

Jerzy W a w r z y n e k (Wrocław)

WPROWADZENIE DO SEKWENCYJNYCH METOD TESTOWANIA HIPOTEZ

1. Wstęp

Postępowanie sekwencyjne. Zadania statystyczne pojawiają się tam, gdzie o rozkładzie pewnej cechy elementów określonej populacji wnioskujemy na podstawie zaobserwowanych wartości tej cechy dla wybranych elementów populacji. Zbiór tych wybranych elementów tworzy próbkę. Metody wnioskowania statystycznego upraszczają się gdy założymy, że próbka została wybrana w sposób zapewniający niezależność poszczególnych obserwacji. To założenie niezależności przyjmujemy również w niniejszej pracy.

W konkretnym zadaniu statystycznym określony jest zbiór możliwych wniosków (decyzji, hipotez) dotyczących rozkładu badanej cechy. Rozwiązanie zadania polega zaś na tym, że na podstawie pewnej liczby obserwacji wartości tej cechy w próbce przyjmuje się jeden wniosek z tego zbioru i odrzuca się wnioski pozostałe. W wyznaczaniu liczby obserwacji, czyli liczebności próbki potrzebnej do ostatecznego rozwiązania zadania statystycznego można wyróżnić dwa rodzaje postępowania.

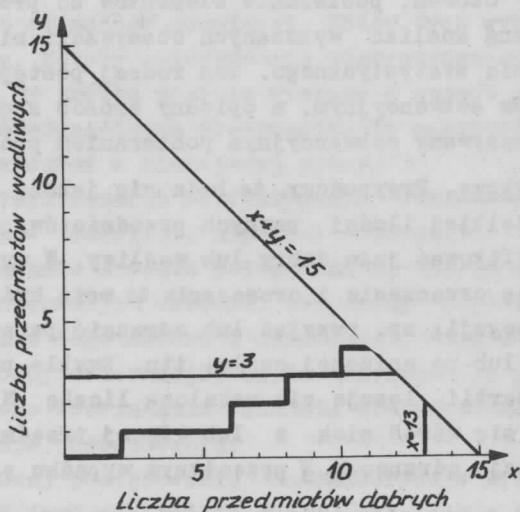
Jeden rodzaj postępowania charakteryzuje się tym, że liczebność próbki N jest wielkością stałą, ustaloną z góry przed rozpoczęciem wybierania elementów populacji do próbki. Do rozwiązania zadania statystycznego przystępuje się po wykonaniu dokładnie N obserwacji. Na podstawie łącznej analizy N obserwacji przyjmuje się jeden wniosek ze zbioru możliwych wniosków i odrzuca się wnioski pozostałe. Opisane postępowanie nazywamy postępowaniem niesekwencyjnym, a pobieranie próbki o z góry ustalonej liczebności nazywamy niesekwencyjnym pobieraniem próbki.

Drugi rodzaj postępowania można opisać następująco. Wybiera się jeden element populacji do próbki. Możliwe są teraz dwie sytuacje. Jeśli wykonana obserwacja upoważnia do przyjęcia jednego wniosku ze zbioru możliwych wniosków i do odrzucenia wniosków pozostałych, to postawione zadanie statystyczne jest rozwiązane i pobieranie próbki jest zakończone. Jeśli wykonana obserwacja nie upoważnia do przyjęcia żadnego wniosku ze zbioru możliwych wniosków, to wybiera się do próbki drugi element. Następuje wtedy łączna analiza obydwu obserwacji. W wyniku tej analizy może się okazać, że zadanie statystyczne jest rozwiązane i tym samym zakończone jest pobieranie próbki, albo że należy pobrać do próbki trzeci element. Ogólnie, po wybraniu n -tego ($n = 1, 2, \dots$) elementu do próbki należy dokonać łącznej analizy n obserwacji. Jeśli analiza ta upoważnia do przyjęcia jednego wniosku ze zbioru możliwych wniosków i do odrzucenia wniosków pozostałych, to pobieranie próbki jest zakończone. W przeciwnym wypadku pobieramy do próbki kolejny element. Słowem, pobieranie elementów do próbki trwa dopóty, dopóki łączna analiza wykonanych obserwacji nie pozwala na rozwiązanie zadania statystycznego. Ten rodzaj postępowania nazywamy postępowaniem sekwencyjnym, a opisany sposób stopniowego pobierania próbki nazywamy sekwencyjnym pobieraniem próbki.

Badanie wyrywkowe. Przypuśćmy, że bada się jakość partii składającej się z wielkiej ilości pewnych przedmiotów. Każdy przedmiot można sklasyfikować jako dobry lub wadliwy. W wyniku badania o partii wydaje się orzeczenie i orzeczenie to może być jedną z dwu dopuszczalnych decyzji; np. przyjąć lub odrzucić partię, sprzedać po normalnej lub po niższej cenie, itp. Zwykle postępuje się wtedy tak, że z partii losuje się ustaloną liczbę N przedmiotów i jeśli znajdzie się wśród nich z lub więcej przedmiotów wadliwych, partia zostaje odrzucona. W przeciwnym wypadku partię przyjmujemy.

Wyniki badania przeprowadzonego w ten sposób można przedstawić graficznie w prostokątnym układzie współrzędnych. Liczbę przedmiotów dobrych x odkładamy na osi poziomej, liczbę przedmiotów wadliwych y odkładamy na osi pionowej. W takim układzie współrzędnych można narysować tzw. ścieżkę próbkową. Ścieżka próbkowa jest linią łamaną zaczynającą się w początku układu współrzędnych. Składa się ona z N odcinków o długości jednostkowej, z których każdy może być poziomy lub pionowy. Pierwszy odcinek ścieżki ma

jeden z końców w początku układu współrzędnych; rysujemy go poziomo, jeśli pierwszy wylosowany przedmiot okazał się dobry; rysujemy go pionowo, gdy pierwszy wylosowany przedmiot okazał się wadliwy. Odcinek n -ty ($n = 2, 3, \dots, N$) rysujemy poczynając od "wolnego" końca odcinka o numerze $n-1$; rysujemy go poziomo, jeśli n -ty wylosowany przedmiot okazał się dobry; rysujemy go pionowo, gdy n -ty wylosowany przedmiot okazał się wadliwy. Koniec ścieżki próbkowej znajduje się w pewnym punkcie prostej o równaniu $x + y = N$, bo z partii losuje się N przedmiotów. Jeśli koniec ścieżki próbkowej znajduje się w takim punkcie tej prostej, w którym $y \geq z$, to partię odrzucamy. Jeśli ścieżka kończy się w punkcie, dla którego $y \leq z - 1$, partię przyjmujemy. Na rys. 1 mamy, że dla $N = 15$ i $z = 3$ ścieżka próbkowa kończy się w punkcie, w którym $y \geq 3$, tzn. badaną partię należy odrzucić.



Rys.1

Łatwo zauważyć, że opisany plan badania wrywkowego może być bardzo niedoskonały. Na przykład, jeśli $N = 15$, $z = 3$, a pierwsze trzy kolejno wylosowane przedmioty okazały się wadliwe, wiadomo już, że partię należy odrzucić. Opisane wyżej postępowanie nakazuje jednak wylosować jeszcze 12 przedmiotów z partii, chociaż to

losowanie nie zmieni już orzeczenia o badanej partii. Ten rodzaj marnotrawstwa bierze się stąd, że w opisanym planie badania wyrzywkowego z góry ustalona jest liczebność N próbki, którą należy pobrać z partii, tzn. mamy tu przypadek niesekwencyjnego pobierania próbki.

Rozważmy teraz nieco inny plan badania wyrzykowego. Według tego planu przedmioty z partii losuje się dopóty, dopóki nie zaobserwuje się wśród wylosowanych przedmiotów albo z przedmiotów wadliwych albo $N-z+1$ przedmiotów dobrych. Losowanie przerywa się z chwilą zaobserwowania jednego z tych dwu zdarzeń. Jeśli w momencie zakończenia losowania w próbce znajduje się z przedmiotów wadliwych, odrzuca się badaną partię. Jeśli w momencie zakończenia losowania w próbce znajduje się $N-z+1$ przedmiotów dobrych, badana partia zostaje przyjęta. Przedstawienie graficzne tego postępowania polega na narysowaniu prostej poziomej $y = z$ i prostej pionowej $x = N-z+1$, co oznacza zakończenie losowania wtedy, kiedy ścieżka próbkowa osiągnie jedną z tych dwu prostych. Np. z rys. 1 widać, że badanie należy zakończyć po wylosowaniu 11 przedmiotów i że badaną partię należy odrzucić.

Przedstawiony teraz plan badania wyrzykowego ma cechy postępowania sekwencyjnego. Po wylosowaniu kolejnego przedmiotu trzeba bowiem sprawdzić, czy zaszło już jedno z dwu zdarzeń, których zaobserwowanie pozwala zakończyć badanie. Liczebność próbki nie jest więc stała, zmienia się, choć nie może przekroczyć N . Dla każdej możliwej ścieżki próbkowej oba przedstawione tu plany badania wyrzykowego dają zgodne orzeczenie o badanej partii. Jednak użycie planu sekwencyjnego pozwala tu często zmniejszyć liczbę badanych przedmiotów.

Omówimy teraz na przykładzie wyżej opisanych planów badania wyrzykowego zalety i wady planów z sekwencyjnym pobieraniem próbki. Używanie planu z sekwencyjnym pobieraniem próbki jest bardziej skomplikowane niż używanie planu, w którym liczebność próbki jest określona z góry. Wynika stąd, że badanie przebiega wolniej, bo należy je zatrzymywać po wylosowaniu każdego nowego przedmiotu, żeby sprawdzić czy spełnione są różne warunki. Przypuśćmy dalej, że stosujemy sekwencyjne badanie wyrzykowe np. gdy $N = 15$, $z = 3$. Jeśli pierwsze trzy wylosowane przedmioty są wadliwe, przerywamy badanie i partię odrzucamy. Mamy wystarczającą informację do wydania właściwego orzeczenia o partii, ale bardzo mało wiemy wtedy np. o procencie przedmiotów wadliwych w partii. Nie można jednak

krytykować sekwencyjnego planu badania wyrywkowego za brak własności, których dla tego planu nie postulowano. Rzeczywista korzyść ze stosowania sekwencyjnego planu badania wyrywkowego jest ta, że podczas gdy dla każdej danej próbki oba opisane plany dają zgodne orzeczenie o partii, to przez stosowanie planu sekwencyjnego można średnio znacznie zmniejszyć liczbę losowanych przedmiotów.

2. Sekwencyjny test ilorazowy

Opis testu. Podstawy teorii znanej dziś pod nazwą analizy sekwencyjnej stworzył około 1943 roku A. Wald. Najważniejszym wynikiem tej teorii jest sekwencyjny test ilorazowy.

Przypuśćmy, że zmienna losowa X ma rozkład prawdopodobieństwa o gęstości (funkcji prawdopodobieństwa) $f(x, p)$, gdzie p jest pewnym parametrem tego rozkładu. Chcemy zweryfikować hipotezę zerową, że $p = p_0$, przeciwko hipotezie alternatywnej, że $p = p_1$. Przyjmujemy jedną z tych hipotez na podstawie niezależnych obserwacji x_1, x_2, \dots zmiennej losowej X . Test ma być tak skonstruowany, że gdy prawdziwa jest hipoteza zerowa, prawdopodobieństwo jej odrzucenia ma być równe α , gdy zaś prawdziwa jest hipoteza alternatywna, prawdopodobieństwo jej odrzucenia ma być równe β .

W przypadku próbki o z góry zadanej liczebności (niesekwencyjne pobieranie próbek) optymalny test dla takiego sagadnienia konstruuje się w oparciu o klasyczny lemat Neymana - Pearsona. Tworzy się mianowicie iloraz największej wiarogodności L_n , gdzie

$$(1) \quad L_n = \frac{\text{Prawdopodobieństwo zaobserwowanych wyników, jeśli } p=p_1}{\text{Prawdopodobieństwo zaobserwowanych wyników, jeśli } p=p_0} \\ = \prod_{i=1}^n \frac{f(x_i, p_1)}{f(x_i, p_0)}$$

i jeśli L_n jest mniejsze od pewnej stałej C , przyjmuje się, że $p = p_0$, a gdy L_n jest większe od tej stałej, przyjmuje się, że $p = p_1$. Przy zadanych prawdopodobieństwach α i β stałą C wyznacza się w zależności od α , zaś liczebność próbki N ustala się na podstawie wielkości prawdopodobieństwa β . Taki test jest testem najmocniejszym, tzn. dla każdego innego testu - przy poprzednio zadanym prawdopodobieństwie α i wyznaczonej liczebności próbki N - prawdopodobieństwo przyjęcia hipotezy zerowej, gdy prawdziwa jest hipoteza alternatywna, nie jest mniejsze od β .

Sekwencyjny test ilorazowy jest analogonem powyższego klasycznego testu. Opiszemy teraz, jak weryfikuje się hipotezy przy pomocy tego testu.

Kontynuujemy losowanie elementów do próbki jeśli

$$(2) \quad B < L_n < A.$$

Losowanie przerywamy, skoro tylko $L_n \leq B$ lub $L_n \geq A$. Jeśli $L_n \leq B$, przyjmujemy hipotezę, że $p = p_0$; jeśli $L_n \geq A$, przyjmujemy hipotezę, że $p = p_1$. Stałe A i B wyznacza się tak, żeby w przybliżeniu otrzymać zadane prawdopodobieństwa α i β błędów I i II rodzaju. Zgodnie z tym Wald udowodnił, że

$$(3) \quad A \approx \frac{1 - \beta}{\alpha}, \quad B \approx \frac{\beta}{1 - \alpha}.$$

Szczegółowy dowód przybliżonych równości (3) znajduje się w [5], str. 40-48, zaś szkice dowodu można znaleźć w [3], str. 491 i [6], str. 15-16.

Stosowanie sekwencyjnego testu ilorazowego dla hipotez złożonych. Rozkład prawdopodobieństwa zmiennej losowej X , o której mówiliśmy wyżej w opisie sekwencyjnego testu ilorazowego, zawierał jeden nieznaną parametr p . Hipotezę dotyczącą parametru p nazywamy hipotezą prostą, jeśli w sposób jednoznaczny określa ona wartość liczbową tego parametru. Natomiast jeśli hipoteza dotyczy większego zbioru wartości liczbowych parametru p , to nazywamy ją hipotezą złożoną. Na przykład hipoteza orzekająca, że $p = 5$ jest hipotezą prostą. Z kolei hipoteza orzekająca, że $3 < p < 5$ jest hipotezą złożoną. Wprowadzone rozróżnienie pozwala na stwierdzenie, że sekwencyjny test ilorazowy przeznaczony jest do weryfikowania prostej hipotezy zerowej przeciwko prostej hipotezie alternatywnej.

W praktyce testowania hipotez rzadko mamy do czynienia z sytuacją, w której zarówno zerowa jak i alternatywna hipoteza o nieznanym parametrze p rozkładu prawdopodobieństwa zmiennej losowej X są hipotezami prostymi. Znacznie częściej spotykamy sytuacje, gdzie przynajmniej jedna z testowanych hipotez jest hipotezą złożoną. Przedstawimy teraz kilka typowych przykładów takich hipotez i sformułujemy reguły, które pokazują jak należy wtedy stosować opisany wyżej sekwencyjny test ilorazowy. Reguły te są ważne dla większości często spotykanych w praktyce rozkładów prawdopodobieństwa. W szczególności można je stosować wtedy, gdy zmienna losowa X ma rozkład normalny o znanej wartości oczekiwanej i nie-

znanej wariancji p , rozkład normalny o nieznannej wartości oczekiwanej p i znanej wariancji, rozkład dwumianowy bądź rozkład Poissona.

Zacznijmy od przykładu następującego. Chcemy zweryfikować prostą hipotezę zerową, że $p = p_0$, przeciwko złożonej hipotezie alternatywnej, że $p \geq p_1$, gdzie $p_1 > p_0$. Test użyty do tego celu ma być taki, że prawdopodobieństwo odrzucenia hipotezy zerowej, gdy jest ona prawdziwa, ma być równe α oraz prawdopodobieństwo odrzucenia hipotezy alternatywnej, gdy jest ona prawdziwa, ma nie przekraczać β . Warunki te spełnia sekwencyjny test ilorazowy dla testowania prostej hipotezy zerowej $p = p_0$ przeciwko prostej hipotezie alternatywnej $p = p_1$ z prawdopodobieństwami błędów α i β .

Całkowicie analogiczna jest reguła użycia sekwencyjnego testu ilorazowego dla zweryfikowania prostej hipotezy zerowej, że $p = p_0$, przeciwko złożonej hipotezie alternatywnej, że $p \leq p_1$, gdzie $p_1 < p_0$.

Zajmijmy się teraz weryfikowaniem złożonej hipotezy zerowej, że $p \leq p_0$, przeciwko złożonej hipotezie alternatywnej, że $p \geq p_1$, gdzie $p_1 > p_0$. Test użyty do zweryfikowania tych hipotez ma być taki, że prawdopodobieństwo odrzucenia hipotezy zerowej, gdy jest ona prawdziwa, ma nie przekraczać α , natomiast prawdopodobieństwo odrzucenia hipotezy alternatywnej, gdy jest ona prawdziwa, ma nie przekraczać β . Wymienione warunki spełnia sekwencyjny test ilorazowy dla testowania prostej hipotezy zerowej $p = p_0$ przeciwko prostej hipotezie alternatywnej $p = p_1$ z prawdopodobieństwami błędów α i β .

Analogicznie używa się też sekwencyjnego testu ilorazowego dla weryfikowania hipotezy zerowej, że $p \geq p_0$, przeciwko hipotezie alternatywnej, że $p \leq p_1$, gdzie $p_1 < p_0$.

Przykład zastosowania sekwencyjnego testu ilorazowego. W medycznych pracach badawczych często chodzi o porównanie skuteczności dwu leków. Weźmiemy pod uwagę dwa sposoby takich badań. Jeden z nich polega na tym, że badanie wykonuje się na jednej grupie pacjentów. Drugi sposób polega na tym, że badanie wykonuje się na dwu równolicznych, połączonych w pary, grupach pacjentów. Jeśli badanie wykonuje się na jednej grupie, to na każdym pacjencie z tej grupy wypróbowuje się w losowej kolejności obydwie leki. Jeśli badanie wykonuje się na parach pacjentów, to obydwie leki wypróbowuje się w każdej parze w ten sposób, że na pierwszym pacjencie wypróbowuje się jeden lek wybrany losowo spośród dwu bada-

nych leków, zaś na drugim pacjencie z tej pary wypróbujemy się lek pozostały. W takich sytuacjach nie należy jednak wprowadzać do doświadczenia wszystkich pacjentów jednocześnie, a to dlatego, że jak zauważa Armitage, rozważania etyczne nakazują przerwać doświadczenie skoro tylko stanie się jasne, który z badanych leków należy preferować. Krótko mówiąc, doświadczenie powinno mieć charakter sekwencyjny.

Oznaczmy badane leki literami K i L. Zakładamy, że dla każdego pacjenta (względnie dla każdej pary pacjentów) możliwe są tylko dwa wyniki badania: 1^o lek K jest gorszy od leku L; 2^o lek K jest lepszy od leku L. Zakładamy dalej, że dla każdego pacjenta (każdej pary pacjentów) prawdopodobieństwo wyniku oznaczonego przez 1^o wynosi p, a prawdopodobieństwo wyniku oznaczonego przez 2^o wynosi 1-p, oraz że wyniki u poszczególnych pacjentów (poszczególnych par pacjentów) są niezależne. Wobec tego prawdopodobieństwo $P(X=x, Y=y)$, że w pewnym momencie doświadczenia wynik oznaczony przez 1^o zaobserwowano u x pacjentów (par pacjentów), zaś wynik oznaczony przez 2^o zaobserwowano u y pacjentów (par pacjentów) jest dane rozkładem dwumianowym i wynosi

$$(4) \quad P(X=x, Y=y) = \binom{x+y}{x} p^x (1-p)^y$$

Za pomocą sekwencyjnego testu ilorazowego chcemy zweryfikować hipotezę zerową (H_0), że skuteczność obu leków jest jednakowa przeciwko hipotezie alternatywnej (H_1), że lek K jest lepszy od leku L co najmniej w 75 % przypadków. Jeśli prawdziwa jest hipoteza zerowa, to obydwa wyniki badania są jednakowo prawdopodobne, czyli mamy $p_0 = 0,5$. Gdy prawdziwa jest hipoteza alternatywna, to prawdopodobieństwo wyniku oznaczonego przez 1^o nie przekracza $p_1 = 1 - 0,75 = 0,25$. Przyjmujemy następujące prawdopodobieństwa błędów I i II rodzaju: $\alpha = 0,02$; $\beta \leq 0,05$.

Należy teraz obliczyć iloraz największej wiarygodności L_n . Ze wzorów (1), (4) oraz z postawionych hipotez otrzymujemy kolejno

$$(5) \quad L_n = L_{x+y} = \frac{\binom{x+y}{x} p_1^x (1-p_1)^y}{\binom{x+y}{x} p_0^x (1-p_0)^y} = \left(\frac{p_1}{p_0} \right)^x \left(\frac{1-p_1}{1-p_0} \right)^y = \left(\frac{0,25}{0,5} \right)^x \left(\frac{1-0,25}{1-0,5} \right)^y = (0,5)^x (1,5)^y$$

Na podstawie wzorów (3) obliczamy teraz stałe A i B podstawiając $\alpha = 0,02$, $\beta = 0,05$.

Otrzymujemy wtedy

$$(6) \quad A \simeq \frac{1 - \beta}{\alpha} = \frac{1 - 0,05}{0,02} = 47,5$$

$$(7) \quad B \simeq \frac{\beta}{1 - \alpha} = \frac{0,05}{1 - 0,02} = 0,051$$

Z opisu sekwencyjnego testu ilorazowego wynika, że w omawianym przykładzie do doświadczenia należy włączyć kolejnego pacjenta (parę pacjentów) jeśli (por. nierówność (2))

$$(8) \quad 0,051 < (0,5)^x (1,5)^y < 47,5.$$

Doświadczenie należy zakończyć, gdy nierówność podwójna (8) nie jest spełniona. Jeśli $(0,5)^x (1,5)^y \leq 0,051$, to przyjmujemy hipotezę zerową o jednakowej skuteczności obu leków. Jeśli $(0,5)^x (1,5)^y \geq 47,5$ przyjmujemy hipotezę alternatywną, że lek K jest lepszy od leku L co najmniej w 75 % przypadków.

Zlogarytmujemy stronami nierówność (8). Mamy wtedy

$$\log 0,051 < x \cdot \log 0,5 + y \cdot \log 1,5 < \log 47,5$$

skąd otrzymujemy

$$-1,2924 < -0,301 \cdot x + 0,1761 \cdot y < 1,6767$$

Do nierówności tej dodajemy stronami wyrażenie $0,301 \cdot x$, a następnie dzielimy ją stronami przez liczbę $0,1761$. Ostatecznie dostaliśmy

$$(9) \quad 1,71 \cdot x - 7,34 < y < 1,71 \cdot x + 9,52$$

Otrzymana nierówność jest równoważna nierówności (8). Jest ona bardziej wygodna do naszych celów. Trzeba tylko pomnożyć aktualną liczbę x pacjentów (parę pacjentów), u których zaobserwowano wynik oznaczony przez 1° , przez współczynnik $1,71$, a do wartości tego iloczynu dodać odpowiednio liczby $-7,34$ i $+9,52$. Otrzymamy w ten sposób dwie liczby graniczne. Jeśli aktualna liczba y pacjentów (parę pacjentów), u których zaobserwowano wynik oznaczony przez 2° , zawiera się między tymi dwiema liczbami granicznymi, to należy włączyć do doświadczenia kolejnego pacjenta (parę pacjentów). Jeśli zaobserwowana liczba y jest większa od mniejszej z dwu liczb granicznych przyjmujemy hipotezę zerową, że skuteczność obu leków jest jednakowa. Jeśli zaobserwowana liczba y nie jest mniejsza od większej z dwu liczb granicznych, to przyjmujemy hipotezę alternatywną, że lek K jest lepszy od leku L co najmniej w 75 % przypadków.

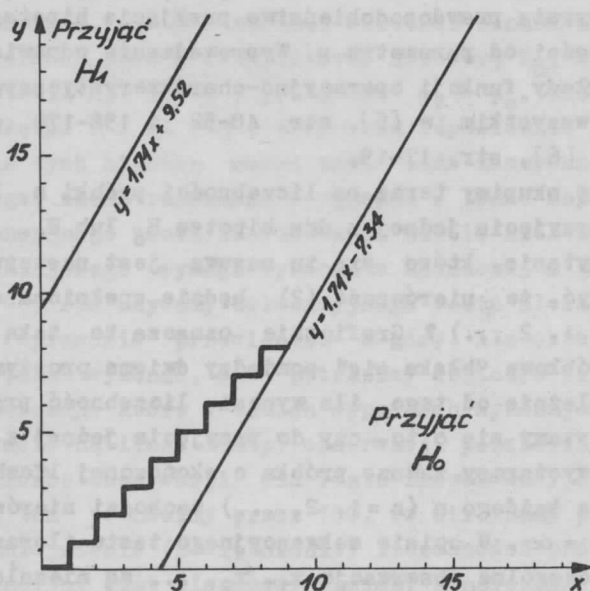
Opisany tu plan badania sekwencyjnego zyskuje jeszcze na przejrzystości, jeśli przedstawimy go graficznie w prostokątnym układzie współrzędnych. Liczbę pacjentów (par pacjentów) x , dla których lek K okazał się gorszy od leku L, odkładamy na osi poziomej; liczbę pacjentów (par pacjentów) y , dla których lek K okazał się lepszy od leku L, odkładamy na osi pionowej. W tym układzie współrzędnych rysujemy (patrz rys. 2) dwie proste równoległe o równaniach

$$(10) \quad y = 1,71 \cdot x - 7,34$$

oraz

$$(11) \quad y = 1,71 \cdot x + 9,52$$

Prawe strony równań (10) i (11) wzięte są z nierówności podwójnej (9).



Rys. 2

W czasie prowadzenia doświadczenia rysujemy w tym układzie współrzędnych ścieżkę próbkową. Kolejny odcinek tej ścieżki rysujemy poziomo, jeśli u danego pacjenta (pary pacjentów) lek K okazał się gorszy od leku L. W przeciwnym wypadku rysujemy ten odcinek pionowo. Należy kontynuować włączanie kolejnego pacjenta do doświadczenia dopóty, dopóki ścieżka próbkowa znajduje się między

prostymi (10) i (11). Doświadczenie należy zakończyć skoro tylko ścieżka próbkowa osiągnie lub przekroczy jedną z tych prostych. Jeśli ścieżka próbkowa osiągnie lub przekroczy prostą o równaniu (10) należy wówczas przyjąć hipotezę H_0 , że skuteczność obu leków jest jednakowa. Jeśli ścieżka próbkowa osiągnie lub przekroczy prostą o równaniu (11), należy przyjąć hipotezę H_1 , że lek K jest lepszy od leku L co najmniej w 75 % przypadków. Np. z rys. 2 wiadać, że po zbadaniu 17 pacjentów (par pacjentów) należy zakończyć doświadczenie i przyjąć hipotezę H_0 .

Niektóre własności sekwencyjnego testu ilorazowego. Przegląd własności sekwencyjnego testu ilorazowego rozpoczęliśmy od stwierdzenia, że w przybliżeniu znana jest funkcja operacyjno-charakterystyczna tego testu. Przy ustalonych α , β , p_0 , p_1 funkcja operacyjno-charakterystyczna $P(p|\alpha, \beta, p_0, p_1)$ sekwencyjnego testu ilorazowego wyraża prawdopodobieństwo przyjęcia hipotezy, że $p = p_0$, w zależności od parametru p . Wyprowadzenie odpowiednich wzorów oraz przykłady funkcji operacyjno-charakterystycznych znaleźć można przede wszystkim w [5], str. 48-52 i 158-170 oraz w [3], str. 487-489 i [6], str. 17-19.

Naszą uwagę skupimy teraz na liczebności próbki n , która jest potrzebna do przyjęcia jednej z dwu hipotez H_0 lub H_1 .

Pierwsze pytanie, które się tu nasuwa jest następujące. Czy może się zdarzyć, że nierówność (2) będzie spełniona dla każdej liczby n ($n = 1, 2, \dots$)? Graficznie oznacza to taką sytuację, że ścieżka próbkowa "błąka się" pomiędzy dwiema prostymi równoległymi niezależnie od tego, ile wynosi liczebność próbki. Inaczej mówiąc, pytamy się o to, czy do przyjęcia jednej z dwu hipotez H_0 lub H_1 wystarczy zawsze próbka o skończonej liczbie elementów. Jeżeli dla każdego n ($n = 1, 2, \dots$) zachodzi nierówność (2), powiemy, że $n = \infty$. W opisie sekwencyjnego testu ilorazowego założono, że poszczególne obserwacje x_1, x_2, \dots są niezależne i pochodzą z tego samego rozkładu o gęstości (funkcji prawdopodobieństwa) $f(x, p)$. Wprowadzając nową zmienną losową

$$(12) \quad Z = \log \frac{f(x, p_1)}{f(x, p_0)}$$

można udowodnić, że jeżeli wariancja zmiennej losowej Z jest różna od zera, to prawdopodobieństwo, że $n = \infty$, jest równe zeru. Dowód tego twierdzenia podany jest w [3], str. 482-483 i [5], str. 157-158.

Liczebność próbki n nie jest ustalona z góry, lecz jest pewną zmienną losową. Niestety o rozkładzie prawdopodobieństwa tej zmiennej losowej wiemy niewiele. Znane są jednak przybliżone wzory na wartość oczekiwaną $E(n|p)$ tej zmiennej losowej, tzn. wzory podające przeciętną ilość obserwacji potrzebnych do zweryfikowania przy pomocy sekwencyjnego testu ilorazowego hipotezy H_0 przeciwko hipotezie H_1 (patrz [3], str. 489-491, [5], str. 52-54, 65-69, 170-180 i [6], str. 19-21). Nie będziemy tu przytaczali tych wzorów, podamy jednak kilka przykładów na to, ile wynosi ta przeciętna ilość obserwacji potrzebnych do zweryfikowania hipotez przy pomocy sekwencyjnego testu ilorazowego. Następnie porównamy te wartości z liczbą obserwacji, które należy wykonać, gdy te same hipotezy weryfikujemy za pomocą klasycznych testów, w których liczebność próbki jest ustalona z góry.

Założmy najpierw, że badana cecha ma rozkład normalny o wariancji równej 1, ale o nieznannej wartości oczekiwanej p . Na podstawie obserwacji chcemy zweryfikować hipotezę H_0 , że $p = p_0$, przeciwko hipotezie H_1 , że $p = p_1$, gdzie $p_1 > p_0$. Niech prawdopodobieństwa błędów I i II rodzaju wynoszą odpowiednio α i β . Do zweryfikowania tych hipotez można użyć bądź klasycznego testu najmocniejszego, skonstruowanego w oparciu o lemat Neymana-Pearsona, bądź sekwencyjnego testu ilorazowego. Użycie niesekwencyjnego testu najmocniejszego wymaga wykonania ustalonej z góry liczby obserwacji. Gdy zaś używamy sekwencyjnego testu ilorazowego, to nie potrafimy wprawdzie przewidzieć z góry, ile obserwacji trzeba w danym wypadku wykonać, ale potrafimy obliczyć przeciętną ilość $E(n|p)$ obserwacji, które w takich wypadkach wykonujemy. Jeżeli teraz tą przeciętną ilość $E(n|p)$ obserwacji podzielimy przez z góry ustaloną liczbę obserwacji dla testu niesekwencyjnego, a następnie iloraz ten pomnożymy przez 100, to otrzymamy przeciętne procentowe zmniejszenie (powiększenie) liczebności próbki gdy stosujemy sekwencyjny test ilorazowy zamiast klasycznego testu najmocniejszego pod warunkiem, że wartość oczekiwana rozkładu badanej cechy wynosi p . Tablice 1 i 2 (zaczepnięte z książki Walda [5], str. 57) podają dla różnych α i β przeciętne zmniejszenie procentowe liczebności próbki przy stosowaniu sekwencyjnego testu ilorazowego zamiast niesekwencyjnego testu najmocniejszego w przypadku, gdy prawdziwa jest odpowiednio hipoteza H_0 (tablica 1) lub hipoteza H_1 (tablica 2). Z tablic tych widać, że przeciętne zmniejszenie liczby obserwacji jest bardzo znaczne i wynosi około 50 %.

T a b l i c a 1

$\beta \backslash \alpha$	0,01	0,02	0,03	0,04	0,05
0,01	58	54	51	49	47
0,02	60	56	53	50	49
0,03	61	57	54	51	50
0,04	62	58	55	52	50
0,05	63	59	55	53	51

T a b l i c a 2

$\beta \backslash \alpha$	0,01	0,02	0,03	0,04	0,05
0,01	58	60	61	62	63
0,02	54	56	57	58	59
0,03	51	53	54	55	55
0,04	49	50	51	52	53
0,05	47	49	50	50	51

Założymy teraz, że badana cecha ma rozkład dwupunktowy o nieznanym prawdopodobieństwie sukcesu p . Na podstawie obserwacji chcemy zweryfikować hipotezę H_0 , że $p = 0,25$, przeciwko hipotezie H_1 , że $p = 0,75$. Niech prawdopodobieństwa błędów I i II rodzaju wynoszą $\alpha = \beta = 0,001$. Dla kilku wartości p obliczono przeciętną ilość $E(n|p)$ obserwacji, które trzeba wykonać, używając sekwencyjnego testu ilorazowego do zweryfikowania postawionych hipotez. Rezultaty tych obliczeń zamieszczono w tabelicy 3, którą zaczerpnięto z książki Wetherilla [6], str. 23. Łatwo pokazać, że gdybyśmy te

T a b l i c a 3

p	0,25	0,37	0,43	0,50	0,57	0,63	0,75
$E(n p)$	12,6	22,0	32,2	39,5	32,2	22,0	12,6

same hipotezy z tymi samymi prawdopodobieństwami błędów $\alpha = \beta = 0,001$ chcieli zweryfikować za pomocą niesekwencyjnego testu najmocniejszego, to potrzebne jest do tego wykonanie 33 obserwacji.

Porównanie tego wyniku z wynikami zamieszczonymi w tabelicy 3 pozwala sformułować następujące wnioski. Jeśli prawdziwe wartości nieznanego parametru p są bliskie hipotetycznym wartościom $p = 0,25$ i $p = 0,75$, wtedy sekwencyjny test ilorazowy jest bardzo efektywny w sensie przeciętnej liczby obserwacji. Jeśli jednak hipotetyczne wartości $p = 0,25$ i $p = 0,75$ zostały wybrane tylko po to, by umożliwić skonstruowanie sekwencyjnego testu ilorazowego, a nie reprezentują nawet w przybliżeniu prawdziwej wartości nieznanego parametru p , to w porównaniu z klasycznym testem najmocniejszym sekwencyjny test ilorazowy nie daje większego przeciętnego zmniejszenia ilości obserwacji, a czasem nawet (gdy p wynosi około 0,5) niekorzystnie ją zwiększa.

3. Test sekwencyjny w przypadku postawienia trzech hipotez

Opis testu. Sposób postępowania opiszemy tu na przykładzie porównywania skuteczności dwu leków. Oznaczmy badane leki literami K i L . Zakładamy, że dla każdego pacjenta (pary pacjentów) możliwe są tylko dwa wyniki: 1° lek K jest gorszy od leku L ; 2° lek K jest lepszy od leku L . Zakładamy dalej, że dla każdego pacjenta (pary pacjentów) prawdopodobieństwo wyniku oznaczonego przez 1° wynosi p , przy czym wyniki u poszczególnych pacjentów (par pacjentów) są niezależne.

Na podstawie obserwacji mamy przyjąć jedną z trzech możliwości: lek K jest gorszy od leku L , lek K jest równie skuteczny jak lek L , lek K jest lepszy od leku L . Te 3 możliwości przedstawimy w postaci trzech hipotez o prawdopodobieństwie p :

$$(13) \quad \begin{cases} H_{-1}: & p \geq p_{-1} > 0,5 \\ H_0: & p = p_0 = 0,5 \\ H_1: & p \leq p_1 = 1 - p_{-1} < 0,5 \end{cases}$$

Z trójki hipotez (13) wybierzmy dwie hipotezy H_{-1} i H_0 . Hipotezę H_{-1} możemy zweryfikować przeciwko hipotezie H_0 za pomocą sekwencyjnego testu ilorazowego. Niech w tym teście, który oznaczymy przez S_{-1} , prawdopodobieństwo przyjęcia hipotezy H_{-1} pod warunkiem, że prawdziwa jest hipoteza H_0 , wynosi $\alpha/2$. Natomiast prawdopodobieństwo przyjęcia hipotezy H_0 pod warunkiem, że prawdziwa jest hipoteza H_{-1} , niech nie przekracza β . W całym rozdziale będziemy zakładać, że prawdopodobieństwa błędów spełniają nierów-

ność $\alpha \leq 2\beta$. Przy pomocy testu S_{-1} można przyjąć albo hipotezę H_{-1} albo hipotezę H_0 .

Z kolei z trójki hipotez (13) wybieramy dwie hipotezy H_0 i H_1 . Hipotezę H_0 zweryfikujemy przeciwko hipotezie H_1 za pomocą sekwencyjnego testu ilorazowego, który oznaczymy tu przez S_1 . Niech test S_1 będzie tak skonstruowany, że prawdopodobieństwo przyjęcia hipotezy H_1 , pod warunkiem, że prawdziwa jest hipoteza H_0 , wynosi $\alpha/2$. Natomiast prawdopodobieństwo przyjęcia hipotezy H_0 pod warunkiem, że prawdziwa jest hipoteza H_1 , niech nie przekracza β . Przy pomocy testu S_1 można przyjąć albo hipotezę H_0 albo hipotezę H_1 .

Do decyzji o przyjęciu jednej z trzech hipotez (13) dochodzimy w następujący sposób. Prowadzimy sekwencyjne badanie skuteczności leków K i L. Obserwacje pochodzące z tego badania służą równocześnie za podstawę testu S_{-1} i testu S_1 . Po wykonaniu pewnej liczby obserwacji okazuje się, że dla jednego z tych testów nie jest spełniona nierówność (2) i tym samym dochodzimy do przyjęcia jednej z dwu hipotez weryfikowanych przy pomocy tego testu. Zakończyliśmy równoczesne operowanie dwoma testami S_{-1} i S_1 , a dalsze kolejne obserwacje służą już tylko za podstawę tego z nich, który spełnia jeszcze nierówność (2). Wreszcie i dla tego testu dochodzimy do przyjęcia jednej z dwu hipotez weryfikowanych przy jego pomocy, wobec czego kończymy sekwencyjne badanie skuteczności leków K i L. Na podstawie wyników testów S_{-1} i S_1 należy teraz powziąć ostateczną decyzję o przyjęciu jednej z trzech hipotez (13). Okazuje się, że opisane postępowanie dopuszcza tylko 3 sytuacje, odpowiadające trzem końcowym decyzjom. Sytuacje te i końcowe decyzje zamieszczono w tabelicy 4.

T a b l i c a 4

Przyjęta hipoteza		Ostateczna decyzja
Test S_{-1}	Test S_1	
H_{-1}	H_0	H_{-1}
H_0	H_0	H_0
H_0	H_1	H_1

Przykład liczbowy i ilustracja graficzna. Postawiono 3 hipotezy o skuteczności leków K i L. Hipoteza H_{-1} orzeka, że lek K jest gorszy od leku L co najmniej w 75 % przypadków. Hipoteza H_0 orzeka, że skuteczność obu leków jest jednakowa. Hipoteza H_1 orzeka, że lek K jest lepszy od leku L co najmniej w 75 % przypadków. Stąd i z (13) mamy

$$(14) \quad \begin{cases} H_{-1}: & p \geq p_{-1} = 0,75 \\ H_0: & p = p_0 = 0,5 \\ H_1: & p \leq p_1 = 1 - 0,75 = 0,25 \end{cases}$$

Niech w naszym przykładzie $\alpha = 0,02$ i $\beta = 0,05$, spełniony jest więc warunek $\alpha \leq 2\beta$. Oznaczmy przez x aktualną liczbę pacjentów (par pacjentów), u których zaobserwowano wynik oznaczony wyżej w opisie przez 1^0 , zaś przez y oznaczmy aktualną liczbę pacjentów (par pacjentów), u których zaobserwowano wynik oznaczony przez 2^0 .

Na podstawie wzorów (3) i zgodnie z tym, co powiedziano o prawdopodobieństwach błędów dla testów S_{-1} i S_1 , obliczamy stałe A i B podstawiając $\alpha/2 = 0,01$ i $\beta = 0,05$. Otrzymujemy wtedy

$$(15) \quad A \simeq \frac{1-\beta}{\alpha/2} = \frac{1-0,05}{0,01} = 95$$

$$(16) \quad B \simeq \frac{\beta}{1-(\alpha/2)} = \frac{0,05}{1-0,01} = 0,0505$$

Ze wzorów (1), (4) oraz (14) obliczamy ilorazy największej wiarygodności dla testu S_{-1} i testu S_1 . Oznaczając ten iloraz przez L_{-1} dla testu S_{-1} , zaś przez L_1 dla testu S_1 , otrzymujemy kolejno

$$(17) \quad L_{-1} = \frac{\binom{x+y}{x} p_{-1}^x (1-p_{-1})^y}{\binom{x+y}{x} p_0^x (1-p_0)^y} = \left(\frac{p_{-1}}{p_0} \right)^x \left(\frac{1-p_{-1}}{1-p_0} \right)^y = \\ = \left(\frac{0,75}{0,5} \right)^x \left(\frac{1-0,75}{1-0,5} \right)^y = (1,5)^x (0,5)^y$$

$$(18) \quad L_1 = \frac{\binom{x+y}{x} p_1^x (1-p_1)^y}{\binom{x+y}{x} p_0^x (1-p_0)^y} = \left(\frac{p_1}{p_0} \right)^x \left(\frac{1-p_1}{1-p_0} \right)^y = \\ = \left(\frac{0,25}{0,5} \right)^x \left(\frac{1-0,25}{1-0,5} \right)^y = (0,5)^x (1,5)^y$$

Z (15), (16) i (17) mamy, że nierówność (2) dla testu S_{-1} wygląda następująco

$$(19) \quad 0,0505 < (1,5)^x (0,5)^y < 95$$

Podobnie z (15), (16) i (18) otrzymujemy, że nierówność (2) dla testu S_1 jest następująca

$$(20) \quad 0,0505 < (0,5)^x (1,5)^y < 95$$

Nierówności (19) i (20) przekształcimy teraz do prostszej postaci. Przekształcając nierówność (19) otrzymujemy kolejno

$$\begin{aligned} \log 0,0505 &< x \cdot \log 1,5 + y \cdot \log 0,5 < \log 95 \\ -1,2967 &< 0,1761 \cdot x - 0,301 \cdot y < 1,9777 \end{aligned}$$

$$\frac{-1,2967 - 0,1761 \cdot x}{-0,301} > y > \frac{1,9777 - 0,1761 \cdot x}{-0,301}$$

skąd ostatecznie

$$(21) \quad 0,58 \cdot x - 6,57 < y < 0,58 \cdot x + 4,31$$

Przekształcając podobnie nierówność (20) otrzymujemy kolejno

$$-1,2967 < -0,301 \cdot x + 0,1761 \cdot y < 1,9777$$

$$\frac{-1,2967 + 0,301 \cdot x}{0,1761} < y < \frac{1,9777 + 0,301 \cdot x}{0,1761}$$

skąd ostatecznie

$$(22) \quad 1,71 \cdot x - 7,36 < y < 1,71 \cdot x + 11,23$$

Przedstawimy teraz ten przykład postępowania sekwencyjnego graficznie w prostokątnym układzie współrzędnych. Liczbę pacjentów (par pacjentów) x , dla których lek K okazał się gorszy od leku L, odkładamy na osi poziomej; liczbę pacjentów (par pacjentów) y , dla których lek K okazał się lepszy od leku L, odkładamy na osi pionowej. W tym układzie współrzędnych test S_{-1} określony jest, zgodnie z (21), przez dwie proste równoległe o równaniach

$$(23) \quad y = 0,58 \cdot x - 6,57$$

oraz

$$(24) \quad y = 0,58 \cdot x + 4,31$$

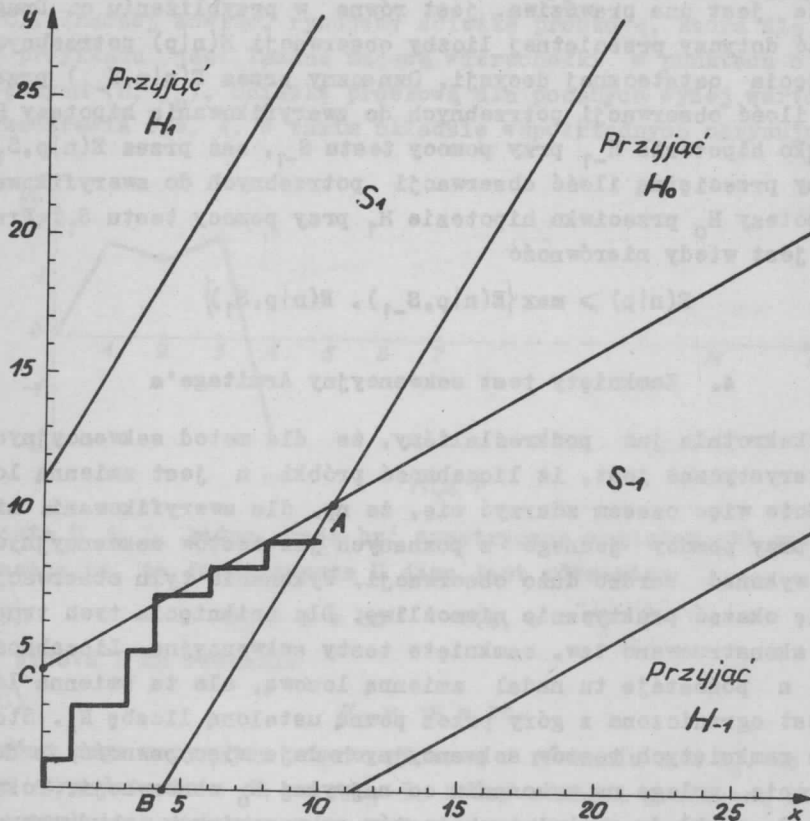
Podobnie test S_1 przedstawiony jest, zgodnie z (22), przez dwie proste równoległe o równaniach

$$(25) \quad y = 1,71 \cdot x - 7,36$$

oraz

$$(26) \quad y = 1,71 \cdot x + 11,23$$

Proste (23), (24) oraz (25) i (26) zaznaczone na rys. 3. Widać z tego rysunku, że oprócz "nożyc" wyznaczonych przez te dwie pary



Rys. 3

prostych znajdują się tam jeszcze 3 obszary. Skoro tylko ścieżka próbkowa znajdzie się w jednym z nich, przerywamy doświadczenie, a ostateczna decyzja polega na przyjęciu jednej z trzech hipotez w sposób wskazany na rys. 3. Zauważyć trzeba, że sposób wskazany na rys. 3 jest zgodny z procedurą podaną w tablicy 4. Zgodnie z tablicą 4 trzeba jednak zwrócić uwagę na jeden wyjątkowy przypadek. Mianowicie, jeśli ścieżka próbkowa (patrz rys. 3) przetnie i odcinek AB i odcinek AC, to natychmiast przerywamy doświadczenie, a ostateczna decyzja polega wtedy na przyjęciu hipotezy H_0 .

Niektóre własności testu. Własnościami opisanego postępowania sekwencyjnego, polegającego na równoczesnym i niezależnym operowaniu dwoma sekwencyjnymi testami ilorazowymi S_{-1} oraz S_1 , zajmowali się w pierwszym rzędzie Sobel i Wald [4]. Wymienimy tu tylko dwie bardzo proste własności. Jedną z nich stwierdza, że dla opisanego testu prawdopodobieństwo odrzucenia hipotezy H_0 pod warunkiem, że jest ona prawdziwa, jest równe w przybliżeniu α . Druga własność dotyczy przeciętnej liczby obserwacji $E(n|p)$ potrzebnych do podjęcia ostatecznej decyzji. Oznaczmy przez $E(n|p, S_{-1})$ przeciętną ilość obserwacji potrzebnych do zweryfikowania hipotezy H_0 przeciwko hipotezie H_{-1} przy pomocy testu S_{-1} , zaś przez $E(n|p, S_1)$ oznaczmy przeciętną ilość obserwacji potrzebnych do zweryfikowania hipotezy H_0 przeciwko hipotezie H_1 przy pomocy testu S_1 . Prawdziwa jest wtedy nierówność

$$(27) \quad E(n|p) \geq \max \{E(n|p, S_{-1}), E(n|p, S_1)\}$$

4. Zamknięty test sekwencyjny Armitage'a

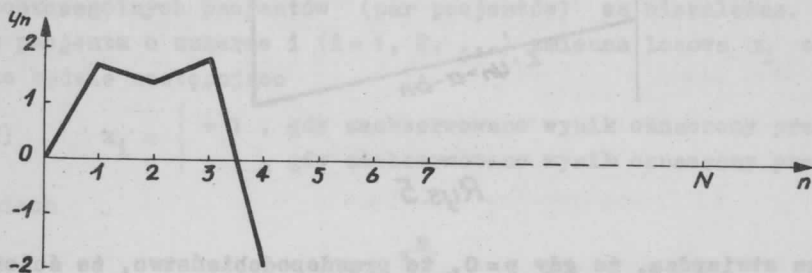
Kilkakrotnie już podkreślaliśmy, że dla metod sekwencyjnych charakterystyczne jest, iż liczebność próbki n jest zmienną losową. Może więc czasem zdarzyć się, że np. dla zweryfikowania hipotez przy pomocy jednego z poznanych już testów sekwencyjnych trzeba wykonać bardzo dużo obserwacji. Wykonanie tylu obserwacji może się okazać praktycznie niemożliwe. Dla uniknięcia tych trudności skonstruowano tzw. zamknięte testy sekwencyjne. Liczebność próbki n pozostaje tu nadal zmienną losową, ale ta zmienna losowa jest ograniczona z góry przez pewną ustaloną liczbę N_0 . Stosowanie zamkniętych testów sekwencyjnych daje więc pewność, że doświadczenie polega na wykonaniu co najwyżej N_0 obserwacji. Pokażemy tu 2 przykłady zamkniętych testów sekwencyjnych, zbudowanych w oparciu o propozycje Armitage'a.

Przykład A. Załóżmy, że badana cecha ma rozkład normalny ze znaną wariancją σ^2 , nieznana jest natomiast wartość oczekiwana μ . Załóżmy dalej, że dane są prawdopodobieństwa α oraz β . Na podstawie sekwencyjnie pobieranej próbki chcemy zweryfikować hipotezę H_0 , że $\mu = 0$, przeciwko hipotezie H_1 , że $|\mu| \geq k\sigma$, gdzie k jest pewną stałą dodatnią. Oznaczmy kolejne obserwacje przez $x_1, x_2, \dots, \dots, x_1, \dots$ i niech

$$(28) \quad y_n = \sum_{i=1}^n x_i,$$

gdzie n oznacza aktualną ilość obserwacji. Np. jeśli $x_1 = 1,7$; $x_2 = -0,3$; $x_3 = 0,4$; $x_4 = -3,7$; ... to zgodnie z (28) mamy $y_1 = 1,7$; $y_2 = 1,4$; $y_3 = 1,8$; $y_4 = -1,9$; ...

Narysujemy teraz prostokątny układ współrzędnych, w którym na osi poziomej odkładac będziemy ilość obserwacji n , natomiast na osi pionowej będziemy odkładać wartości y_n obliczone według wzoru (28). Inaczej mówiąc, rysujemy ścieżkę próbkową, która dla naszego przykładu jest łamaną mającą wierzchołki w punktach o współrzędnych (n, y_n) . Ścieżkę próbkową dla podanych wyżej wartości y_n przedstawia rys. 4. W takim układzie współrzędnych narysujemy dwie



Rys. 4

proste U i L, które mają być symetryczne względem osi poziomej. Oznacza to, że jeśli prosta U dana jest równaniem

$$(29) \quad y_n = a + bn \quad (a, b > 0)$$

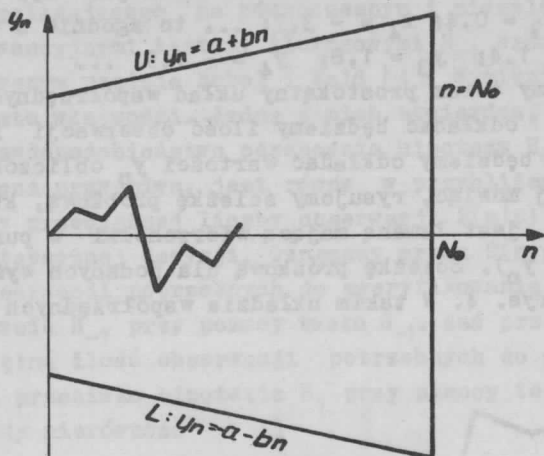
to prosta L ma równanie

$$(30) \quad y_n = -a - bn.$$

Oprócz tego narysujemy prostą pionową o równaniu $n = N_0$. Te 3 proste przedstawione są na rys. 5, który stanowi graficzną ilustrację zamkniętego testu sekwencyjnego Armitage'a.

Postępowanie przy użyciu zamkniętego testu sekwencyjnego Armitage'a można, posługując się rys. 5, opisać następująco. Jeśli ścieżka próbkowa znajduje się wewnątrz obszaru ograniczonego prostymi U i L oraz prostymi o równaniach $n = 0$ i $n = N_0$ to losujemy kolejny element do próbki. Jeśli ścieżka próbkowa osiągnęła bądź prostą U, bądź prostą L, to przyjmujemy hipotezę H_1 . Jeśli ścieżka próbkowa osiągnęła prostą $n = N_0$, to przyjmujemy hipotezę H_0 . Widać stąd, że dla tego testu liczebność próbki rzeczywiście może być co najwyżej N_0 .

Opisane tu postępowanie ma posiadać dwie własności. Własność



Rys. 5

pierwsza stwierdza, że gdy $p=0$, to prawdopodobieństwo, że ścieżka próbkowa osiągnie prostą U lub prostą L , ma być równe α . Oznacza to, że prawdopodobieństwo przyjęcia hipotezy H_0 wynosi wtedy $1 - 2\alpha$. Własność druga stwierdza, że gdy $p = k\sigma$ (bądź $-k\sigma$), to prawdopodobieństwo, że ścieżka próbkowa osiągnie wtedy U (bądź L), ma być równe $1 - \beta$.

Do wymienionych własności dołączmy jeszcze jeden warunek. Warunek ten mówi, że wzdłuż prostej U iloraz największej wiarygodności dla hipotez $p = k\sigma$ oraz $p = 0$ ma być wielkością stałą równą liczbie $(1 - \beta)/\alpha$. Wtedy łatwo pokazać, że współczynniki a i b w równaniach (29) i (30) mają wynosić

$$(31) \quad \begin{cases} a = \frac{\sigma}{k} \log \frac{1 - \beta}{\alpha} , \\ b = \frac{k\sigma}{2} \end{cases}$$

Tym samym wyznaczyliśmy już proste U oraz L , a pozostaje jeszcze wyznaczyć maksymalną liczbę obserwacji N_0 .

Z tymi samymi prawdopodobieństwami błędów 2α i β można hipotezę H_0 zweryfikować przeciwko hipotezie H_1 także przy pomocy klasycznego testu, który wymaga użycia próbki o z góry ustalonej liczebności. Oznaczmy liczebność takiej próbki przez N . Armitage pokazał, że gdy prawdopodobieństwa 2α i β zawarte są w przedziale

od 0,01 do 0,1 to wartość ilorazu N_0/N zmienia się od 1,25 do 1,55. W szczególności dla $2\alpha = \beta = 0,05$ mamy $N_0/N = 1,37$. Z tych wskaźników można w praktyce wyznaczyć w przybliżeniu liczbę N_0 na podstawie znajomości liczby N .

Przykład B. Przypuśćmy, że porównujemy skuteczność dwu leków K i L. Zakładamy, że dla każdego pacjenta (względnie dla każdej pary pacjentów) możliwe są tylko dwa wyniki badania: 1° lek K jest gorszy od leku L; 2° lek K jest lepszy od leku L. Zakładamy dalej, że dla każdego pacjenta (każdej pary pacjentów) prawdopodobieństwo wyniku oznaczonego przez 1° wynosi p , a prawdopodobieństwo wyniku oznaczonego przez 2° wynosi $1 - p$, przy czym wyniki u poszczególnych pacjentów (par pacjentów) są niezależne. Niech dla pacjenta o numerze i ($i = 1, 2, \dots$) zmienna losowa x_i określona będzie następująco

$$(32) \quad x_i = \begin{cases} +1, & \text{gdy zaobserwowano wynik oznaczony przez } 1^\circ \\ -1, & \text{gdy zaobserwowano wynik oznaczony przez } 2^\circ \end{cases}$$

i niech

$$(33) \quad y_n = \sum_{i=1}^n x_i,$$

gdzie n oznacza aktualną ilość zbadanych pacjentów (par pacjentów). Z określenia zmiennej losowej x_i łatwo zauważyć, że

$$(34) \quad |y_{n+1} - y_n| = 1.$$

Przy pomocy zamkniętego testu sekwencyjnego Armitage'a chcemy zweryfikować hipotezę H_0 , że skuteczność obu leków jest jednako-wa, przeciwko hipotezie H_1 , że prawdopodobieństwo, iż jeden z tych leków jest lepszy od drugiego, wynosi conajmniej p_1 , gdzie $p_1 > 0,5$. Mamy więc

$$H_0: p = 0,5, \quad H_1: p \geq p_1 \quad \text{albo} \quad 1 - p \geq p_1.$$

Dalej postępujemy już analogicznie jak w przykładzie A. Rysujemy prostokątny układ współrzędnych, w którym na osi poziomej odkładać będziemy ilość zbadanych pacjentów (par pacjentów) n , natomiast na osi pionowej będziemy odkładać wartości y_n obliczone według wzoru (33). Oprócz tego należy w tym układzie współrzędnych narysować proste U i L o równaniach (29) i (30) oraz prostą $n = N_0$. Dla rozpatrywanego przykładu współczynniki a i b w rów-

naniu prostych U i L wynoszą (por. np. [2])

$$(35) \quad \left\{ \begin{array}{l} a = \frac{2 \log \frac{1-\beta}{\alpha}}{\log \frac{p_1}{1-p_1}} \\ b = \frac{\log \frac{1}{4p_1(1-p_1)}}{\log \frac{p_1}{1-p_1}} \end{array} \right.$$

gdzie 2α oznacza prawdopodobieństwo błędu I rodzaju, a β oznacza prawdopodobieństwo błędu II rodzaju. Liczbę N_0 można wyznaczyć w oparciu o wskazówki zamieszczone na końcu przykładu A.

Dla ilustracji załóżmy, że $p_1 = 0,8$ oraz, że $2\alpha = \beta = 0,05$. Ze wzoru (35) mamy wtedy

$$(36) \quad a = \frac{2 \log \frac{1-0,05}{0,025}}{\log \frac{0,8}{1-0,8}} = \frac{2 \log 38}{\log 4} = 5,25$$

$$(37) \quad b = \frac{\log \frac{1}{4 \cdot 0,8(1-0,8)}}{\log \frac{0,8}{1-0,8}} = \frac{-\log 0,64}{\log 4} = 0,32$$

Stąd prosta U ma równanie

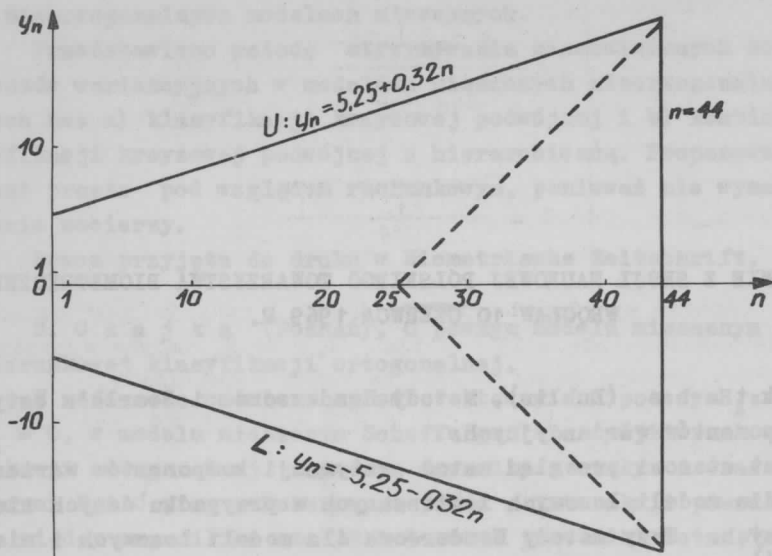
$$(38) \quad y_n = 5,25 + 0,32n,$$

zaś prosta L ma równanie

$$(39) \quad y_n = -5,25 - 0,32n.$$

Zweryfikowanie wymienionych hipotez na podstawie testu niesekwencyjnego wymaga użycia próbki o liczebności $N = 32$, skąd otrzymujemy, że $N_0 = 1,37 \cdot N = 1,37 \cdot 32 \approx 44$. Ten zamknięty test sekwencyjny, gdzie proste U i L mają odpowiednio równania (38) i (39), zaś $N_0 = 44$, przedstawiony jest na rys. 6.

Tutaj może się zdarzyć, że badanie przerywamy i przyjmujemy hipotezę H_0 , zanim jeszcze ścieżka próbkowa osiągnie prostą $n = N_0$. Mianowicie z (34) wynika, że badanie należy zakończyć i przyjąć hipotezę H_0 , skoro tylko ścieżka próbkowa osiągnie jeden z dwu odcinków, zaznaczonych na rys. 6 linią przerywaną.



Rys. 6

Literatura cytowana

- [1] Armitage, P., Sequential medical trials, Oxford 1960 (cytowane za [6]).
- [2] Billewicz, W. Z., Metody sekwencyjne w medycznych pracach badawczych; Zastosowanie metod matematycznych w medycynie, Wrocław-Warszawa-Kraków 1966, str. 5-35.
- [3] Fisz, M., Rachunek prawdopodobieństwa i statystyka matematyczna, Warszawa 1958.
- [4] Sobel, M., Wald, A., A sequential decision procedure for choosing one of three hypotheses concerning the unknown mean of a normal distribution, Annals of Mathematical Statistics, 20 (1949), str. 502-522.
- [5] Wald, A., Sequential analysis, New York 1947.
- [6] Wetherill, G.B., Sequential methods in statistics, London 1966.