TESTY STATYSTYCZNE W PROCESIE PODEJMOWANIA DECYZJI

CZESŁAW DOMAŃSKI
DOROTA PEKASIEWICZ
ALEKSANDRA BASZCZYŃSKA
ANNA WITASZCZYK

WYDAWNICTWO
UNIWERSYTETU
ŁÓDZKIEGO

TESTY STATYSTYCZNE W PROCESIE PODEJMOWANIA DECYZJI

WYDAWNICTWO
UNIWERSYTETU
ŁÓDZKIEGO

TESTY STATYSTYCZNE W PROCESIE PODEJMOWANIA DECYZJI

CZESŁAW DOMAŃSKI

DOROTA PEKASIEWICZ

ALEKSANDRA BASZCZYŃSKA

ANNA WITASZCZYK

WYDAWNICTWO
UNIWERSYTETU
ŁÓDZKIEGO

ŁÓDŹ 2014

Czesław Domański, Dorota Pekasiewicz, Aleksandra Baszczyńska, Anna Witaszczyk
Uniwersytet Łódzki, Wydział Ekonomiczno-Socjologiczny, Katedra Metod Statystycznych
90-214 Łódź, ul. Rewolucji 1905 r. nr 41/43

RECENZENT
Mirosław Szreder

REDAKTOR WYDAWNICTWA UŁ
Iwona Gos

SKŁAD KOMPUTEROWY
AGENT PR

PROJEKT OKŁADKI
Stämpfli Polska Sp. z o.o.

Zdjęcie na okładce: © Shutterstock.com

Praca naukowa finansowana ze środków Narodowego Centrum Nauki przyznanych na podstawie decyzji numer DEC-2011/01/B/HS4/02746

Wydane przez Wydawnictwo Uniwersytetu Łódzkiego
Wydanie I. W.06577.14.0.K

ISBN (wersja drukowana) 978-83-7969-358-0
ISBN (ebook) 978-83-7969-763-2

Wydawnictwo Uniwersytetu Łódzkiego
90-131 Łódź, ul. Lindleya 8
www.wydawnictwo.uni.lodz.pl
e-mail: ksiegarnia@uni.lodz.pl
tel. (42) 665 58 63, faks (42) 665 58 62

SPIS TREŚCI

Przedmowa 7

1. Testy statystyczne i decyzje statystyczne [*Czesław Domański*] 11
1.1. Uwagi ogólne i podstawowe pojęcia 11
1.2. Weryfikacja hipotez statystycznych 13
1.3. Statystyczne problemy decyzyjne 26
1.4. Uwagi o testach statystycznych wykorzystujących próby z brakującą informacją 28

2. Wybrane klasyczne testy statystyczne [*Czesław Domański*] 37
2.1. Uwagi wstępne 37
2.2. Testy dla jednej zmiennej 37
2.3. Testy dla dwóch i więcej zmiennych 42
2.4. Analiza wariancji (ANOVA) 55
2.5. Wielowymiarowa analiza wariancji (MANOVA) 58
2.6. Wybrane testy zgodności dla rozkładów dochodów 62

3. Testy statystyczne w porównaniach wielokrotnych i modelach symulacyjnych [*Czesław Domański*] 67
3.1. Uwagi wstępne 67
3.2. Klasyfikacja porównań wielokrotnych 67
3.3. Wielokrotne procedury decyzyjne 71
3.4. Podejście Neymana-Pearsona 74
3.5. Porównania wielokrotne 77
3.6. Weryfikacja hipotez dla modeli symulacyjnych 79

4. Bayesowskie testy statystyczne [*Dorota Pekasiewicz*] 87
4.1. Uwagi wstępne 87
4.2. Idea konstrukcji testów bayesowskich 87
4.3. Rozkłady *a priori* parametrów zmiennych losowych i zasady ich określania 93
4.4. Testy bayesowskie przy niezależnym schemacie losowania próby 94
4.5. Bayesowska weryfikacja hipotez statystycznych przy zależnym schemacie losowania próby 101
4.6. Analiza własności bayesowskich procedur testowych 104
4.7. Przykłady zastosowań testów bayesowskich 110

5. Bootstrapowe testy statystyczne [*Dorota Pekasiewicz*] 119
5.1. Uwagi wstępne 119
5.2. Istota konstrukcji bootstrapowych testów statystycznych 120
5.3. Nieparametryczne testy bootstrapowe 122
5.4. Parametryczne i semiparametryczne testy bootstrapowe 125
5.5. Testy bootstrapowe dla hipotez o wartościach średnich populacji 126
5.6. Analiza własności wybranych testów bootstrapowych 135

6. Sekwencyjne testy statystyczne [*Dorota Pekasiewicz*] ... 141
6.1. Uwagi wstępne ... 141
6.2. Idea konstrukcji ilorazowego testu sekwencyjnego ... 141
6.3. Ilorazowe testy sekwencyjne przy niezależnym schemacie losowania próby ... 148
6.4. Ilorazowe testy sekwencyjne dla schematów losowania próby innych niż losowanie niezależne ... 160
6.5. Nieparametryczne testy sekwencyjne ... 165

7. Testy statystyczne oparte na metodzie jądrowej [*Aleksandra Baszczyńska*] ... 171
7.1. Uwagi wstępne ... 171
7.2. Metoda jądrowa ... 171
7.3. Jądrowe testy zgodności, niezależności i symetryczności ... 176
7.4. Jądrowe testy w analizie regresji ... 187
7.5. Jądrowe testy w badaniu obserwacji nietypowych ... 198

8. Testy statystyczne dotyczące rozkładów wielowymiarowych [*Anna Witaszczyk*] ... 205
8.1. Uwagi wstępne ... 205
8.2. Macierze losowe i przekształcenie Stieltjesa ... 205
8.3. Wybrane twierdzenia graniczne dla macierzy losowych ... 212
8.4. Testy dla wektorów wartości oczekiwanych ... 217
8.5. Weryfikacja hipotez dotyczących macierzy kowariancji ... 224
8.6. Testy wielowymiarowej normalności ... 236

9. Testy statystyczne dla danych cenzurowanych [*Aleksandra Baszczyńska*] ... 245
9.1. Uwagi wstępne ... 245
9.2. Podstawowe pojęcia ... 245
9.3. Testy zgodności dla dwóch lub więcej populacji dla danych cenzurowanych ... 249
9.4. Testy zgodności z rozkładem teoretycznym dla danych cenzurowanych ... 255
9.5. Testy w analizie regresji dla danych cenzurowanych ... 258

10. Weryfikacja hipotez statystycznych dla szeregów czasowych [*Czesław Domański*] ... 263
10.1. Uwagi wstępne i podstawowe pojęcia ... 263
10.2. Testy pierwiastka jednostkowego ... 265
10.3. Testy szczytów ... 268
10.4. Weryfikacja parametrów modeli ARMA ... 277
10.5. Weryfikacja parametrów modeli VAR ... 282

Zakończenie ... 289

Statistical Tests in the Decision Making Process (Summary) ... 291

Literatura ... 293

Tablice statystyczne wybranych rozkładów prawdopodobieństwa ... 299

Wybrane oznaczenia ... 313

Indeks ... 317

PRZEDMOWA

Ukazanie się książki Johna Graunta *Naturalne i polityczne obserwacje poczynione na biuletynach śmiertelności* w 1662 r. to moment, od którego zauważalny jest rozwój statystyki. Jednak rodowód statystyki, podobnie jak matematyki, sięga odległej starożytności. Już w XXXII w. p.n.e. plemię Ashipu, mieszkające między Eufratem a Tygrysem, zajmowało się udzielaniem konsultacji w zakresie ryzyka i niepewności oraz podejmowania trudnych decyzji.

Statystyka została wyodrębniona w oddzielną dyscyplinę jako metoda wydobywania informacji z zaobserwowanych danych oraz jako logika podejmowania decyzji w warunkach niepewności.

Wiedza statystyczna jest cenna dla przedstawicieli wszystkich zawodów. Wiele złożonych problemów naszego życia wyglądałoby prościej, gdyby przed podjęciem działań najpierw stawiać pytania, a następnie uzyskiwać właściwe informacje. Formułowanie pytań uważa się często za kłopotliwe, gdyż wymaga analizy, myślenia i precyzowania wniosków. Działania takie zabierają nam czas i energię. Mogą też prowadzić do niepożądanej dezorientacji i zdenerwowania. W wielu przypadkach, aby uniknąć takich sytuacji, opieramy się na mądrości innych lub działamy emocjonalnie, co może prowadzić do nieporozumień, złego wyboru momentu działania i pomyłek. Porady mogą być pomocne, ale raczej jako punkt odniesienia, a nie samowystarczalne podejście.

W dzisiejszym świecie, jak nigdy wcześniej, istnieje potrzeba myślenia statystycznego. Jesteśmy otoczeni wyzwaniami różnorodnych banków danych (chociaż rzadko zgodnych z oczekiwaniami), które wymagają coraz lepszych metod statystycznych, algorytmów, modeli systemów przetwarzania.

Statystyka zajmuje się kolekcjonowaniem informacji liczbowych oraz ich analizą i interpretacją. Prezentowane w opracowaniu metody pozwalają odpowiedzieć na pytanie, co te informacje liczbowe, które traktujemy jako dane, mówią nam o populacji i o zjawiskach, których dotyczą. Odpowiedź zależeć będzie nie tylko od samych informacji, tzn. od obserwacji, ale również od wiedzy *a priori*. Ta wiedza jest formalizowana za pomocą założeń przy konstrukcji metod. Rozróżniane są najczęściej trzy podejścia oparte na różnych zasadach. Należą do nich:

– analiza danych,
– klasyczne wnioskowanie i teoria decyzji,
– analiza bayesowska.

W pierwszym podejściu informacje statystyczne są analizowane jako dane, w istocie rzeczy bez żadnych dodatkowych założeń. Głównym celem jest ich obróbka i prezentacja graficzna lub tabelaryczna umożliwiająca wykrycie najważniejszych własności i wyjaśnienie struktur danych.

W drugim podejściu obserwowane dane są traktowane jako wartości przyjęte przez zmienne losowe, dla których przyjmuje się, że mają pewien łączny rozkład P z klasy $\mathcal{P}$. Często rozważane rozkłady indeksowane są parametrem θ lub Θ.

W analizie bayesowskiej zakłada się dodatkowo, że sam parametr jest zmienną losową o pewnym znanym rozkładzie. Ten rozkład, zwany rozkładem *a priori*, zdefiniowany przed zapoznaniem się z danymi, jest modyfikowany za pomocą danych do rozkładu *a posteriori* parametru θ pod warunkiem zaobserwowanych danych. Rozkład *a posteriori* w pewnym sensie syntetyzuje to, co można powiedzieć o parametrze θ na podstawie danych i wiedzy wstępnej *a priori*.

Wspomniane trzy podejścia pozwalają na formułowanie coraz mocniejszych wniosków, a zarazem mniej pewnych założeń. Często pożądane jest korzystanie z kombinacji tych różnych podejść, np. planując badanie, uwzględnia się wybór liczebności próby przy bardziej szczegółowych założeniach i przeprowadza analizę wyników przy słabszych, ale za to bardziej przekonujących założeniach. W niektórych zastosowaniach często pożyteczne jest formułowanie różnych modeli do danego problemu. Wówczas zgodność wniosków daje dodatkowy argument na rzecz poprawności analizy i odwrotnie, rozbieżności we wnioskach wskazują na konieczność dokładniejszego przyjrzenia się założeniom różnych modeli.

Problemy statystyczne charakteryzują się tym, że mamy w nich do czynienia nie z pojedynczymi rozkładami prawdopodobieństwa, ale z rodzinami $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ rozkładów określonych na pewnej wspólnej przestrzeni mierzalnej (χ, A).

Zasadniczym materiałem badań statystycznych jest zbiór wyników obserwacji, będących wartościami zmiennej losowej X, której rozkład P_θ jest przynajmniej częściowo znany. Przyjmujemy, że o parametrze θ wiemy tylko tyle, że należy on do pewnego zbioru Θ, zwanego przestrzenią parametrów.

Potrzeba analizy statystycznej wynika z faktu, że rozkład zmiennej losowej X, a zatem pewne elementy sytuacji stanowiącej podstawę modelu matematycznego nie są znane, co powoduje trudności w wyborze najlepszego postępowania.

Książka przedstawia w zwartej formie różne klasy testów statystycznych, które mogą być stosowane w procesie podejmowania decyzji dotyczących zjawisk ekonomicznych, społecznych, demograficznych, technicznych i medycznych. Klasyczne procedury testowe prezentowane w literaturze przedmiotu nie zawsze można wykorzystać ze względu na założenia ich stosowalności. Dotyczyć to może niespełnienia określonych założeń o rozkładzie zmiennych losowych, z którymi utożsamiane są badane cechy statystyczne, braku dostatecznej liczby elementów próby lub też stosowanego w badaniu schematu losowania próby, odmiennego od losowania niezależnego.

Rozważane grupy testów charakteryzują się odmiennymi procedurami testowymi, np. przy zastosowaniu testów bayesowskich parametr rozkładu zmiennej losowej jest traktowany jako zmienna losowa, natomiast w testach sekwencyjnych liczebność próby jest zmienną losową. W testach jądrowych można wykorzystywać różne funkcje jądra i parametry wygładzania, co wpływa w znacznym stopniu na rezultaty zastosowanej procedury, natomiast w testach bootstrapowych procedura wnioskowania jest oparta na tzw. próbach bootstrapowych. Oprócz rozważań teoretycznych zaprezentowane są również wyniki przeprowadzonych badań, dotyczących własności analizowanych procedur weryfikacji hipotez statystycznych wraz ze wskazaniem obszarów ich zastosowań.

Praca składa się z dziesięciu rozdziałów. Punktem wyjścia do rozważań dotyczących testów statystycznych opisywanych w dalszej części książki są trzy pierwsze rozdziały. Obejmują one zagadnienia związane z klasycznym i teoriodecyzyjnym podejściem do weryfikacji hipotez statystycznych. Związek między testami statystycznymi a podejmowaniem decyzji zaprezentowany jest w rozdziale pierwszym i trzecim. Rozdział drugi przedstawia wybrane klasyczne testy statystyczne z uwzględnieniem warunków, które muszą być spełnione, by dany test mógł być stosowany w praktyce.

W kolejnym rozdziale prezentowane są testy bayesowskie charakteryzujące się tym, że parametr rozkładu jest traktowany jako zmienna losowa o znanym rozkładzie *a priori*. Stosując je, podejmujemy decyzję o akceptacji hipotezy o mniejszym ryzyku *a posteriori*, które wyznacza się na podstawie rozkładu *a priori* i ustalonej funkcji straty. Rozważane testy bayesowskie dotyczą weryfikacji hipotez statystycznych o parametrach rozkładu zmiennych losowych i wskazują na możliwość zastosowania różnych schematów losowania próby.

Testy bootstrapowe, którym poświęcony jest piąty rozdział książki, zasługują na uwagę, ponieważ nie wymagają informacji o klasie rozkładu badanej zmiennej losowej. Zastosowanie metod bootstrapowych do aproksymacji rozkładów statystyk testowych pozwala na weryfikację hipotez o parametrach rozkładu populacji w oparciu o małe próby, co jest dużą zaletą tych metod.

Testy sekwencyjne rozważane w rozdziale szóstym to kolejna grupa testów nieklasycznych. W testach tych liczebność próby jest zmienną losową. Sekwencyjne zwiększanie liczby elementów próby losowej pozwala podjąć decyzję o akceptacji jednej z weryfikowanych hipotez z przyjętymi prawdopodobieństwami błędów I i II rodzaju. Zaletą stosowania testów należących do tej klasy jest nawet dwukrotnie mniejsza wartość oczekiwana liczebności próby niezbędnej do podjęcia decyzji w porównaniu z testami klasycznymi dla identycznych błędów I i II rodzaju, co wpływa na koszt przeprowadzanego badania statystycznego.

W rozdziale siódmym przedmiotem rozważań jest klasa testów jądrowych. Metoda jądrowa, wywodząca się z estymacji funkcji gęstości, stanowi typowo nieparametryczne podejście w procedurach wnioskowania statystycznego. W rozdziale tym rozważane są procedury weryfikacji hipotez dotyczących rozkładu

zmiennej losowej, w tym: normalności, zgodności dwóch i więcej rozkładów, hipotez o postaci funkcji regresji i hipotez mówiących o niezależności zmiennych losowych.

Rozdział ósmy poświęcony jest podejściu wielowymiarowemu w weryfikacji hipotez statystycznych. Analizie poddane są testy służące do weryfikacji hipotez o wektorach wartości oczekiwanych oraz hipotez dotyczących macierzy kowariancji, zarówno klasyczne, jak i konstruowane w oparciu o twierdzenia graniczne teorii macierzy losowych.

Rozdział dziewiąty dotyczy procedur wnioskowania statystycznego stosowanych w sytuacji, gdy dane mają charakter przekrojowo-czasowy i brak jest informacji dla pewnych okresów lub momentów czasu. W rozdziale tym przedstawione są najważniejsze klasy testów dla danych cenzurowanych, m.in. testy dotyczące zgodności rozkładów dwóch lub więcej populacji oraz testy zgodności rozkładu badanej populacji z rozkładem hipotetycznym.

Specjalna grupa testów stosowanych w analizach szeregach czasowych jest przedmiotem rozważań w rozdziale dziesiątym, ze szczególnym uwzględnieniem analizy stacjonarności i niestacjonarności procesu stochastycznego oraz weryfikacji parametrów modeli VAR i ARMA.

Serdecznie dziękuję wszystkim tym, których życzliwe uwagi przyczyniły się do udoskonalenia tej książki, przede wszystkim Panu Profesorowi Mirosławowi Szrederowi za wnikliwą recenzję.

Czesław Domański

1. TESTY STATYSTYCZNE I DECYZJE STATYSTYCZNE

1.1. Uwagi ogólne i podstawowe pojęcia

Zasadniczym materiałem badań statystycznych jest zbiór wyników obserwacji. Obserwacje są podstawowym źródłem wiedzy o otaczającym świecie. Wiedzę dotyczącą każdego zjawiska można „magazynować" w postaci modeli tego zjawiska. Modelem nazywamy sformalizowane ujęcie pewnej teorii lub sytuacji przyczynowej, o której zakładamy, że generuje obserwowane dane.

Każdą analizę statystyczną pewnego rzeczywistego zjawiska musimy oprzeć na jego modelu matematycznym (tj. modelu wyrażonym w postaci zależności matematycznych), w którym uwzględniony został sposób pozyskania obserwacji. Dążyć należy do tego, aby stosowany model stanowił oszczędny opis natury. Oznacza to, że postać funkcyjna modelu powinna być prosta, a liczba jego parametrów i składników – jak najmniejsza.

Łatwo zauważyć, że nie istnieją modele doskonałe, czyli w idealny sposób odwzorowujące zachowanie modelowanego obiektu. Każda nowa obserwacja oraz analiza niezgodności modelu matematycznego i rzeczywistego obiektu prowadzą do nowych, bardziej dokładnych, modeli matematycznych. Jako główną przyczynę braku zgodności pomiędzy modelem a modelowanym zjawiskiem należy wymienić:

1) aktualny stan wiedzy o badanym zjawisku;

2) wysoki stopień zależności modelowanego zjawiska, który uniemożliwia zastosowanie modelu matematycznego ujmującego wszystkie cechy tego obiektu;

3) rozmaitość i zmienność wpływów środowiska, którym podlega obiekt, co sprawia, że modelowanie rzeczywistych przyczyn stanu obiektu staje się niemożliwe;

4) barierą złożoności modelu bywa także koszt związany z jego wykorzystaniem. Może się zdarzyć, że model prostszy, choć mniej dokładny, okaże się lepszy, bowiem zysk związany z rezygnacją ze skomplikowanych pomiarów często przewyższa straty wynikające ze stosowania modelu mniej dokładnego.

Modele matematyczne można podzielić na trzy klasy: modele deterministyczne, modele deterministyczne z prostymi wielkościami losowymi i modele stochastyczne.

Modele deterministyczne – każda obserwacja jest tu wartością pewnej funkcji parametrów tego modelu oraz funkcją takich wielkości, jak czas, położenie w przeszłości czy wielkość pewnego bodźca. Innymi słowy, model deterministyczny nie zawiera elementów losowych, a przyszłość systemu jest zdeterminowana przez jego pozycję, prędkość itp. w pewnym ustalonym momencie.

Modele deterministyczne z prostymi wielkościami losowymi – każda obserwacja jest pewną funkcją wielkości deterministycznych, a także wielkości losowych, które są związane z błędami pomiarów, z obserwacjami, z wielkościami początkowymi oraz ze zmiennością próbkową. Przyjmuje się tu założenie o niezależności składników losowych różnych obserwacji.

Modele stochastyczne – zbudowane na bazie pewnych zdarzeń losowych lub wielkości losowych. Takie modele pozwalają opisać zjawiska dynamiczne lub ewolucyjne: od schematu Bernoulliego (matematyczny model rzutu monetą) do procesu urodzin i śmierci (matematyczny model wielkości populacji biologicznej). W modelach stochastycznych każda obserwacja może zależeć w pewnym stopniu od obserwacji poprzedzających ją w czasie lub sąsiadujących z nią w przestrzeni.

Punktem wyjścia w naszych rozważaniach będzie zawsze pewien element losowy X (zmienna losowa, skończony lub nieskończony ciąg zmiennych losowych). Będziemy często nazywali go wynikiem eksperymentu, wynikiem pomiaru, wynikiem obserwacji lub po prostu obserwacją. Zbiór wszystkich wartości elementu losowego X jest przestrzenią próby oznaczoną przez χ. Przestrzeń χ będzie zbiorem skończonym lub przeliczalnym, albo pewnym obszarem w skończenie wymiarowej przestrzeni R^n.

Niech Ω będzie zbiorem zdarzeń elementarnych i niech $\aleph$ będzie σ-ciałem podzbiorów zbioru Ω. Trójkę uporządkowaną $(\Omega, \aleph, P)$ nazywamy **przestrzenią probabilistyczną**, gdzie P oznacza prawdopodobieństwo.

Niech A będzie wyróżnionym σ-ciałem podzbiorów zbioru $X \subset R^n$, zaś X jest mierzalnym przekształceniem $(\Omega, \aleph) \to (\chi, A)$. Rozkład $P^X(A) = P(X^{-1}(A))$ jest miarą na przestrzeni (χ, A). W problemach statystycznych zakłada się, że rozkład P należy do pewnej określonej klasy rozkładów $\mathcal{P}$ na (χ, A). Znając tę klasę oraz mając dane wyniki obserwacji zmiennej losowej X, chcemy wysnuć poprawne wnioski o nieznanym rozkładzie. Wobec tego matematyczną podstawą badań statystycznych jest przestrzeń mierzalna (χ, A) i rodzina rozkładów $\mathcal{P}$. Przestrzeń probabilistyczna $(\Omega, \aleph, P)$ odgrywa rolę pomocniczą. Sformułowanie: dana jest przestrzeń probabilistyczna $(\Omega, \aleph, P)$, oznacza, że znany jest model probabilistyczny pewnego zjawiska lub doświadczenia, czyli wiemy, jakie są możliwe wyniki tego doświadczenia, jakie zdarzenia wyróżniamy oraz jakie prawdopodobieństwa tym zdarzeniom przypisujemy. Reasumując, wiedza *a priori* o przedmiocie badań jest sformułowana w postaci pewnych modeli probabilistycznych. Probabilistyka może wynikać z samego charakteru badanego zjawiska lub też być wprowadzana przez badacza.

Zauważmy, że $\mathcal{P} = \{P_\theta: \theta \in \Theta\}$ jest rodziną rozkładów prawdopodobieństwa na odpowiednim σ-ciele zdarzeń losowych w χ.

Przestrzeń próby wraz z rodziną rozkładów $\mathcal{P}$, tzn. obiekt:

$$(\chi, \{P_\theta: \theta \in \Theta\}) \tag{1.1}$$

nazywamy **modelem statystycznym** (przestrzenią statystyczną), natomiast odwzorowania z χ w R^k – statystykami lub k-wymiarowymi statystykami.

Jeżeli $\mathbf{X} = (X_1, X_2, ..., X_n)^T$, przy czym $X_1, X_2, ..., X_n$ są niezależnymi zmiennymi losowymi o jednakowym rozkładzie, to będziemy stosować też oznaczenie:

$$(\chi, \{P_\theta: \theta \in \Theta\})^n, \tag{1.2}$$

w którym χ jest zbiorem wartości zmiennej losowej X (a więc każdej ze zmiennych $X_1, X_2, ..., X_n$) oraz P_θ to rozkład tej zmiennej losowej. Używa się wtedy również terminologii: $X_1, X_2, ..., X_n$ jest próbą z rozkładu P_θ lub próbą z populacji P_θ dla pewnego $\theta \in \Theta$.

Będziemy zawsze zakładali, że jeżeli $\theta_1 \neq \theta_2$, to $P_{\theta_1} \neq P_{\theta_2}$. Takie modele określa się jako identyfikowalne: znając rozkład P_θ, znamy wartość parametru θ). Wprowadzenie parametru θ do rozważań ułatwia sformułowanie wielu problemów, a dopóki nie wprowadzamy ograniczeń na zbiór Θ, odbywa się to bez straty ogólności rozważań, bo każdą rodzinę $\mathcal{P}$ rozkładów prawdopodobieństwa możemy „sparametryzować", przyjmując za parametr θ rozkładu P ten sam rozkład P.

Modele statystyczne możemy podzielić na parametryczne i nieparametryczne.

Parametryczny model statystyczny powstaje wówczas, gdy Θ jest przestrzenią skończenie wymiarową.

Nieparametrycznym modelem statystycznym nazywamy z kolei taki model, w którym nie istnieje skończenie wymiarowa parametryzacja rodziny rozkładów prawdopodobieństwa, czyli parametryzacja za pomocą pewnego $\theta \in \Theta \subset R^k$, dla $k \in N$.

1.2. Weryfikacja hipotez statystycznych

Przypomnijmy podstawowe pojęcia dotyczące weryfikacji hipotez statystycznych.

Populacją generalną nazywamy zbiór elementów powiązanych ze sobą logicznie i jednocześnie nieidentycznych ze względu na badaną cechę.

Próba jest to podzbiór populacji podlegający bezpośrednio obserwacji w celu zbadania własności całej populacji.

Próba losowa to taka próba, którą otrzymaliśmy w drodze losowania, tzn. jedynie przypadek decyduje o tym, który element populacji generalnej wejdzie do próby, a który nie.

Innymi słowy, przez losowy dobór próby rozumiemy taki sposób pobierania próby, który spełnia dwa następujące warunki (por. np. Szreder [2004]):

1) każda jednostka populacji ma dodatnie i znane prawdopodobieństwo dostania się do próby;

2) dla każdego zespołu jednostek populacji można ustalić prawdopodobieństwo tego, że w całości znajdzie się on w próbie.

Próbą prostą *n*-elementową nazywamy próbę wylosowaną z populacji w taki sposób, że przed jej pobraniem każdy podzbiór składający się z *n*-elementów populacji będzie mieć jednakowe prawdopodobieństwo wylosowania.

Rozkładem populacji nazywamy rozkład badanej zmiennej w tej populacji. Modelem matematycznym rozkładu populacji jest rozkład prawdopodobieństwa pewnej zmiennej losowej skokowej lub ciągłej. Odpowiednie prawdopodobieństwa interpretujemy jako częstość względną występowania w populacji elementów o określonych wartościach badanej cechy. Rozważamy jedynie badania częściowe, tzn. takie, które umożliwiają ocenę rozkładu populacji na podstawie pobieranej z niej próby statystycznej. Uwzględniając reprezentacyjny charakter próby, możemy uogólnić jej wyniki na całą populację, gdyż dopuszczamy jedynie losowy dobór próby. Losowość próby umożliwia bowiem otrzymywanie prób reprezentatywnych, tzn. charakteryzujących się rozkładem badanej zmiennej nieistotnie różniącym się od rozkładu zbiorowości. Ponadto, daje podstawę do wnioskowania o populacji na gruncie rachunku prawdopodobieństwa, pozwalającą ocenić dokładność wnioskowania.

Próbę losową ze skończonych populacji otrzymuje się drogą odpowiedniego losowania elementów tej populacji, natomiast z populacji nieskończonych, np. w badaniach przyrodniczych bądź technicznych, uzyskuje się drogą obserwacji niezależnych powtórzeń wykonywanych w określonych warunkach uwzględniających różne czynniki wpływające na wyniki eksperymentu. W przypadku populacji skończonych korzysta się często przy losowaniu elementów do próby z tablic liczb losowych bądź generatorów liczb losowych.

W praktyce zasadniczym kryterium doboru próby są tzw. tablice liczb losowych. Zbudowane są one z kolumn i wierszy liczb dwu- cztero- lub sześciocyfrowych, występujących po sobie w sposób przypadkowy. Procedura korzystania z tych tablic polega na tym, że:

1) wszystkim elementom zbiorowości N-elementowej przyporządkowuje się numery od 1 do *N*;

2) poczynając od dowolnie wybranej liczby z tablic liczb losowych, otrzymujemy *n* numerów, tzn. tyle, ile elementów ma być wylosowanych do próby. Jeżeli raz odczytany numer uwzględnimy jeszcze tyle razy, ile razy natrafimy na niego przy dalszym czytaniu liczb losowych, to wówczas otrzymujemy próbę prostą. Postępowanie takie nazywamy **losowaniem niezależnym** lub ze **zwracaniem**.

Przez losowanie lub wybór przypadkowy będziemy zawsze rozumieć losowanie zgodne z rozkładem jednostajnym. Stąd wynika, że skład próby jest przypadkowy, a więc i wartości badanej cechy wylosowanych elementów są przy-