1.6*

Rozkład Poissona

Innym ważnym rozkładem dla cechy skokowej jest rozkład Poissona.

W rozkładzie tym prawdopodobieństwo, że zmienna losowa X przyjmie wartość k, obliczamy ze wzoru:

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda},$$

gdzie:

$\lambda > 0$ — to stały parametr rozkładu.

W rozkładzie tym wartość oczekiwana i odchylenie standardowe wynoszą odpowiednio:

$$E(X) = n\,p = \lambda, \qquad D(X) = \sqrt{n\,p} = \sqrt{\lambda}.$$

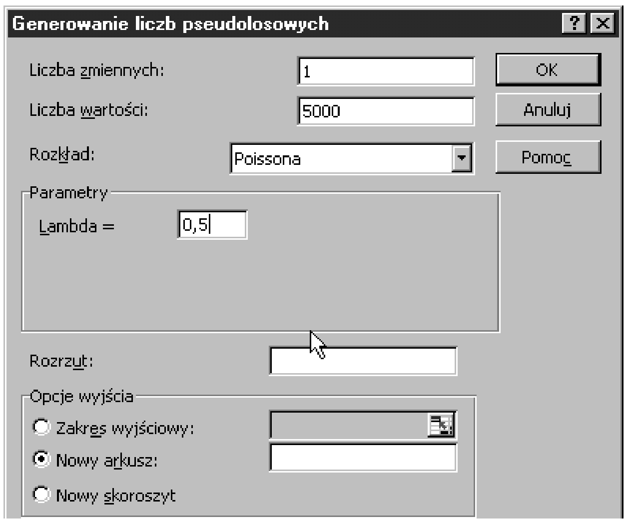
Wykres histogramu wykonanego dla zmiennej losowej o rozkładzie Poissona charakteryzuje się różnym stopniem asymetrii w zależności od wartości parametru λ. Zobaczymy to w poniższym ćwiczeniu.

Ćwiczenie 1.7.

Wygeneruj 5000 liczb losowych według rozkładu Poissona dla parametru $\lambda = 0,5$. Następnie w oparciu o uzyskany zbiór liczb zbuduj szereg rozdzielczy punktowy i przedstaw go na histogramie. Czynności te powtórz jeszcze dwukrotnie, raz dla parametru $\lambda = 5$, a drugi dla $\lambda = 50$.

1. Utwórz nowy, pusty skoroszyt.
2. Wygeneruj 5000 liczb losowych wg rozkładu Poissona o zadanym parametrze.

   Z paska *Narzędzi* wybierz polecenie *Analiza danych*, a następnie wybierz *Generowanie liczb pseudolosowych*. Wprowadź wartości odpowiednich pól wg rysunku 1.23.



Rysunek 1.23. *Rysunek pomocniczy do ćwiczenia 1.7*

3. Wyznacz najmniejszą i największą wygenerowaną liczbę.

   Do komórek *B1* i *B2* nowo utworzonego arkusza wpisz kolejno: `=MIN(A1:A5000)`, `=MAX(A1:A5000)` (wartość najmniejszą i największą zbioru możesz również ustalić, korzystając z funkcji sortowania).

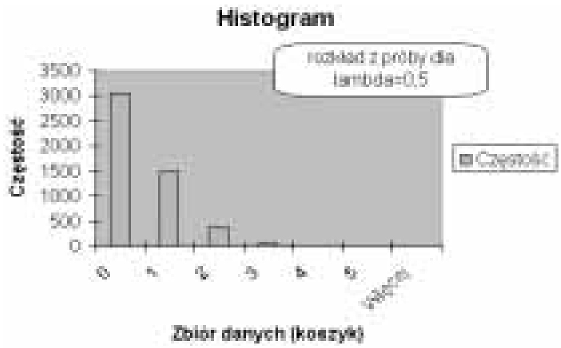
4. Określ przedziały klasowe dla szeregu rozdzielczego.

   Do komórki *C1* wpisz wartość najmniejszej liczby. Następnie do kolejnych komórek w kolumnie *C* wpisuj liczby zawsze większe o jeden od poprzedniej, aż otrzymasz liczbę odpowiadającą liczbie największej w całym wygenerowanym zbiorze liczb.

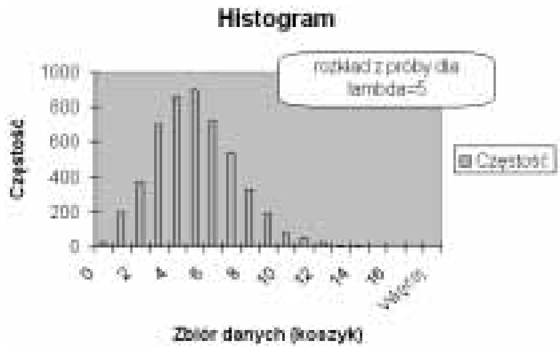
5. Wykonaj histogram rozkładu empirycznego.

   Z paska menu wybierz *Narzędzia/Analiza danych*. Następnie znajdź i wybierz *Histogram*. W polu *Zakres komórek* wpisz `$A$1:$A$5000`, a w polu *Zakres zbioru* wpisz zakres komórek z przygotowanymi przedziałami klasowymi. W opcjach wyjścia zaznacz *Nowy arkusz* oraz *Wykres wyjściowy*.

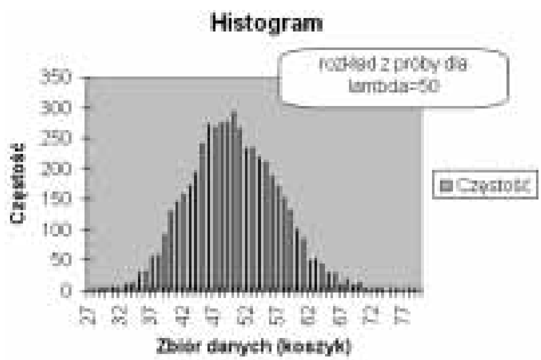
   Powtórz procedurę od punktu 2 dla parametru $\lambda = 5$, a następnie dla parametru $\lambda = 50$.



Rysunek 1.24. *Fragment arkusza przedstawiający przykładowe rozwiązanie ćwiczenia 1.7*



Rysunek 1.25. *Fragment arkusza przedstawiający przykładowe rozwiązanie ćwiczenia 1.7*



Rysunek 1.26. *Fragment arkusza przedstawiający przykładowe rozwiązanie ćwiczenia 1.7*

Obliczenia numeryczne wykonywane dla rozkładu dwumianowego wg podanego wcześniej wzoru przy dużych wartościach n nie są wygodne. Można wówczas skorzystać z twierdzenia Poissona. Mówi ono, że jeżeli $n \to \infty$ i $p \to 0$ tak, że $np = \lambda > 0$, to

$$\lim_{n\to\infty} P\{S_n = k\} = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Wzór ten pozwala w praktyce już dla niewielkich n (rzędu kilkudziesięciu), przy małych p (dla których iloczyn $\lambda = np$ nie przekracza 10), obliczać prawdopodobieństwo k sukcesów w serii n niezależych doświadczeń. A zatem

$$P(X = k) = \binom{n}{k} p^k q^{n-k} \approx \frac{\lambda^k}{k!} e^{-\lambda}.$$

Ostatnie ćwiczenie tego rozdziału zilustruje to twierdzenie.

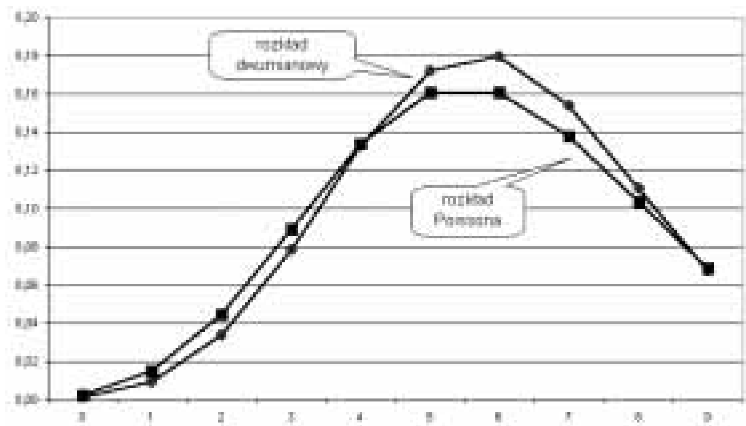
Ćwiczenie 1.8.

Prawdopodobieństwo zajścia pewnego zdarzenia (sukcesu) wynosi $p = 0{,}2$. Przeprowadzono serię 30 niezależnych doświadczeń. Oblicz prawdopodobieństwo łącznej ilości k sukcesów, gdzie $k = 0{,}1,\ldots 9$. W rozwiązaniu zastosuj wzór Bernoulliego oraz Poissona. Zinterpretuj uzyskane wyniki.

1. Otwórz skoroszyt *Ćwiczenie1_8.xls*.
2. Oblicz prawdopodobieństwa uzyskania kolejno: 0, 1, 2, itd. aż do 9 sukcesów, stosując wzór Bernoulliego.

   Do komórki *B2* wpisz `=KOMBINACJE(30;A2)*(0,2)^A2*(0,8)^(30-A2)`. Następnie przekopiuj jej zawartość do obszaru *B3 – B11*.
3. Oblicz ponownie prawdopodobieństwa uzyskania tych samych ilości sukcesów, stosując wzór Poissona.

   Do komórki *C2* wpisz `=6^A2*EXP(-6)/SILNIA(A2)` i przekopiuj jej zawartość do komórek *C3 – C11*.



Rysunek 1.27. *Rysunek pomocniczy do ćwiczenia 1.8*