Etap trzeci realizacji procesu analizy danych statystycznych w zasadzie powinien rozwiązać nasz zasadniczy problem związany z identyfikacją cechy populacji generalnej w języku teorii prawdopodobieństwa. Powinniśmy wiedzieć, jakiego typu jest to rozkład i dysponować oszacowaniem jego parametrów.

Statystyka dysponuje jednak jeszcze jednym narzędziem, którego rola polega między innymi na weryfikacji wiedzy, jaką pozyskaliśmy stosując zaprezentowane wyżej metody 2 i 3. Narzędziem tym jest *teoria testów statystycznych*.

Niech dane będzie populacja generalna będąca przedmiotem obserwacji z punktu widzenia cechy $\mathbb{X}$. Przez *hipotezę statystyczną* będziemy rozumieli każde przypuszczenie odnoszące się do rozkładu tej cechy formułowane w kategoriach bądź jakościowych–dotyczących postaci tego rozkładu, bądź ilościowych–związanych z jego parametrami. Będziemy wtedy mówili, że mamy do czynienia odpowiednio z *hipotezą nieparametryczną* albo *hipotezą parametryczną*. Podstawą do formułowania przypuszczeń na temat cechy populacji generalnej jest próba prosta. Proces weryfikacji sformułowanej hipotezy nazywamy *testem statystycznym*. Stąd też procedura testowania hipotez dzieli się na: *testy parametryczne* i *nieparametryczne*. U podstaw każdego testu znajduje się założenie, że obok hipotezy weryfikowanej, zwanej też *hipotezą zerową* i oznaczaną przez H_o, formułuje się hipotezę będącą w swojej treści wynikiem zaprzeczenia hipotezy H_o, którą nazywamy z tego powodu *hipotezą alternatywną* i oznaczamy przez H_1. Wtedy, w przypadku odrzucenia hipotezy H_o w wyniku przeprowadzonego testu, za prawdziwą uznaje się hipotezę H_1.

Jak zaznaczyliśmy wyżej, podstawą do przeprowadzenia weryfikacji hipotezy H_o przeciwko hipotezie H_1 jest próba prosta. Z teoretycznego punktu widzenia istnieje więc możliwość popełnienia błędu. Mogą zdarzyć się co najmniej dwie sytuacje: H_o jest prawdziwa albo H_o jest fałszywa. Jeśli w sytuacji pierwszej odrzucimy H_o (a więc przyjmiemy H_1), to mówimy wtedy o popełnionym *błędzie I–rodzaju*, a jego miarą jest wtedy pewna liczba α. W sytuacji drugiej, jeśli H_o przyjmujemy (a więc odrzucamy H_1), to mówimy o *błędzie II–rodzaju* i jego miarę oznaczamy przez β. Aby zrozumieć znaczenie obu zdefiniowanych błędów musimy dokładniej opisać przebieg procesu testowania hipotez.

Niech

$$(x_1, \dots, x_n) = (\mathbb{X}_1, \dots, \mathbb{X}_n)(\omega_o)$$

dla pewnego zdarzenia elementarnego ω_o oznacza próbę prostą naszej populacji generalnej. Wtedy

$$(x_1, \dots, x_n) \in W = (\mathbb{X}_1, \dots, \mathbb{X}_n)(\Omega) \subset \mathbf{R}^n,$$

co oznacza, że próba prosta jest jedną z wielu realizacji procesu obserwacji populacji generalnej, którą opisuje tutaj mnogość W. Zadaniem testu statystycznego jest skonstruowanie takiego podzbioru $Q \subset W$, że jeśli $(x_1, \dots, x_n) \in Q$, to hipotezę H_o odrzucimy na korzyść hipotezy H_1 oraz w przypadku gdy $(x_1, \dots, x_n) \in W \backslash Q$, to H_o przyjmujemy. Z tego względu zbiór Q nazywamy *obszarem krytycznym* (mówimy też o obszarze odrzucenia), w przeciwieństwie do zbioru $W \setminus Q$ zwanego *obszarem przyjęcia* hipotezy H_o. Wtedy znaczenie ilościowe zdefiniowanych wyżej błędów α i β jest następujące:

$$\alpha = P((x_1, \dots, x_n) \in Q | H_o), \ \ \beta = P((x_1, \dots, x_n) \in W \setminus Q | H_1).$$

Z metodologicznego punktu widzenia pierwszą z powyższych równości traktuje się jako równanie z niewiadomą Q, przyjmując, że α jest małą liczbą równą zazwyczaj $0,1$ albo $0,05$. Nazywamy ją wtedy *poziomem istotności* (wtedy $1-\alpha$ nazywa się *poziomem ufności*). Testy statystyczne, które kontrolują tylko błąd I–rodzaju nazywamy *testami istotności*.

Można zapytać, co dzieje się z błędem II–rodzaju, jeśli błąd I–rodzaju został zminimalizowany. Czy on również przyjmuje minimalną wartość? Okazuje się, że tak nie jest. Jest akurat na odwrót, czyli wartości β rosną, jeśli wartości α maleją (patrz np. [19]). Wobec tego testy istotności jako metoda ignorująca z założenia kontrolę błędu II-rodzaju z jednej strony kontrolują na poziomie α odrzucenie hipotezy H_o (gdy $(x_1, x_2, \dots, x_n) \in Q$), ale nie wypowiadają się na temat jej przyjęcia. Dlatego wnioskowanie w przypadku przeprowadzenia takiego testu sprowadza się do następującej konkluzji:

1. odrzucenia hipotezy H_o i przyjęcia H_1 albo
2. braku podstaw do odrzucenia hipotezy H_o.

Zaczniemy od omówienia podstaw związanych z tzw. *testami parametrycznymi* (patrz też [4], [19]). Przypuśćmy, że na podstawie próby prostej

$$(x_1 \dots x_n) = (\mathbb{X}_1 \dots \mathbb{X}_n)(\omega_o)$$

oraz metod opartych na pojęciu dystrybuanty empirycznej (histogramu) i teorii estymacji otrzymaliśmy następujące wyniki:

1. poznaliśmy typ rozkładu cechy $\mathbb{X}$, czyli jej dystrybuantę F, ewentualnie dysponujemy wiedzą na temat istnienia momentów tego rozkładu i parametrów, od których rozkład ten zależy,
2. zakładając, celem uproszczenia sytuacji, że rozkład F zależy od jednego parametru, powiedzmy τ, określiliśmy metodą estymacji przedziałowej jego przedział na zadanym poziomie ufności $1-\alpha$.

Pojawia się problem wyboru wartości tego parametru. Faza wstępna testu parametrycznego polega na zdefiniowaniu dwóch *hipotez*:

hipotezy zerowej $\mathbb{H}_o$ przeciwko *hipotezie alternatywnej* $\mathbb{H}_1$, gdzie

$$\mathbb{H}_o : \ \tau = \tau_o, \ \mathbb{H}_1 : \ \tau \mathcal{R} \tau_o,$$

gdzie $\tau \mathcal{R} \tau_o$ na ogół oznacza jedną z trzech relacji:

$$\tau \neq \tau_o, \quad \tau < \tau_o, \quad \tau > \tau_o.$$

Rola testu statystycznego, jak pisaliśmy wyżej, polega na tym, aby *zweryfikować* hipotezę zerową przeciwko hipotezie alternatywnej. Pozytywna weryfikacja kończy się na ogół stwierdzeniem, że *nie ma powodów do jej odrzucenia*, co nie musi oznaczać, że hipotezę zerową przyjmuje się. Zazwyczaj w tej sytuacji cały proces analizy statystycznej zaczyna się od początku, a więc od pobrania nowej próby prostej.

W przypadku negatywnej weryfikacji hipotezy zerowej albo przy aktualnych wynikach zmienia się jej treść i powtarza weryfikację albo przyjmuje się hipotezę alternatywną.

Podstawą procesu weryfikacji jest, jak pisaliśmy wyżej, *obszar krytyczny* Q, który w praktyce definiuje się jako podzbiór prostej rzeczywistej, a nie przestrzeni $\mathbf{R}^n$. Aby go wyznaczyć potrzebna jest odpowiednia statystyka $\mathbb{Z}$, czyli

$$\mathbb{Z} = f(\mathbb{X}_1 \ldots \mathbb{X}_n),$$

gdzie

$$(x_1 \ldots x_n) = (\mathbb{X}_1 \ldots \mathbb{X}_n)(\omega_o)$$

jest próbą prostą cechy $\mathbb{X}$ populacji generalnej.

Wtedy obszar Q wyznaczony jest przez warunek (patrz definicja błędu I rodzaju)

$$P(\{\omega \in \Omega : \ \mathbb{Z}(\omega) \in Q\}) = \alpha, \text{ przy założeniu, że zaszła hipoteza } \mathbb{H}_o.$$

Jeśli teraz

$$z_{obs} = \mathbb{Z}(\omega_o) \in Q,$$

gdzie z_{obs} jest *wartością zaobserwowaną* statystyki $\mathbb{Z}$, to hipotezę $\mathbb{H}_o$ należy odrzucić na korzyść hipotezy $\mathbb{H}_1$. W przeciwnym razie mówimy, że nie ma podstaw do jej odrzucenia.

Dalej w kolejnych przykładach pokażemy sposoby konstruowania statystyki $\mathbb{Z}$ i odpowiadających jej obszarów krytycznych (patrz też [10]).

Przykład 6.4.6 *Cecha $\mathbb{X}$ populacji generalnej ma rozkład typu $\mathcal{N}(m, \sigma^2)$, gdzie parametr σ jest znany. Wtedy dla zweryfikowania hipotezy zerowej*

$$\mathbb{H}_o \ : \ m = m_o$$

posługujemy się styatystyką

$$\mathbb{Z} = \frac{\overline{\mathbb{X}_n} - m}{\sigma}\sqrt{n}.$$

Natomiast hipoteza alternatywna $\mathbb{H}_1$ i obszar krytyczny mają wtedy postać:

1.

$$\mathbb{H}_1 \ : \ m \neq m_o, \ Q = (-\infty, -n_\alpha) \cup (n_\alpha, \ +\infty),$$

gdzie n_α, przy założeniu hipotezy $\mathbb{H}_o$ określone jest równaniem

$$P(\{\omega \in \Omega : \ | \ \mathbb{Z}_o \ |(\omega) \geqslant n_\alpha\}) = 2(1 - \Phi(n_\alpha)) = \alpha,$$

bowiem statystyka

$$\mathbb{Z}_o = \frac{\overline{\mathbb{X}_n} - m_o}{\sigma}\sqrt{n}$$

ma przy założeniu hipotezy zerowej rozkład standardowy normalny.

2.

$$\mathbb{H}_1 \ : \ m > m_o, \ Q = (n_\alpha^+, \ +\infty)$$

gdzie n_α^+, przy założeniu hipotezy $\mathbb{H}_o$ określone jest równaniem

$$P(\{\omega \in \Omega \colon \mathbb{Z}_o(\omega) > n_\alpha^+\}) = 1 - \Phi(n_\alpha^+) = \alpha,$$

gdzie statystyka $\mathbb{Z}_o$ jest identyczna, jak w sytuacji poprzedniej.

3.

$$\mathbb{H}_1 \ : \ m < m_o, \ Q = (-\infty, \ -n_\alpha^+)$$

gdzie n_α^+ i statystyka $\mathbb{Z}_o$ są jak wyżej.

Przeprowadźmy teraz symulację liczbową. W tym celu odwołamy się do wyników z przykładu 6.4.1.

Dostaliśmy tam

$$m \in (1{,}838, \ 2{,}328), \ \overline{x_4} = 2{,}083, \ \sigma = 0{,}25, \ \alpha = 0{,}05.$$

Postawmy hipotezę zerową

$$\mathbb{H}_o : \ m = 2 \ \textit{przeciwko} \ \mathbb{H}_1 : \ m \neq 2.$$

Wtedy, przy założeniu hipotezy zerowej $n_\alpha = 1{,}96$*, i przedział krytyczny ma postać*

$$Q = (-\infty,\ -1{,}96) \cup (1{,}96,\ +\infty).$$

Tymczasem $z_{obs} = \frac{2{,}083-2}{0{,}25} = 0{,}684 \notin Q$*, co oznacza, że nie ma powodu do odrzucenia hipotezy zerowej.*

Z drugiej strony biorąc

$$\mathbb{H}_o:\ m = 1{,}9 \text{ przeciwko } \mathbb{H}_1:\ m < 1{,}9$$

dostalibyśmy

$$\Phi(n_\alpha^+) = 1 - \alpha = 0{,}95, \text{ skąd } Q = (-\infty,\ -1{,}65).$$

Ponieważ teraz $z_{obs} = 1{,}46$*, zatem również nie ma podstaw, aby nie przyjąć* $m = 1{,}9$*.*

Okazuje się, że dopiero

$$\mathbb{H}_o:\ m = 2{,}3 \text{ przeciwko } \mathbb{H}_1:\ m < 2{,}3$$

daje $z_{obs} = -1{,}736 \in Q = (-\infty,\ -1{,}65)$*, co oznacza, że hipotezę tę należy odrzucić. Jeśli pobrana próba była reprezentatywna, to oznacza to, że najprawdopodobnie wartość średnia cechy* $\mathbb{X}$ *jest mniejsza od 2,3.*

Przykład 6.4.7 *Cecha* $\mathbb{X}$ *populacji generalnej ma rozkład typu* $\mathcal{N}(m, \sigma^2)$*, gdzie oba parametry są nieznane. Wtedy dla zweryfikowania hipotezy zerowej*

$$\mathbb{H}_o\ :\ m = m_o$$

wykorzystamy statystykę

$$t = \frac{\overline{\mathbb{X}_n} - m}{\mathbb{S}}\sqrt{n-1}.$$

Natomiast hipoteza alternatywna $\mathbb{H}_1$ *i obszar krytyczny mają wtedy postać:*

1.

$$\mathbb{H}_1\ :\ m \neq m_o,\ Q = (-\infty, -t_\alpha) \cup (t_\alpha,\ +\infty),$$

gdzie t_α*, przy założeniu hipotezy* $\mathbb{H}_o$ *określone jest równaniem*

$$P(\{\omega \in \Omega:\ |\ t_o\ |(\omega) \geqslant t_\alpha\}) = \alpha,$$

bowiem statystyka

$$t_o = \frac{\overline{\mathbb{X}_n} - m_o}{\mathbb{S}}\sqrt{n-1}$$

ma przy założeniu hipotezy zerowej rozkład t–Studenta o $n-1$ *stopniach swobody.*