Weryfikacja hipotez

Każde badanie naukowe rozpoczyna się od sformułowania problemu badawczego oraz najbardziej prawdopodobnego (na gruncie wiedzy badającego) ogólnego rozwiązania, czyli hipotezy badawczej. Poprawne sformułowanie hipotezy w dużej mierze przesądza o sukcesie badawczym. Ogólnie mówiąc hipotezy mogą dotyczyć:

- wartości badanych zmiennych (np. średnia waga zwierząt chorych na pewną chorobę wynosi 45 kilogramów)
- różnicy między cechami opisującymi badaną grupę (populację) (np. lek A skuteczniej obniża ciśnienie od leku B; istnieje różnica między czterema metodami czasu krzepnięcia osocza)
- zależności między badanymi zmiennymi (np. istnieje silna korelacja dodatnia dla psów między poziomem glukozy w surowicy a prawdopodobieństwem wystąpienia cukrzycy lub zapalenia nerek)
- "kształtu" zależności badanych zmiennych (np. istnieje zależność logarytmiczna między wzrostem a wagą zwierząt chorych chronicznie na schorzenie Y)
- porównania rozkładów zmiennych (np. rozkład zmiennej "cholesterol" jest rozkładem normalnym)

Weryfikacja hipotez statystycznych to zasadnicza domena statystyki. W dalszych rozważaniach przez **hipotezę statystyczną** rozumieć będziemy dowolne przypuszczenie dotyczące rozkładu populacji generalnej (jego postaci lub wartości parametrów) bez przeprowadzenia badania całkowitego. Prawdziwość hipotezy będziemy weryfikować na podstawie wyników próby losowej. Najczęściej hipoteza jest formułowana w taki sposób, aby na podstawie wyników z próby można ją było łatwo odrzucić.

Tradycyjnie dzielimy hipotezy statystyczne na dwie grupy:

- **parametryczne**, gdy dotyczą wartości parametrów statystycznych populacji (np. średnia, wariancja)
- **nieparametryczne**, gdy dotyczą postaci rozkładu cech lub losowości próby.

Proces weryfikacji hipotezy przebiega według pewnego schematu, którego najważniejsze etapy zostały omówione poniżej.

Etap I. Formułowanie hipotezy zerowej H_0 oraz związanej z nią hipotezy alternatywnej H_1.

Weryfikacja hipotez rozpoczyna się zwykle od postawienia tej hipotezy, która będzie podlegała sprawdzeniu. Taką hipotezę nazywamy **hipotezą zerową** i oznaczamy $\mathbf{H_0}$. Następnie formułujemy (konkurencyjną) hipotezę, którą jesteśmy skłonni przyjąć, gdy odrzucamy hipotezę zerową. Taką hipotezę nazywamy **hipotezą alternatywną** i oznaczamy $\mathbf{H_1}$. W problemie testowania zawsze **muszą** być sformułowane obie hipotezy

Przykładowo chcemy stwierdzić, że paciorkowce odpowiedzialne są za cięższe przypadki choroby niż gronkowce. Choć wyniki z próby mogą częściowo wskazywać, że tak jest, musimy być ostrożni. Mogło to być dziełem przypadku - np. nietypowego doboru próby. Dlatego jako hipotezę zerową przyjmujemy zazwyczaj brak różnic. Natomiast to, co nas interesuje, to hipoteza alternatywna mówiąca o zachodzeniu różnic np. między działaniem różnych bakterii. Odrzucając bowiem hipotezę zerową przyjmujemy alternatywną, a jej nieodrzucenie stawia nas w trudnej sytuacji. Nie można oczywiście przyjąć hipotezy alternatywnej, ani zerowej, bo jej nieodrzucenie wynikać może np. z niewłaściwego doboru grupy próbnej lub z jej zbyt małej liczebności.

Etap II. Przyjmujemy odpowiedni poziom istotności

W wyniku testowania hipotezy statystycznej jak i estymacji podejmujemy pewną decyzję – jest to decyzja względnie subiektywna. Stosując odpowiednie procedury statystyczne nie orzekamy o prawdziwości bądź fałszywości hipotezy, ani o tym, że **prawdziwa wartość** szacowanego parametru jest taka jak obliczona z próby. Podejmujemy decyzję z wszelkimi konsekwencjami wynikającymi z decyzji błędnej. Oczywiście weryfikacja powinna przebiegać tak, aby zapewnić jak najmniejsze prawdopodobieństwo pomyłki. W trakcie weryfikacji można popełnić dwa błędy:

- **błąd pierwszego rodzaju**, polegający na odrzuceniu hipotezy zerowej, mimo że jest ona prawdziwa.
  Prawdopodobieństwo popełnienia błędu pierwszego rodzaju nazywamy poziomem istotności i oznaczamy przez α. Najczęściej przyjmowane są wartości 0,05 oraz 0,01 i 0,001.
- **błąd drugiego rodzaju**, polegający na przyjęciu hipotezy zerowej, gdy jest ona w rzeczywistości fałszywa.
  Prawdopodobieństwo popełnienia błędu drugiego rodzaju oznaczać będziemy przez β.

Zestawienie omawianych błędów przedstawione jest w poniższej tabeli:

Hipoteza zerowa	**DECYZJE**	
	Nie Odrzucić H_0	**Odrzucić H_0**
Hipoteza zerowa prawdziwa	**decyzja prawidłowa**	***błąd I rodzaju*** α
Hipoteza zerowa fałszywa	***błąd II rodzaju*** β	**decyzja prawidłowa**

Wartości α i β są ze sobą powiązane. Zmniejszenie prawdopodobieństwa α powoduje wzrost prawdopodobieństwa β. Pewnym kompromisem w tej sytuacji są tzw. testy istotności, które dla wybranego przez nas z góry poziomu istotności α zapewniają możliwie najmniejszą wartość prawdopodobieństwa β. W tym i pozostałych artykułach będziemy mówić tylko o takich testach. Wartość $1 - \beta$ jest nazywana mocą testu.

W naukach biologicznych przyjmujemy wartość $\alpha = 0,05$ lub mniejszą. Poziom istotności α powinien być ustalony w zasadzie przed rozpoczęciem testowania. Jaki jest sens tej liczby? Przypominamy, że poziom istotności wskazuje na jak mały błąd "wyrażamy" zgodę, np. poziom 0,01 świadczy, że jesteśmy skłonni popełnić jeden błąd na 100 badań. Pamiętajmy także, że wybierając niższy poziom istotności uzyskujemy wyższy poziom wiarygodność hipotezy alternatywnej (jej przyjęcie jest jakby mocniej uzasadnione), ale będzie nam trudniej odrzucić hipotezę zerową.

Przy weryfikacjach hipotez za pomocą pakietów komputerowych ważne staje się zrozumienie poziomu prawdopodobieństwa, oznaczanego przez p. Autorzy badań naukowych podają zwykle tą wartość p. Jest to najmniejsze prawdopodobieństwo, przy którym hipoteza zerowa może być odrzucona.

Pamiętajmy, jeżeli $\alpha > p$, to na danym poziomie istotności α odrzucamy hipotezę zerową, natomiast gdy $\alpha < p$, to na danym poziomie istotności α nie ma podstaw do odrzucenia hipotezy zerowej. Porównanie tych dwóch wartości to podstawa przy podejmowaniu decyzji weryfikacyjnych.

Etap III. Wybór - odpowiedniego do postawionej hipotezy zerowej - testu i obliczenie jego wartości w oparciu o dane pochodzące z próby.

Jest to kolejna ważna decyzja podejmowana w trakcie weryfikacji hipotez. Wybór bowiem niewłaściwego testu przekreśli wartość całego późniejszego rozumowania. Musimy wiedzieć, jakie jest odpowiednie "narzędzie" dla naszego problemu i naszych badanych zmiennych. Musimy również zawsze sprawdzić czy założenia wybranego przez nas testu są w pełni spełnione. Możemy dobierać wiele statystyk jako podstawę testu dla zweryfikowania tej samej hipotezy H_0. Na decyzję o wyborze konkretnej statystyki ma wpływ, między innymi, hipoteza alternatywna. Staramy się wybierać test najmocniejszy (w pewnej klasie testów), to jest taki, gdzie przy ustalonym α wartość β przybiera wartość najmniejszą.

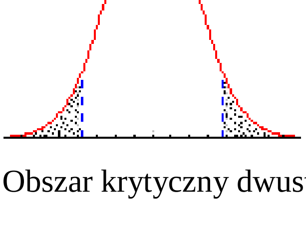
Etap IV. Przy ustalonym poziomie istotności znajdujemy obszary krytyczne i w oparciu o nie podejmujemy decyzję o odrzuceniu lub nie hipotezy zerowej

Jaka jest zasadnicza idea tworzenia obszarów krytycznych? Otóż robiąc bardzo ważne założenie, że hipoteza H_0 jest prawdziwa oraz posługując się matematyczną teorią (opisującą naszą zmienną) tworzy się pewną zmienną losową (statystykę) Z. Następnie określa się wartości jakie musiałaby przyjąć ta zmienna losowa, aby było to "mało prawdopodobne", to znaczy prawdopodobieństwo zaistnienia tych wartości byłoby równe poziomowi istotności. Te "mało prawdopodobne" wartości tworzą tzw. obszar krytyczny. Następnie, jeśli wartość testu obliczona dla grupy próbnej znalazła się w obszarze krytycznym, to wystąpiło zdarzenie bardzo mało prawdopodobne. Zdarzenie takie praktycznie nie powinno zaistnieć. Skoro jednak zaszło (a mając zaufanie do obliczeń w grupie próbnej) coś nie tak z prawdziwością hipotezy zerowej. Nie jest więc spełnione założenie o prawdziwości hipotezy zerowej (wykorzystane do tworzenia obszaru krytycznego). Ostatecznie hipotezę zerową odrzucamy i przyjmujemy hipotezę alternatywną.

Lokalizacja obszaru krytycznego zależy od postaci hipotezy alternatywnej. Przypuśćmy, że efektywność leku mierzymy czasem jego działania. Hipoteza H_0 teraz brzmi - średni czas działania nowego leku ($\overline{X}$) równy jest pewnej wartości A (czas działania stosowanych do tej pory specyfików). Formalizując mamy - H_o: $\overline{X} = A$. Możemy sformułować trzy hipotezy alternatywne:

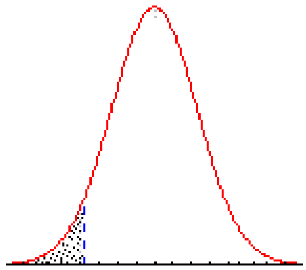
- $\overline{X} \neq A$
- $\overline{X} < A$
- $\overline{X} > A$

W pierwszym przypadku obszar krytyczny obejmowałby wszystkie wartości testu "dużo" większe od A i "dużo" mniejsze od A. Taki obszar nazywany jest obszarem dwustronnym (rysunek poniżej)



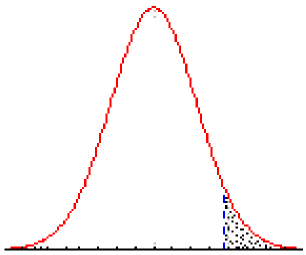
Rys. 1 Obszar krytyczny dwustronny

W drugim przypadku obszar krytyczny obejmowałby wszystkie wartości testu “dużo” mniejsze od A. Taki obszar nazywany jest obszarem lewostronnym. (rysunek poniżej).



Rys. 2 Obszar krytyczny lewostronny

W trzecim przypadku obszar krytyczny obejmowałby wszystkie wartości testu “dużo” większe od A. Taki obszar nazywany jest obszarem prawostronnym. (rysunek poniżej). I taki obszar krytyczny przyjęlibyśmy w naszym przypadku. Oczekujemy bowiem, iż nowe lekarstwo wspomaga proces leczenia, czyli działa dłużej. Jest to przykład hipotezy zerowej. Oczekujemy bowiem wzrostu współczynnika zdrowia, jeśli pacjenci zażywają lekarstwo.



Rys. 3 Obszar krytyczny prawostronny

Stosowanie w praktyce metod statystyki matematycznej nie może ograniczać się do mechanicznego stosowania gotowych „recept” i wzorów. Musi być poprzedzone rzetelną analizą i sprawdzeniem założeń stosowanej, w konkretnym przypadku, metody. Tylko przy znajomości wszystkich wymogów teorii można z powodzeniem stosować ją w praktyce, unikając tzw. „błędu trzeciego rodzaju” polegającego na wyciąganiu fałszywych wniosków z prawdziwych danych statystycznych.

Pamiętajmy też, że wyniki weryfikacji hipotez zależą również od liczebności grupy próbnej (liczby dokonanych pomiarów). Im grupa mniejsza tym większe prawdopodobieństwo przypadkowości. Bardziej rzeczywiste będą różnice zaobserwowane w eksperymencie dla 60 osób niż dla 10. Poniżej podsumowaliśmy kolejne etapy wnioskowania statystycznego na “piechotę” i za pomocą programów statystycznych Przeprowadzanie “na piechotę” weryfikacji hipotez wymaga dużo wprawy, cierpliwości i znajomości tablic statystycznych.

Jak pokazuje poniższe zestawienie programy komputerowe (przykładowo pakiet statystyczny *STATISTICA*) uwalnia nas od najbardziej pracochłonnego etapu obliczeń. Pozostawia nam interpretację otrzymanych wyników (co jest sprawą ważną jak i trudną).

Etapy wnioskowania statystycznego

“na piechotę”	**z użyciem pakietu statystycznego**
1. formułowanie hipotezy zerowej i alternatywnej	1. formułowanie hipotezy zerowej i alternatywnej
2. przyjęcie poziomu istotności	2. przyjęcie poziomu istotności

3. sprawdzenie założeń wybranego testu	3. sprawdzenie założeń wybranego testu
4. obliczanie wartości testu na podstawie wyników z próby **5. znajdowanie wartości krytycznej z tablic statystycznych dla ustalonego poziomu istotności**	
6. podjęcie decyzji o odrzuceniu lub nie hipotezy zerowej na danym poziomie istotności	4. Porównanie p i α umożliwia podjęcie decyzji o odrzuceniu lub nie hipotezy zerowej na danym poziomie istotności α.
7. interpretacja otrzymanych wyników	5. interpretacja otrzymanych wyników

Rys. 4 Etapy wnioskowania statystycznego

Etap V. Interpretacja wyników

To jeden z najtrudniejszych i najważniejszych kroków w przeprowadzanej analizie statystycznej. Poprawna interpretacja analizy statystycznej nie może być niezależna od wiedzy na temat charakteru i sposobu otrzymywania danych. Suche liczby to za mało. Najlepiej, gdy interpretacji wyników dokonuje lekarz łącznie ze statystykiem. Współpraca badacza ze statystykiem powinna obejmować:

- badacz formułuje hipotezę badawczą i wyjaśnia ją możliwie dokładnie statystykowi,
- statystyk tłumaczy ją na hipotezę statystyczną (problem statystyczny) i rozwiązuje ten problem metodami statystycznymi,
- statystyk wyjaśnia badaczowi uzyskane rezultaty i ich interpretację, najlepiej w terminach badania (nie w terminach statystycznych).

Pamiętajmy też, że test statystyczny nie jest dowodem prawdziwości czy fałszywości hipotezy. Za pomocą testu można albo odrzucić hipotezę zerową, albo też orzec, że wyniki doświadczenia nie przeczą tej hipotezie. Ocena testu statystycznego ma na ogół postać zdania: "Na ustalonym poziomie istotności α = ... hipotezę zerową H_0 .... odrzucamy lub nie mamy podstaw do jej odrzucenia".

Nieodrzucenie hipotezy zerowej nie jest równoważne z jej przyjęciem. Wynik "nieistotny" nie oznacza nieważny lub nieistniejący Wyniki te niosą też pewną informację o efektach mających znaczenie kliniczne. Najlepiej traktować te wyniki jako "nieudowodnione". Być może np. zwiększenie liczebności próby zmieni sytuację. Natomiast przyjęcie hipotezy alternatywnej ma charakter wniosku pozytywnego. Orzekamy bowiem zachodzenie pewnych różnic, a nie tylko brak podstaw do założenia różnicy lub identyczności populacji.

Przykład 1

Dla przykładu przeanalizujmy za pomocą programu *STATISTICA* dane dotyczące poziomu sodu dla 40 psów z naszej przykładowej bazy. Będziemy chcieli sprawdzić hipotezę, że średnia poziom sodu wynosi 134 mmol/l.

Etap I - Formułowanie hipotez.

Formalnie możemy to zapisać jako:

H_0: $\mu = 134$ – średni poziom sodu w surowicy jest równy 134 mmol/l

Hipotezą alternatywną wobec H_0 jest

H_1: $\mu \neq 134$ – średni poziom sodu w surowicy jest różny od 134 mmol/l

Etap II – Przyjęcie poziomu istotności.

Jako poziom istotności przyjmujemy wartość 0,05

Etap III – Dobranie odpowiedniego testu i wyliczenie jego wartości w oparciu o dane pochodzące z próby.

Do weryfikacji hipotezy wykorzystujemy testy istotności dla nieznanej średniej populacji. W programie *STATISTICA* do weryfikacji hipotez dotyczących wartości średniej służy opcja **Test t dla pojedynczych średnich** w module S**tatystyki podstawowe i tabele**. Musimy tylko sprawdzić założenie o normalności rozkładu poziomu sodu w organizmie. Obecnie omówimy otrzymane wyniki i przedstawimy najciekawsze interpretacje geometryczne otrzymanych wyników.

Dla naszych przykładowych danych otrzymamy następujący arkusz wyników:

	Test średnich względem stałej wartości odniesienia (Baza danych)							
Zmienna	Średnia	Odch.st.	Ważnych	Bł. std.	Odniesienie Stała	t	df	p
Sód	138,5500	6,868546	40	1,086012	134,0000	4,189639	39	0,000155

Rys. 5 Arkusz wyników testu t dla pojedynczych średnich

Etap IV – Podjęcie decyzji o odrzuceniu lub nie hipotezy zerowej na danym poziomie istotności.

Porównujemy otrzymaną wartość p z przyjętym poziomem istotności. Ponieważ $p = 0{,}000155 < 0{,}05$, więc mamy podstawy do odrzucenia hipotezy zerowej. Reasumując na poziomie istotności 0,05 odrzucamy hipotezę zerową i przyjmujemy, że przeciętny poziom sodu w surowicy jest różny od 134 mmol/l.

Rozkład - normalny

Często, w celu uzyskania matematycznego opisu wyników obserwacji lub eksperymentu, usiłujemy przypisać każdemu z nich pewną liczbę rzeczywistą. Określamy w ten sposób zmienną losową. Mówiąc precyzyjniej zmienną losową nazywamy funkcję, określoną na przestrzeni zdarzeń elementarnych, która każdemu zdarzeniu elementarnemu przyporządkowuje liczbę rzeczywistą z określonym prawdopodobieństwem. Wartości jej nie możemy więc z góry przewidzieć, bowiem zależy ona od przyczyn losowych. Zmienne losowe oznaczamy zazwyczaj dużymi literami X, Y, natomiast ich realizacje, czyli wartości jakie one przyjmują odpowiednio małymi literami x, y.

Jeżeli zbiór wartości zmiennej losowej jest zbiorem przeliczalnym (lub skończonym), wówczas zmienną losową nazywamy dyskretną. Jeżeli natomiast zmienna losowa przyjmuje wszystkie wartości z pewnego przedziału liczbowego, to nazywamy ją zmienną losową ciągłą.

Rozkład prawdopodobieństwa zmiennej losowej dyskretnej można przedstawić w postaci funkcji prawdopodobieństwa. Przypuśćmy, że zmienna losowa X przyjmuje wartości x_1, x_2, ...,

x_n z prawdopodobieństwami, odpowiednio p_1, p_2, ..., p_n. Wówczas funkcja prawdopodobieństwa zmiennej losowej X przedstawia się następująco:

$$P(X = x_i) = p_i \text{ dla } i = 1,2,...,n \text{ oraz } \sum_{i=1}^{n} p_i = 1$$

Zestawienie wszystkich możliwych par (x_i, p_i) - wartości zmiennej z ich wartościami prawdopodobieństw - nazywamy rozkładem prawdopodobieństwa zmiennej losowej dyskretnej X. Rozkład prawdopodobieństwa można przedstawić w postaci tabeli.

Wartości zmiennej dyskretnej x_i	**x_1**	**x_2**	**.....**	**x_n**
Prawdopodobieństwo każdej powyższej wartości p_i	**p_1**	**p_2**	**......**	**p_n**

Przykładowo rzucamy raz kostką do gry. Zbiór zdarzeń elementarnych to wszystkie możliwe wyniki rzutu kostką (od wyrzucenia jednego oczka do wyrzucenia sześciu oczek). Prawdopodobieństwo zajścia każdego zdarzenia elementarnego wynosi oczywiście 1/6. Na tym zbiorze określamy zmienną losową X, która każdemu elementarnemu zdarzeniu przyporządkuje liczbę rzeczywistą, równą ilości wyrzuconych oczek. Rozkład tej zmiennej można przedstawić w postaci tabeli:

Wartość zmiennej x_i	1	2	3	4	5	6
Prawdopodobieństwo p_i	1/6	1/6	1/6	1/6	1/6	1/6

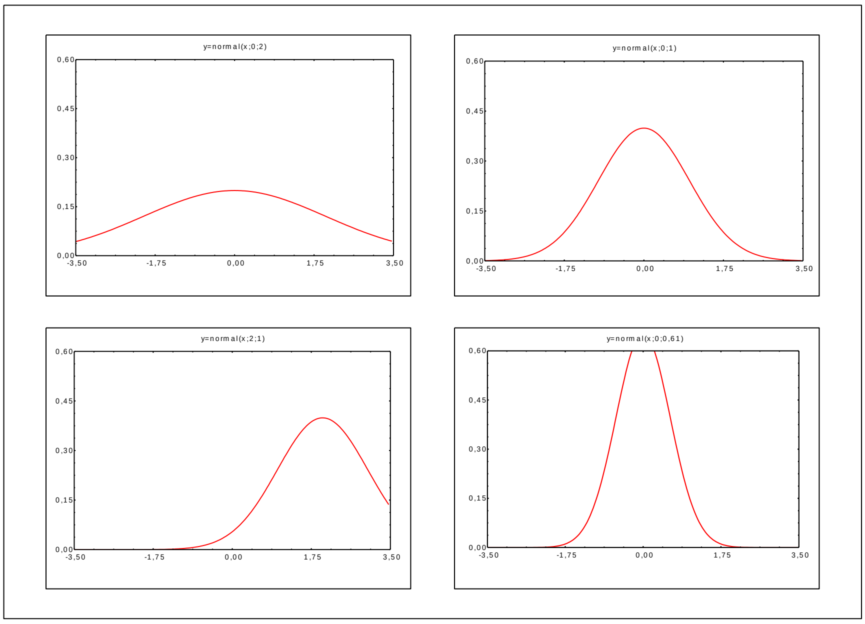
Dla zmiennej losowej ciągłej liczba wszystkich możliwych i wzajemnie się wykluczających zdarzeń elementarnych jest nieskończona i dlatego prawdopodobieństwo w punkcie odpowiadającym x_i równa się zero. Opis rozkładu zmiennej losowej ciągłej musi zatem przebiegać inaczej niż w przypadku dyskretnym. Najważniejszą rolę odgrywa tu pojęcie funkcji gęstości prawdopodobieństwa oznaczane przez f(x).

Wśród rozkładów dotyczącej zmiennej losowej ciągłej rozkład normalny zajmuje ważną pozycję. Wynika to z faktu, że większość zjawisk zależnych od wielu czynników jednocześnie (a taka sytuacja występuje często w medycynie) kształtuje się zgodnie z rozkładem normalnym. Rozkładowi normalnemu poświęcimy więcej uwagi, ponieważ wiele zagadnień statystycznych ma „prostsze rozwiązanie”, jeśli analizowana cecha ma rozkład normalny. Wiele analiz statystycznych i testów wymaga też spełnienia założenia o normalności rozważanej zmiennej (testy t-Studenta, analiza wariancji, analiza regresji, analiza kanoniczna). Dlatego często musimy przeprowadzić weryfikację charakteru rozkładu ilekroć chcemy zastosować analizy statystyczne wymagające danych o określonym rozkładzie.

Mówimy, że zmienna losowa X ma rozkład normalny o parametrach m i σ (co zapisujemy X~N(m, σ)), jeśli jej funkcja gęstości wyraża się wzorem:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} \quad , -\infty < x < \infty$$

Ze względu na swój kształt krzywa rozkładu normalnego nazywana bywa krzywą „dzwonową). Jest to rozkład symetryczny, którego kształt zależy od dwóch parametrów: μ i σ. Parametr μ to wartość średnia populacji, względem której rozkład jest symetryczny. W punkcie tym funkcja gęstości osiąga maksimum. Parametr σ to odchylenie standardowe, stanowiące miarę rozrzutu wokół średniej μ. Najczęściej nie znamy prawdziwych wartości tych parametrów, lecz oceniamy je na podstawie obliczeń średniej i odchylenia standardowego z próby. Parametr μ decyduje o położeniu krzywej względem osi poziomej (0x), natomiast od parametru σ zależy „wysmukłość” krzywej. Wpływ tych parametrów na położenie i postać krzywej normalnej ilustruje rysunek 6.



Rys. 6 Przykładowe rozkłady normalne o różnych parametrach

Rozkład normalny jest rozkładem, do którego zmierzają inne rozkłady zmiennej losowej, gdy liczebność próby rośnie do nieskończoności. Szczególny przypadek stanowi rozkład normalny ze średnią $\mu = 0$ oraz odchyleniem standardowym $\sigma = 1$ nazywany standaryzowanym rozkładem normalnym i oznaczamy przez N(0,1).

Testy zgodności

Testy te dotyczą postaci rozkładu teoretycznego badanej zmiennej losowej skokowej lub ciągłej. Weryfikują one stawiane przez badaczy hipotezy dotyczące kształtu rozkładu zmiennej losowej. Celem tych testów jest porównanie rozkładów dwóch cech w jednej populacji lub jednej cechy w dwóch populacjach. Są to oczywiście dwa różne zagadnienia, jednakże metody obliczeniowe są w obu przypadkach podobne. Idea tych testów jest oczywista – jeśli jakaś cecha w dwóch populacjach ma taki sam rozkład, to wartości liczbowe pewnych statystyk dla obu populacji powinny się niewiele różnić. Jeśli jednak wartości te będą istotnie różnie, to mamy prawo sądzić, że cecha ma odmienny rozkład w różnych populacjach.

W praktyce często potrzebujemy stwierdzić, jaki rozkład ma interesująca nas cecha w populacji generalnej. Najczęściej pytamy, czy badana cecha posiada rozkład normalny. Jest to bardzo często warunek stosowania omówionych dalej testów parametrycznych. Istnieją testy statystyczne, zwane testami normalności, które pozwalają sprawdzić, na podstawie rozkładu w próbie, czy populacja ma rozkład normalny. Wszystkie te testy weryfikują hipotezę zerową postaci:

H_0: Rozkład danej zmiennej jest rozkładem normalnym

względem

H_1: Rozkład danej zmiennej nie jest rozkładem normalnym.

Jak widzimy w hipotezie zerowej zakładamy, że n- elementowa próba losowa pochodzi ze zbiorowości generalnej, w której rozkład obserwowanej zmiennej losowej jest normalny. Zatem odrzucenie hipotezy zerowej, (jeśli $p < 0,05$) jest równoważne z tym, że dana zmienna (cecha) nie ma rozkładu normalnego.

Do najczęściej stosowanych testów weryfikujących normalność rozkładów zaliczamy testy:

- Test Kołmogorowa – Smirnowa

Test ten opiera się na porównaniu zaobserwowanych procentów skumulowanych z procentami oczekiwanymi. Jako wartość testu podawana jest wiec maksymalna różnica absolutna pomiędzy zaobserwowanymi i oczekiwanymi procentami skumulowanymi. Test ten wymaga jednak znajomości parametrów rozkładu (średniej i odchylenia standardowego całej populacji). Gdy tego nie znamy, a tak jest najczęściej stosujemy test Kołmogorowa – Smirnowa z poprawką Lillieforsa.

- Test W Shapiro – Wilka

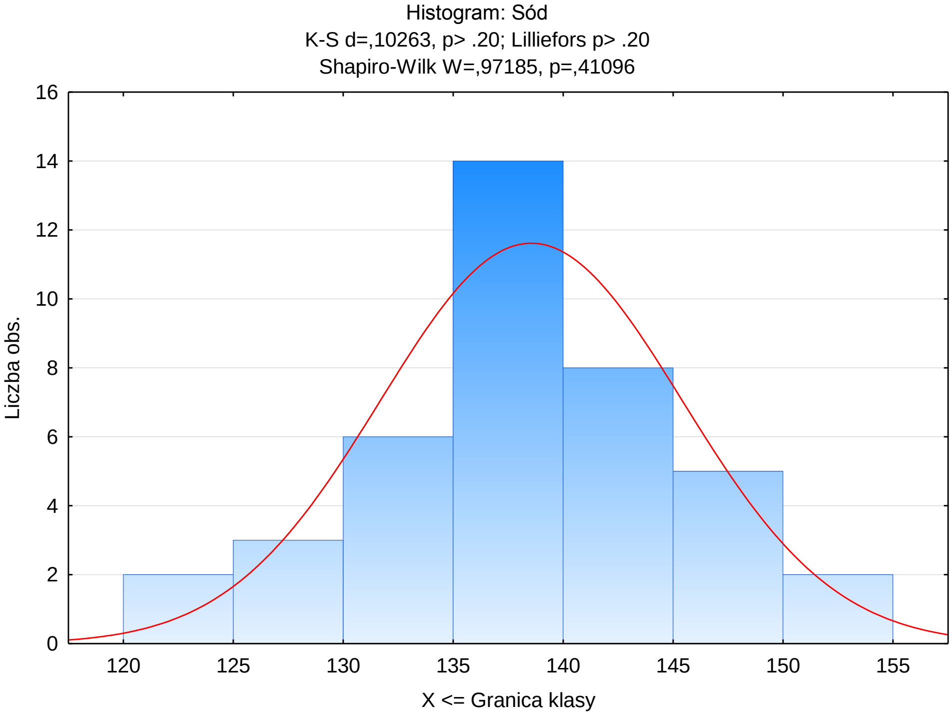
Test ten najbardziej polecany ze względu na jego dużą moc. Test ten możemy również stosować dla małych prób.

- Test χ^2 Pearsona

W celu weryfikacji hipotezy o normalności rozkładu wyniki próby dzielone są na rozłączne klasy a następnie porównywana jest liczebność obserwowana i oczekiwana w każdej z tych klas. Istotne różnice tych liczebności mówią nam, że prawdopodobnie dana próba nie pochodzi z populacji, w której rozkład obserwowanej zmiennej losowej jest normalny. Test ten wymaga próby o dużej liczebności (N>100).

Przykład 2

Sprawdzimy normalność rozkładu poziomu sodu w surowicy dla danych z naszej przykładowej bazie leczonych psów. W tym celu wykorzystamy pakiet statystyczny *STATISTICA*.

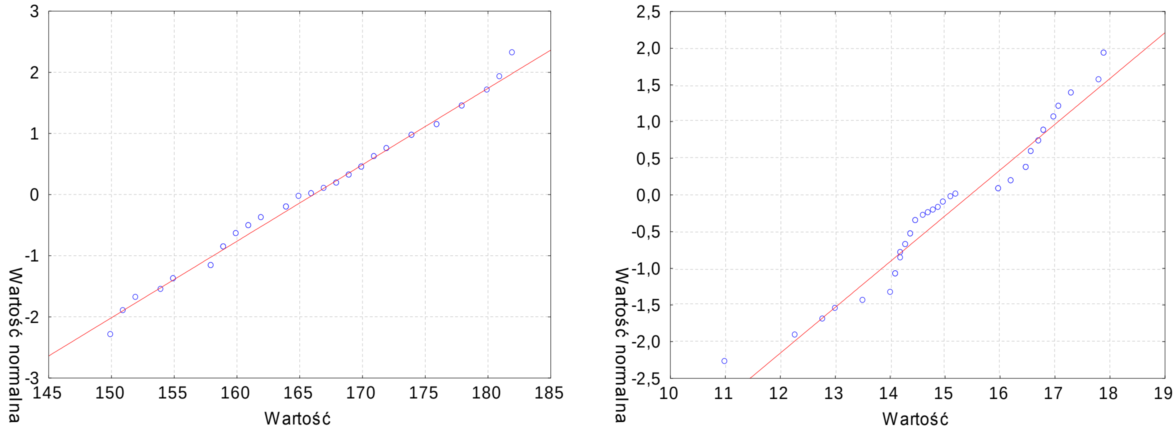


Rys. 7 Histogram z wynikami testów normalności

Otrzymujemy następujący histogram z wynikami testów normalności. Jak widzimy, nie mamy podstaw (p = 0,41096) do odrzucenia hipotezy zerowej. Możemy zatem powiedzieć, że rozkład zmiennej SÓD jest zbliżony do rozkładu normalnego.

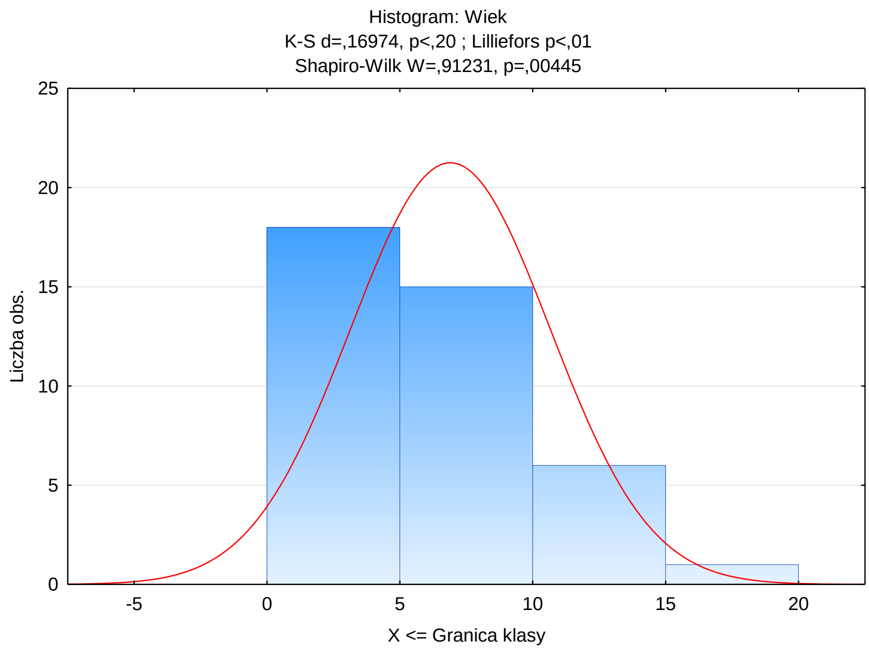
Mamy też możliwość „wzrokowej" oceny normalności rozkładu. Wykorzystujemy w tym celu wykres normalności. Na wykresie tym umieszczona jest prosta dopasowana do przedstawianych punktów metodą najmniejszych kwadratów. Im bardziej wszystkie punkty układają się na prostej, tym bardziej mamy prawo sądzić, że dany rozkład jest rozkładem normalnym.

Znaczące odchylenia od prostej wskazują, że analizowana zmienna podlega rozkładowi innemu niż normalny. Na wykresie takim wyraźnie widoczne są też ewentualne wartości odstające. Jeżeli mamy do czynienia z wyraźnym brakiem dopasowania punktów do prostej, a punkty układają się w jakiś prosty wzór (np. w kształt litery S), to zmienna może wymagać jakiegoś przekształcenia, zanim zostanie zastosowana w procedurach wymagających normalności. Poniższy rysunek pokazuje wykres normalności gdy zmienna ma rozkład normalny (po lewej stronie) i w przypadku braku normalności (po prawej stronie).



Rys. 8 Wykresy normalności

W kolejnym przykładzie sprawdzimy, czy cecha WIEK w naszej przykładowej bazie ma również rozkład normalny. Tym razem otrzymujemy następujący histogram z wynikami testów normalności.



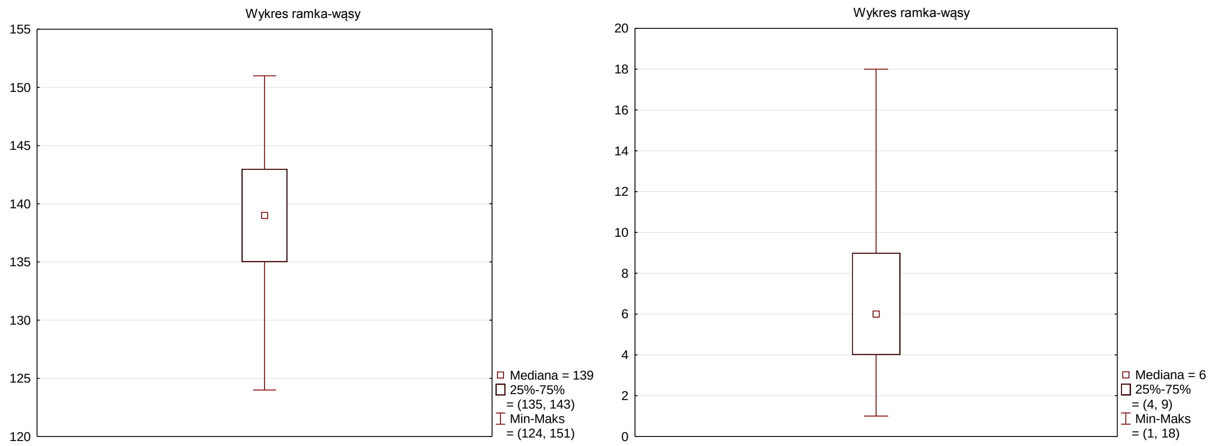
Rys. 9 Histogram dla cech WIEK z wynikami testów normalności

Tu sytuacja jest odmienna. Tym razem test okazał się istotny (p = 0,00445 - wartość mniejsza od 0,05), więc odrzucamy hipotezę zerową o normalności zmiennej WIEK. Histogram wskazuje na dużą asymetrię prawostronną.

Na zakończenie przypominamy jeszcze jeden popularny wykres. Są to wykresy ramka wąsy (*Box And Whiskers*) - wykres pokazujący zakresy wybranej zmiennej (zmiennych) oraz statystyki opisowe (średnia, mediana, odchylenie standardowe lub błąd standardowy). Na wykresie mogą również być wykreślone odstające punkty danych. Wykres ten wprowadził w 1977 roku J. Tukey. Do czasu pojawienia się pakietów statystycznych mało popularny. Dopiero szybkie komputery spopularyzowały takie prezentowanie statystyk pozycyjnych. W pakiecie *STATISTICA* możemy utworzyć następujące cztery grupy wykresów ramka-wąsy:

- punkt centralny - mediana, ramka-kwartyle, wąsy-rozstęp
- punkt centralny - średnia, ramka-błąd standardowy, wąsy-odchylenie standardowe
- punkt centralny - średnia, ramka-odchylenie standardowe, wąsy-1,96*odchylenie standardowej (95% przedział ufności dla poszczególnych obserwacji wokół średniej)
- punkt centralny - średnia, ramka-błąd standardowy, wąsy-1,96*błąd standardowy (95% przedział ufności dla wartości średniej)

Przykładowy wykres „skrzynka z wąsami" dla Zmiennych Wiek i Sód widoczny jest na poniższym rysunku.



Rys. 5 Wykres ramka –wąsy zmiennej Sód i Wiek

Widzimy różnicę w długości „wąsów" i niesymetryczne położenie mediany dla zmiennej Wiek (po prawej) wskazującą na dużą asymetrię prawostronną. Asymetria taka wyklucza oczywiście normalność. Rozkład normalny jest to bowiem rozkład idealnie symetryczny, dla którego współczynnik asymetrii jest równy zero.