Janusz L. Wywiał

PRÓBY LOSOWE W AUDYCIE FINANSOWYM

PRÓBY LOSOWE
W AUDYCIE FINANSOWYM

PODRĘCZNIK

Janusz L. Wywiał

PRÓBY LOSOWE W AUDYCIE FINANSOWYM

KATOWICE 2014

ISBN 978-83-7875-160-1

WYDAWNICTWO UNIWERSYTETU EKONOMICZNEGO W KATOWICACH

ul. 1 Maja 50, 40-287 Katowice, tel. +48 32 257-76-35, faks +48 32 257-76-43

www.wydawnictwo.ue.katowice.pl e-mail: wydawnictwo@ue.katowice.pl

Spis treści

WSTĘP

W ostatnich latach można zauważyć wzmożone zainteresowanie metodą reprezentacyjną, która znajduje zastosowanie w różnych obszarach badań naukowych, a w szczególności tradycyjnie w statystyce oficjalnej, badaniach opinii publicznej i marketingu, a także w rolnictwie. Nowe dziedziny korzystające z metody reprezentacyjnej, to badania przestrzenne, geologiczne, a także audytowe. Metody statystyczne są bardzo przydatne szczególnie podczas wnioskowania o wynikach kontroli bardzo licznych populacji dokumentów księgowych. Dotyczy to zwłaszcza technik losowania prób dokumentów księgowych, zeznań podatkowych itd. Przekonamy się dalej, że stosowane w audycie techniki losowania prób są w istocie niemalże powszechnie znanymi schematami losowania prób, których własności od dawna wykorzystuje się w metodzie reprezentacyjnej.

Metody statystyczne zwykle są stosowane w procedurach kontrolnych znanych w rewizji finansowej i do tego aspektu badań audytowych ograniczyno się w niniejszej pracy. Zrozumienie istoty metod statystycznych stosowanych w audycie finansowym wymaga wprowadzenia minimum potrzebnych metod doboru losowego prób oraz elementów wnioskowania statystycznego. Dlatego zaraz po krótkim naświetleniu problemów dotyczących badań audytowych wprowadzono podstawy schematów losowania prób z populacji skończonej i ustalonej. Potem wyłożono tylko te wiadomości z teorii estymacji i weryfikacji hipotez statystycznych, które będą potrzebne przy prezentacji metod podejmowania decyzji, będących etapem końcowym postępowania audytowego wykorzystującego losowane próby obiektów kontrolowanych. Od razu należy podkreślić, że niniejsza praca dotyczy zastosowania w audycie tzw. podejścia randomizacyjnego wyróżnianego w metodzie reprezentacyjnej. Charakteryzuje się tym, że wszelkie wnioskowanie statystyczne, czyli estymacja lub weryfikacja hipotez statystycznych, jest prowadzone na podstawie prób z ustalonych i skończonych populacji. Znane w literaturze tzw. podejście modelowe do wnioskowania statystycznego jest rzadziej stosowane w procesach rewizji finansowej i nie będzie tutaj rozwijane, bo wymagałoby to znacznego rozszerzenia wykorzystywanej dalej metodologii statystycznej.

W literaturze dotyczącej audytu finansowego dość jednolicie są wykładane

metody estymacji branych pod uwagę parametrów. W niniejszej pracy położono nacisk na stosowanie zasad weryfikacji hipotez statystycznych w postępowaniu prowadzącym do wyciągania wniosków z przeprowadzanych kontroli z wykorzystaniem danych z prób losowych. W szczególności zwracano uwagę na związki pomiędzy definiowanymi w teorii weryfikacji hipotez statystycznych pojęciami prawdopodobieństw popełnienia błędów I albo II rodzaju, a rodzajami ryzyka podejmowania decyzji w postępowaniu audytowym. Zawartość rozdziałów 2 i 3 stanowi zasadniczą treść niniejszej pracy. Obydwa dotyczą prezentacji podstawowych procedur używanych w audycie finansowym, lecz w kontekście podejścia randomizacyjnego metody reprezentacyjnej. Trzymając się definicji Karlińskiego (2005, s. 25), rozdział 2 dotyczy badania zgodności, czyli innymi słowy skuteczności systemu kontroli wewnętrznej, a w rozdziale 3 zajęto się badaniem wiarygodności np. ksiąg lub sprawozdań rachunkowych. W każdym z tych rozdziałów działanie prezentowanych procedur starano się ilustrować przykładami. Obliczenia prowadzono za pomocą specjalnie do tych celów stworzonych programów komputerowych napisanych w języku R, który jest coraz bardziej znany nie tylko wśród statystyków. Ponadto, ten język programowania jest powszechnie dostępny, a jego używanie jest bezpłatne.

Treść podręcznika jest efektem studiów m.in. podręczników dotyczących stosowania metod statystyczych w audycie autorstwa: Guy, Carmichael i Whittingston (2002), Hołdy i Pociechy (2004, 2009), Karlińskiego (2005), Klima (2005). Niniejsza praca ma za zadanie wzbogacić literaturę dotyczącą zastosowań statystyki w audycie. Powinna być przydatna dla studentów kierunków: Finanse i rachunkowość, Analityka gospodarcza lub Informatyka i ekonometria. Powinna być też użyteczna dla audytorów zajmujących się w praktyce rewizją finansową. W pracy zamieszczano oprogramowanie pozwalające na podejmowanie decyzji finalizujących prace kontrolne bez korzystania z tablic statystycznych, które są zamieszczane w innych podręcznikach dla audytorów. Prezentowane przykłady powinny ułatwić korzystanie z tego oprogramowania.

Bardzo dziekuję Panu profesorowi dr. hab. Stanisławowi Heilpernowi za wnikliwą recenzję niniejszej pracy.

Rozdział 1

WPROWADZENIE DO WNIOSKOWANIA STATYSTYCZNEGO O CHARAKTERYSTYKACH DOKUMENTACJI RACHUNKOWEJ

Wiadomo, że studenci drugiego roku kierunków ekonomicznych mają za sobą podstawowy kurs statystyki, który oprócz statystyki opisowej powinien także obejmować elementy wnioskowania statystycznego. Dlatego w niniejszym rozdziale ograniczymy się do przypomnienia podstaw wnioskowania statystycznego z jednoczesnym nawiązaniem do problemów audytu finanasowego. Zwłaszcza chodzi tutaj o uświadomienie sobie faktu, że pewne pojęcia występujące w audycie mają swoje odpowiedniki w teorii wnioskowania statystycznego. W różnych sytuacjach poziom istotności testu statystycznego raz może być traktowany jako ryzyko aprobaty nieskutecznej kontroli, a z innego punktu widzenia jako ryzyko dezaprobaty skutecznego systemu kontroli.

1.1. Rodzaje ryzyka w audycie

Problem ryzyka w badaniach audytowych jest jednym z kluczowych zagadnień prawidłowego przeprowadzenia postępowania sprawdzającego szeroko rozumiane czynności księgowe. Różnorodne podejścia dotyczące tego problemu są prezentowane m.in. w pracy Pfaffa (2008). W niniejszym opracowaniu zostanie nakreślony problem ryzyka z punktu widzenia stosowanych w audycie metod statystycznych.

Ryzyko związane ze stosowaniem procedur statystycznego wnioskowania na tle innych czynników ryzyka jest następujące. *Ryzyko postępowania (badania) audytowego* jest definiowane jako iloczyn:

$$r_b = r_n r_k r_s, \tag{1.1}$$

gdzie przez r_n, oznaczono tzw. *ryzyko nieodłączne (inherent risk)*, które charakteryzuje specyficzne działania księgowe. Z kolei r_k jest *ryzykiem za-*

wodności systemów kontroli wewnętrznej (control risk). W końcu przez r_n oznaczono tzw. *ryzyko niewykrycia (przeoczenia)* przez kontrolera *(detection risk)*, przy czym

$$r_n = r_p r_s. \tag{1.2}$$

Symbolem r_p oznaczono *ryzyko spowodowane wykorzystaniem dodatkowych procedur kontrolnych (audit procedures)* niezwiązanych z losowaniem prób. Z kolei r_s jest *ryzykiem statystycznym*, które najbardziej interesuje nas w niniejszej pracy. W istocie ryzyko $r_s = \kappa$ jest prawdopodobieństwem popełnienia tzw. *błędu II rodzaju* κ polegającego na akceptacji występujących nieprawidłowości *(risk of incorrect acceptance).* κ jest w szczególności ryzykiem błędnego wydania opinii pozytywnej o systemie kontroli wewnętrznej, który działa nieprawidłowo. W innym przypadku κ jest ryzykiem błędnej akceptacji sprawozdania finansowego (np. bilansu) obarczonego niedopuszczalnym błędem.

Ze wzorów (1.1) i (1.2) wynika

$$r_b = r_n r_k r_p r_s,$$

co prowadzi do wyrażenia:

$$r_s = \kappa = \frac{r_b}{r_n r_k r_p}. \tag{1.3}$$

Sposoby ustalania w praktyce poziomów opisanych ryzyk zależą od specyfiki kontrolowanych obiektów oraz od wiedzy i doświadczenia audytora. W wielu pracach spotykamy się z różnymi zaleceniami, dotyczącymi sposobów wyznaczania poziomów tych ryzyk. Owczarek (1996) pisze, iż zwykle przyjmuje się, że

$$r_b = 0.05, \qquad r_n = 1.00, \qquad r_p \in [0.50;\ 1.00],$$

$r_k \in [0.01;\ 0.10]$, gdy kontrola wewnetrzna jest bardzo dobra,
$r_k \in (0.10;\ 0.30]$, jeśli kontrola jest dobra,
$r_k \in (0.30;\ 0.50]$, jeśli kontrola jest wystarczajca,
$r_k \in (0.50;\ 0.70]$, gdy kontrola jest słaba,
$r_k \in (0.70;\ 1.00]$, jeżeli kontrola jest żadna (niedostateczna).

Na przykład gdy $r_b = 0.05, \quad r_n = r_k = 1, \quad r_p = 0.5$, to $\kappa = r_s = 0.1$. Jeśli $r_b = 0.05, \quad r_n = 1, \quad r_p = r_k = 0.5$, to $\kappa = r_s = 0.2$.

Z przedstawionego krótkiego opisu czynników składających się na ryzyko badania audytowego wynika, że nie ma ścisłych zasad ustalania ich poziomów. Mamy tutaj do czynienia z dość dużym subiektywizmem, ponieważ każdy audytor, oprócz wiedzy i doświadczenia, ma swoje skłonności do większego lub mniejszego optymizmu przy ustalaniu poziomów wyróżnionych ryzyk. W konsekwencji dotyczy to również ryzyka statystycznego r_s wyznaczanego ze wzoru (1.3). Jak się jednak okaże w dalszej treści pracy

i co należy podkreślić, procedury wnioskowania statystycznego pozwalają podejmować decyzje ściśle zachowujące ustalony poziom ryzyka statystycznego. Innymi słowy, po ustaleniu ryzyka r_s dalsze postępowanie jest ścisłe, co gwarantują znane własności metod statystyki matematycznej. Tu nie ma już miejsca na subiektywizm badacza.

Rozważany model ryzyka jest uproszczoną wersją znanego w auditingu tzw. podejścia probabilistycznego. W literaturze przedmiotu są jeszcze brane pod uwagę podejścia lingwistycze oraz tzw. teoria Dempstera–Shafera, które są analizowane m.in. przez Hołdę i Pociechę (2009).

1.2. Populacja i parametry zmiennych

W rachunkowości mamy głównie do czynienia z populacjami tradycyjnych dokumentów papierowych lub dokumentów zapisanych na odpowiednich nośnikach elektronicznych. Populację będziemy oznaczać przez U. Liczbę elementów tworzących populację U (liczebność) oznaczamy przez N. Każdy element tego typu populacji ma swój numer identyfikacyjny, którym może być np. nr faktury. Mamy tutaj zatem do czynienia z tzw. *populacjami identyfkowalnymi*, co znacznie ułatwi losowanie prób. Identyfikowalność populacji w praktyce oznacza, że dysponujemy listą elementów populacji, co pozwala nadać każdemu elementowi populacji w sposób wzajemnie jednoznaczny liczbę naturalną. Taki spis elementów ewentualnie rozbudowany o dodatkowe informacje o strukturze wartości tzw. zmiennych pomocniczych w populacji nazywamy *operatem losowania*.

Każdy dokument ma zwykle kilka zapisów, czyli na fakturze mamy wyszczególnione oprócz jej numeru kwotę netto sprzedawanego towaru, kwotę brutto, ewentualny rabat, ilość towaru, jego cenę jednostkową itd. Te zapisy będziemy traktowali jako wartości odpowiednich *zmiennych (cech)* charakteryzujących dokumenty. W szczególności niech x i y będą zmiennymi oznaczającymi odpowiednio ilość sprzedanego przez rolników zboża i jego wartość. Z kolei przez x_i oraz y_i oznaczamy odpowiednio ilość oraz wartość zboża sprzedanego przez ustalonego i-tego rolnika, $i = 1, \ldots, N$. Krótko mówiąc, para (x_i, y_i) jest wartością (obserwacją) dwuwymiarowej zmiennej (x, y). Wektor wszystkich wartości (obserwacji) zmiennej x oznaczamy przez $\mathbf{x} = [x_1\ x_2 \ldots x_N]$. Ten wektor jest również nazywany parametrem populacji. Zakładamy, że wartości zmiennych (parametrów, cech) są liczbami ze zbioru liczb rzeczywistych, czyli $\mathbf{x} \in R^N$. Z kolei wartości dwuwymiarowej zmiennej (x, y) możemy zapisać jako macierz (parametr) o wymiarach $2 \times N$ itd.

Przykład 1.1

W przypadku audytu finansowego, dysponujemy nominalnymi obserwacjami wartości faktur, które oznaczamy przez x_i, $i = 1, \ldots, N$. Każda

faktura ma swój numer i. Zatem operat losowania składa się z ciągu par wartości (i, x_i), $i = 1, \ldots, N$, gdzie i jest kolejnym numerem faktury, a x_i jest nominalną obserwacją jej wartości.

Na podstawie dokumentów wyliczane są np. sumy ich określonych pozycji księgowych, czyli formalnie rzecz biorąc wartość sumy

$$x_U = \sum_{i=1}^{N} x_i,$$

która jest nazywana *wartością globalną* lub *wartością łączną* zmiennej x.

Wartość przeciętną zmiennej x oznaczamy przez

$$\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i = \frac{x_U}{N}.$$

Przypuśćmy, że badana cecha dokumentu może być typu dychotomicznego, co w szczególności oznacza, iż zmienna x jest zero-jedynkowa, czyli $x = 0$ albo $x = 1$. Przykładowo, gdy rachunek dotyczy zapłaty za pracę akordową, to $x = 0$, a gdy za inny rodzaj pracy, to $x = 1$. *Liczbę elementów populacji z cechą wyróżnioną*, czyli np. w naszym przypadku liczbę błędnych dokumentów oznaczymy przez $m_{U,x}$ i wyliczymy ze wzoru:

$$m_{U,x} = x_U = \sum_{i=1}^{N} x_i.$$

Częstość względną występowania cechy wyróżnionej w populacji wyliczamy ze wzoru:

$$p_{U,x} = \frac{1}{N} \sum_{i=1}^{N} x_i = \frac{m_{U,x}}{N}.$$

Wariancję zmiennej w populacji pozwala wyliczyć wzór:

$$v_{*U,x} = \frac{\sum_{k \in U} (x_k - \bar{x})^2}{N - 1}$$

Określone symbolami x_U, $\bar{x}$, $m_{U,x}$, $p_{U,x}$ i $v_{U,x}$ funkcje wartości cechy x (parametru populacji) nazywamy *opisowymi parametrami populacji* lub krótko parametrami opisowymi.

Zauważmy, że z zapisanych wzorów wynikają następujące, jak okaże się, pożyteczne tożsamości.

$$x_U = N\bar{x} \qquad m_{U,x} = Np_{U,x}. \tag{1.4}$$

Przykład 1.2

Załóżmy, że 5 faktur ma następujące wartości w zł: $x_1 = 100, x_2 = 281, x_3 = 29, x_4 = 150, x_5 = 20$. Zatem populacja składa się z $N = 5$ elementów. W wyniku przeprowadzonego pełnego audytu uzyskano następujące poprawione, czyli prawdziwe wartości tych dokumentów: $y_1 = 100, y_2 = 275, y_3 = 25, y_4 = 150, y_5 = 20$. Populacja jest więc identyfikowalna, bowiem każdej fakturze jest wzajemnie jednoznacznie przypisany jej numer. Ponadto, z góry są znane wartości nominalne faktur, które wraz z jej numerem tworzą operat losowania próby. Wektory $\mathbf{x} = [100\ 281\ 29\ 150\ 20]$ i $\mathbf{y} = [100\ 275\ 25\ 150\ 20]$ są parametrami naszej populacji lub innymi słowy wartościami zmiennych (cech) odpowiednio x i y. Stąd wynika, że wektor błędów w wartościach faktur ma postać: $\mathbf{e} = \mathbf{x} - \mathbf{y} = [0\ 6\ 4\ 0\ 0]$. Z kolei wartości zero-jedynkowej zmiennej identyfikującej występowanie błędu oznaczamy przez $\mathbf{z} = [0\ 1\ 1\ 0\ 0]$. Stąd wyliczamy następujące parametry opisowe populacji. Nominalna wartość globalna faktur wynosi $x_U = 580$ zł, a prawdziwa wartość globalna faktur wynosi $y_U = 570$ zł, natomiast wartość globalna błędów jest równa $e_U = x_U - y_U = 10$ zł. Średnia wartość nominalna faktury wynosi $\bar{x} = 116$ zł, natomiast przeciętna prawdziwych wartości faktur jest równa $\bar{y} = 114$ zł. Średnia błędów w wartościach dokumentów wynosi $\bar{e} = 2$ zł. Z kolei liczba błędnych faktur to $m_{z,U} = 2$, a częstość ich występowania jest równa $p_{z,U} = 0.4$.

Zmienna zero-jedynkowa z gra rolę zmiennej diagnostycznej, czyli identyfikuje nieprawidłowości, gdy jest równa jedności, a ich brak, jeśli jest równa zeru.

Inny ważny problem związany z badaniem tzw. wiarygodności sprawozdań wymaga pewnych założeń modelowych generowania się błędów w wartościach np. pozycji księgowych. Niech x będzie nominalną pozycją księgową (np. faktury, zadeklarowanej kwoty podatku), czyli x_k, $k = 1, \ldots, N$ jest zaobserwowaną wartością pozycji księgowej (np. wpłaconą kwotą podatku). Z kolei przez y oznaczymy zmienną, której wartość y_k, $k = 1, \ldots, N$ jest prawdziwą wartością pozycji księgowej, czyli skorygowaną w wyniku przeprowadzonego audytu. Zatem błąd określa różnica: $e_k = x_k - y_k$. Wówczas związek między zmiennymi można modelować następująco:

$$x_k = y_k + e_k, \qquad k = 1, \ldots, N. \tag{1.5}$$

Parametry tego addytywnego modelu są następujące:

$$v_{*U,x} = v_{*U,y} + v_{*U,e} + 2v_{*U,y,e}, \tag{1.6}$$

gdzie $v_{*U,y}$, $v_{*U,x}$, $v_{*U,e}$ oraz

$$v_{*U,y,e} = \frac{1}{N-1} \sum_{k \in U} (y_k - \bar{y})(e_k - \bar{e}) \tag{1.7}$$

są odpowiednio wariancjami zmiennych y, x, e oraz kowariancją zmiennych y i e. Wartości zmiennej e są błędami addytywnie nakładającymi się na odpowiednie wartości zmiennej y. Jeśli zmienne y i e nie są skorelowane, to

$$v_{*U,x} = v_{*U,y} + v_{*U,e}. \tag{1.8}$$

Z kolei, gdy zmienne y i e są skorelowane dodatnio, to

$$v_{*U,x} > v_{*U,y} + v_{*U,e}, \tag{1.9}$$

W końcu, gdy zmienne y i e są skorelowane ujemnie, to

$$v_{*U,x} < v_{*U,y} + v_{*U,e}. \tag{1.10}$$

Można łatwo pokazać, że współczynnik korelacji zmiennych x i y ma postać:

$$r_{x,y} = \frac{v_{*U,x,y}}{\sqrt{v_{*U,x} v_{*U,y}}} = \frac{v_{*U,y} + v_{*U,y,e}}{\sqrt{v_{*U,y}(v_{*U,y} + v_{*U,e} + 2v_{*U,y,e})}}. \tag{1.11}$$

Parametr $r_{x,y}$ jest dodatni, gdy zmienne y i e nie są ujemnie skorelowane. W szczególności, jeśli y i e nie są skorelowane, to

$$r_{x,y} = r = \sqrt{\frac{v_{*U,y}}{v_{*U,y} + v_{*U,e}}}. \tag{1.12}$$

Ten wskaźnik jest nazywany współczynnikiem rzetelności obserwacji x zmiennej y (por. np. Guilford, 1960). Bliskie jedności wartości współczynnika r świadczą o wysokiej jakości obserwacji zmiennej, ponieważ wariancja błędów obserwacji jest mała w stosunku do zmiennej y. Trzeba jednak zwrócić uwagę na szczególny przypadek, gdy $r = 1$, co bynajmniej nie świadczy o braku występowania błędów, ponieważ taka sytuacja zachodzi, gdy niezerowy błąd obserwacji jest taki sam dla wszystkich wartości y.

Multiplikatywny model generowania się obserwacji ma postać:

$$x_k = y_k u_k, \qquad k = 1, \ldots, N, \tag{1.13}$$

przy czym $u_k > 0$. Ten model można sprowadzić do następującej postaci addytywnej, przyjmując, iż $u_k = 1 + z_k$. Wówczas:

$$x_k = y_k(1 + z_k) = y_k + y_k z_k, \qquad k = 1, \ldots, N, \tag{1.14}$$

przy czym wobec założenia, że wartości zmiennej y są nieujemne, należy postulować, aby $z \geq -1$. W tym przypadku wartość błędu $e_k = y_k z_k$ jest proporcjonalna do wartości y_k.

1.3. Metody losowania próby

W praktyce badań statystycznych jest popularne pojęcie próby losowanej bezzwrotnie jako n-elementowego podzbioru s populacji U, czyli $s \subseteq U$. Z kombinatorycznego punktu widzenia próba s jest n-elementową kombinacją bez powtórzeń, wybraną ze zbioru N-elementowego. Zatem zbiór $\mathbf{S}$ wszystkich możliwych do wybrania z populacji U prób wynosi $\binom{N}{n}$. Zbiór $\mathbf{S}$ jest zwany również *przestrzenią prób*.

Definicja 1.1

Plan losowania próby $P(s)$ jest następującym rozkładem prawdopodobieństwa określonym na zbiorze (przestrzeni) prób $\mathbf{S}$, przy czym

$$0 \leq P(s) \leq 1 \quad \text{dla każdej} \quad s \in \mathbf{S}, \quad \text{oraz} \quad \sum_{s \in \mathbf{S}} P(s) = 1. \tag{1.15}$$

Zatem plan losowania próby przyporządkowywuje każdej możliwej do wylosowania próbie pewną wartość prawdopodobieństwa. W szczególności próba s nie będzie losowana z populacji, gdy $P(s) = 0$. Jeśli $P(s) = 1$, to z populacji jest losowana tylko próba s, która jest nazywana *próbą celową*. Próba celowa występuje w praktyce, gdy audytor wskazuje na podstwie swojej wiedzy lub doświadczenia taką próbę, uważając, że ona właśnie najlepiej reprezentuje populację. Konstrukcje planów losowania mogą być dowolne, jakkolwiek istnieją pewne racjonalne przesłanki ich definiowania, o czym będzie mowa dalej.

Przykład 1.3

Załóżmy, że populacja składa się z $N = 4$ elementów i oznaczmy ją zbiorem: $U = \{1, 2, 3, 4\}$. Wszystkie możliwe do wybrania próby dwuelementowe tworzą zbiór:

$$\mathbf{S} = \{(1,2), (1,3), (1,4), (2,3), (2,4), (3,4)\}.$$

Określmy następujący plan losowania: $P(1,2) = 0.05$, $P(1,3) = 0.05$, $P(1,4) = 0.01$, $P(2,3) = 0.09$, $P(2,4) = 0.2$, $P(3,4) = 0.6$.

Prawdopodobieństwo inkluzji rzędu pierwszego dla k-tego elementu populacji jest prawdopodobieństwem wylosowania tego elementu do próby, czyli:

$$\pi_k = \sum_{\{s: k \in s\}} P(s).$$

Prawdopodobieństwo inkluzji rzędu drugiego dla k-tego oraz t-ego elementu populacji jest prawdopodobieństwem wylosowania tych elementów do próby, czyli:

$$\pi_{k,t} = \pi_{t,k} = \sum_{\{s: k \in s, t \in s\}} P(s).$$

Jeśli próba o ustalonym rozmiarze n jest losowana bezzwrotnie, to

$$\sum_{k=1}^{N} \pi_k = n, \qquad \sum_{k=1}^{N} \sum_{t=1,t\neq k}^{N} \pi_{k,t} = n(n-1). \tag{1.16}$$

Przykład 1.4

Niech plan losowania próby trójelementowej z populacji 5-elementowej ma postać:

$$P(1,2,3) = P(1,2,4) = P(1,3,4) = P(1,3,5) = P(2,3,5) = P(2,4,5) = 0,$$

$$P(1,2,5) = 0.3,\ P(1,4,5) = 0.2,\ P(2,3,4) = 0.1,\ P(3,4,5) = 0.4.$$

Wówczas, prawdopodobieństwa inkluzji rzędu pierwszego mają postać:

$$\pi_1 = 0.5,\ \pi_2 = 0.4,\ \pi_3 = 0.5,\ \pi_4 = 0.7,\ \pi_5 = 0.9,$$

przy czym np.

$$\pi_1 = P(1,2,3) + P(1,2,4) + P(1,2,5) + P(1,3,4) + P(1,3,5) + P(1,4,5) =$$

$$= 0 + P(1,2,5) + P(1,4,5) = 0.3 + 0.2 = 0.5.$$

Z kolei prawdopodobieństwa inkluzji rzędu drugiego są następujące:

$$\pi_{1,2} = 0.3,\ \pi_{1,3} = 0,\ \pi_{1,4} = 0.2,\ \pi_{1,5} = 0.5,\ \pi_{2,3} = 0.1,\ \pi_{2,4} = 0.1,$$

$$\pi_{2,5} = 0.3,\ \pi_{3,4} = 0.5,\ \pi_{3,5} = 0.4,\ \pi_{4,5} = 0.6,$$

przy czym np.

$$\pi_{4,5} = P(1,2,5) + P(1,3,5) + P(1,4,5) + P(2,3,5) + P(2,4,5) + P(3,4,5) =$$

$$= P(1,4,5) + P(3,4,5) = 0.2 + 0.4 = 0.6.$$

Sprawdzamy, że

$$\sum_{k=1}^{5} \pi_k = 3, \qquad \sum_{k=1}^{5} \sum_{t=1,t\neq k}^{5} \pi_{k,t} = 6.$$

W sensie technicznym, po to aby dobrać kolejno elementy populacji do próby zgodnie z założonym planem losowania, trzeba określić *schemat losowania próby*. Innymi słowy schemat losowania próby jest pewnym algorytmem doboru elementów populacji do próby, realizującym postulowany plan jej losowania.

1.3.1. Próba prosta

Powszechnie znane określenie próby prostej jest następujące.

Definicja 1.2

Prosta próba o liczebności n losowana bezzwrotnie jest definiowana wyrażeniem:

$$P_{bz}(s) = 1/\binom{N}{n}, \quad \text{gdzie} \quad s \in \mathbf{S}.$$

Parametry tego planu są następujące: $\pi_k = \frac{n}{N}$, $\pi_{k,t} = \frac{n(n-1)}{N(N-1)}$ dla $k = 1, \ldots, N$, $t = 1, \ldots, N$, $k \neq t$.

Zatem, każda możliwa do wybrania z populacji próba prosta jest losowana z tym samym prawdopodobieństwem. Podobnie, prawdopodobieństwa inkluzji rzędu pierwszego są takie same. To również dotyczy prawdopodobieństwa inkluzji rzędu drugiego.

Algorytm (schemat) losowania rozważanej próby prostej jest następujący:

1. Pierwszy element próby jest losowany z populacji z prawdopodobieństwem $p_1 = \frac{1}{N}$, czyli np. należy wygenerować liczbę losową u o rozkładzie jednostajnym określonym na przedziale $[0;\ 1]$ i wtedy jest wybierany j-ty element populacji, gdy $\frac{j-1}{N} \leq u < \frac{j}{N}$, $j = 1, \ldots, N$.
2. $k-$ty element próby jest podobnie losowany z populacji pomniejszonej o już wcześniej wylosowane z niej $(k-1)$ elementy z prawdopodobieństwem $p_k = \frac{1}{N-k+1}$, $k = 2, \ldots, n$.

Przykład 1.5

Gdy $N = 5$ i $n = 2$, to plan bezzwrotnego losowania próby prostej określa wyrażenie: $P_{bz}(s) = 0.1$, gdzie jego prawdopodobieństwa inkluzji są postaci: $\pi_k = 0.4$, $\pi_{k,t} = 0.1$ dla $k = 1, \ldots, 5$, $t = 1, \ldots, 5$, $k \neq t$.

Algorytm losowania rozważanej próby jest następujący. Pierwszy element próby należy losować z prawdopodobieństwem $p_1 = 0.2$ z całej pięcioelementowej populacji, a drugi element wybieramy z prawdopodobieństwem $q_2 = 0.25$ z populacji czteroelementowej, czyli pomniejszonej o już wcześniej wylosowany element.

Definicja 1.3

Prosta próba o liczebności n losowana zwrotnie jest definiowana wyrażeniem (por. Tillé, 2006, s. 54):

$$P_z(\mathbf{s}) = \frac{n!}{N^n \prod_{k \in U} n_k!}, \quad \text{gdzie} \quad s \in \underline{\mathbf{S}}, \quad 0 \leq n_k \leq n, \quad \sum_{k=1}^{N} n_k = n$$

przy czym n_k jest krotnością występowania k-tego elementu populacji w próbie s, natomiast $\underline{\mathbf{S}}$ jest zbiorem wszystkich możliwych prób n-elementowych, w których elementy populacji mogą wielokrotnie występować.

Prawdopodobieństwa inkluzji tego planu określają wzory: $\pi_k = 1 - \left(\frac{N-1}{N}\right)^n$, $\pi_{k,t} = 1-2\left(\frac{N-1}{N}\right)^n+\left(\frac{N-2}{N}\right)^n$, przy czym $k = 1, \ldots, N, t = 1, \ldots, N$, $k \neq t$.

Schemat losowania rozważanej próby jest następujący. Każdy element populacji jest losowany zwrotnie z prawdopodobieństwem $p = 1/N$ do próby $n-$krotnie. Zatem elementy populacji są losowane do próby niezależnie.

Przykład 1.6

Gdy $N = 4$ i $n = 2$, to plan zwrotnego losowania próby prostej ma postać: $P_z(\mathbf{s}) = 0.1$, gdzie $\mathbf{s} \in \underline{\mathbf{S}}$. Parametry tego planu są następujące: $\pi_k = 7/16$, $\pi_{k,t} = 1/8$, $k = 1, \ldots, 4$, $t = 1, \ldots, 4$, $k \neq t$.

Po to aby wylosować rozważaną próbę, należy z całej populacji wylosować (zwrotnie) dwa elementy, przy czym za każdym razem prawdopodobieństwo wylosowania elementu populacji ma wynosić 0.25.

Teraz rozważymy najprostszy przypadek systematycznego losowania próby, gdy jej liczebność n jest liczbą naturalną otrzymaną z ilorazu N/q, gdzie parametr q jest nazywany odstępem (interwałem) losowania systematycznego. Schemat losowania próby s polega na tym, że spośród $q-$pierwszych elementów populacji losuje się jeden z nich, a następnie wybiera się te elementy, których numery są odległe od numeru pierwszego wylosowanego elementu o krotności liczby q. Jeśli np. wylosowano i_e-ty element populacji spośród $q-$pierwszych elementów, wówczas do próby s_e wejdą elementy o numerach $i_e, i_{e+q}, i_{e+2q}, \ldots, i_{e+tq}, \ldots i_{e+(n-1)q}$ lub w zapisie bardziej syntetycznym $s_e = (i_{e+eq}; e = 0, 1, \ldots, n-1)$. Przestrzeń wszystkich możliwych do wylosowania prób składa się z $q-$prób tworzących zbiór $\mathbf{S} = \{s_1, \ldots, s_q\}$. W związku z tym mamy następujące określenie.

Definicja 1.4

Plan losowania prostej próby systematycznej określa wyrażenie:

$$P_{sys}(s) = \frac{1}{q} \qquad dla \qquad s \in \mathbf{S}.$$

Prawdopodobieństwa inkluzji rzędu pierwszego i drugiego są takie same i wynoszą:

$$\pi_i = \frac{1}{q} \qquad dla \qquad i = 1, \ldots, N,$$

$$\pi_{i,j} = \begin{cases} \frac{1}{q} & dla \quad i \neq j, \quad i \in s, \quad j \in s, \\ 0 & dla \quad i \neq j, \quad i \notin s, \quad lub \quad j \notin s. \end{cases}$$

Algorytm schematu losowania próby prostej systematycznej o liczebności n, z interwałem losowania q jest następujący. Należy wygenerować liczbę losową u z rozkładu jednostajnego prawdopodobieństwa określonego na przedziale $[0;\ 1]$. Próbę wyznacza ciąg numerów elementów w populacji: $s =$

$(i_1, i_1 + q, \ldots, i_1 + (n-1)q)$, przy czym element populacji o numerze i_1 wyznacza nierówność:

$$\frac{i_1 - 1}{q} < u \leq \frac{i_1}{q}. \tag{1.17}$$

Przykład 1.7

Gdy $N = 12$, $n = 3$, $p_k = 1/4$. Wtedy interwał losowania systematycznego prostego wynosi $q = \frac{N}{n} = 4$ Niech liczba losowa o rozkładzie jednostajnym na przedziale $[0, 1]$ wynosi $u = 0.62$. Wówczas powyższy algorytm prowadzi do wyboru próby $s = (3, 7, 11)$, gdyż dla pierwszego elementu próby, będącego jednocześnie elementem poplacji o numerze $i_1 = 3$ zachodzi nierówność:

$$\frac{2}{q} = 0.5 < u = 0.62 < 0.75 = \frac{3}{q}.$$

1.3.2. Próba monetarna

Specyfiką badań audytowych, zwłasza wiarygodności np. bilansów, jest fakt, że zwykle jest nieograniczony dostęp do nominalnych wartości dokumentów księgowych, bowiem zazwyczaj np. wartości faktur są przechowywane wraz z jej numerem (identyfikatorem) w jednym miejscu, którym może być dysk magnetyczny. Te informacje, jak się okaże, można w sposób efektywny wykorzystać do podniesienia dokładności wnioskowania statystycznego związanego z zadaniami kontrolnymi.

Zmienną, której wartości są z góry znane w populacji dokumentów, np. wspomniane wartości nominalne faktur, oznaczymy przez x, a jej wartości przez x_k, $k = 1, \ldots, N$. Jednym ze sposobów wykorzystania tej danej jest konstrukcja schematu losowania próby, który jak to okaże się później, może znacznie podnieść m.in. precyzję estymacji charakterystyk rozkładu skorygowanych wartości faktur.

Postuluje się, aby elementy populacji losować do próby z prawdopodobieństwami inkluzji rzędu pierwszego, proporcjonalnymi do wartości nominalnych przypisanych dokumentom, czyli aby $\pi_k \propto x_k$, przy czym $x_k > 0$ dla $k = 1, \ldots, N$.

Wtedy zgodnie z żądaniem, aby próba była losowana bezzwrotnie i miała ustaloną liczebność, powinna być spełniona własność: $\sum_{k=1}^{N} \pi_k = n$. Zatem należałoby przyjąć, że:

$$\pi_k^{(0)} = \frac{n x_k}{\sum_{i=1}^{N} x_i}, \qquad k = 1, \ldots, N. \tag{1.18}$$

Nie da się jednak wykluczyć, że dla pewnego e wartość x_e jest na tyle duża, iż $\pi_h > 1$. Wówczas przyjmuje się (por. np. Tillé, 2006 lub Ardilly i Tillé, 2006), że $\pi_h = 1$. To oznacza, że $h-$ty element populacji będzie celowo dobierany do próby, a prawdopodobieństwa doboru pozostałych elementów

należy skorygować. W tej sytuacji algorytm postępowania jest następujący. Wyznaczamy taki wskaźnik h spośród $k = 1, \ldots, N$, że

$$\pi_h^{(0)} = \text{maximum}_{k \in U}\{\pi_k^{(0)}\},$$

gdzie, przypomnijmy, przez U oznaczono populację, a dokładnie zbiór numerów elementów populacji. Jeśli $\pi_h^{(0)} \leq 1$, to przyjmujemy, że $\pi_k = \pi_k^{(0)}$, dla $k = 1, \ldots, N$ i to już koniec algorytmu wyznaczania wartości prawdopodobieństw inkluzji.

W przypadku gdy $\pi_h^{(0)} > 1$, przyjmujemy, że $\pi_h = 1$ i wyznaczamy następujące ilorazy:

$$\pi_k^{(1)} = \frac{(n-1)x_k}{\sum_{i=1}^{N} x_i - x_h}, \qquad k = 1, \ldots, N; \quad k \neq h.$$

Potem znów wyznaczamy wskaźnik g, tak aby

$$\pi_g^{(1)} = \text{maximum}_{k \in \{U-h\}}\{\pi_k^{(1)}\}$$

Jeśli $\pi_g^{(1)} \leq 1$, to przyjmujemy, że $\pi_h = 1$, $\pi_k = \pi_k^{(1)}$ dla $k = 1, \ldots, N$, przy czym $k \neq h$. To kończy algorytm wyznaczania wartości prawdopodobieństw inkluzji.

W przeciwnym przypadku, jeśli $\pi_g^{(1)} > 1$, to przyjmujemy, że $\pi_h = 1$, $\pi_g = 1$ i ponownie wyznaczamy odpowiednie ilorazy.

Uogólniając przedstawione rozumowanie, przypuśćmy, że mamy do czynienia z $e-$tym etapem algorytmu, przy czym $e = 0, 1, \ldots, r < n$, $\pi_{k_i} = 1$ dla $i = 1, \ldots, e-1$ oraz

$$\pi_k^{(e)} = \frac{(n-e)x_k}{\sum_{i=1}^{N} x_i - \sum_{h=1}^{e} x_{k_h}}, \qquad k \in \{U - \{k_1, \ldots, k_e\}\}, \tag{1.19}$$

gdzie $U - \{k_0\} = U$, czyli $\{k_0\}$ jest zbiorem pustym. Wyznaczamy wskaźnik k_{e+1}, tak aby

$$\pi_{k_{e+1}}^{(e)} = \text{maximum}_{k \in \{U - \{k_1, \ldots, k_e\}\}}\{\pi_k^{(e)}\}.$$

Jeśli $\pi_{k_{e+1}}^{(e)} \leq 1$, to przyjmujemy, że $\pi_{k_i} = 1$ dla $i = 1, \ldots, e$ oraz $\pi_k = \pi_{k_i}^{(e)}$ dla $k \in \{U - \{k_1, \ldots, k_e\}\}$. Wtedy algorytm wyznaczania prawdopodobieństw inkluzji zatrzymujemy.

W przeciwnym przypadku, gdy $\pi_{k_{e+1}}^{(e)} > 1$, to kontynuujemy algorytm wyznaczania tych prawdopodobieństw, przyjmując, że $\pi_{k_{e+1}} = 1$ i wyliczamy

$$\pi_k^{(e+1)} = \frac{(n-e-1)x_k}{\sum_{i=1}^{N} x_i - \sum_{h=1}^{e+1} x_{k_h}}, \qquad k \in \{U - \{k_1, \ldots, k_{e+1}\}\}. \tag{1.20}$$

Przykład 1.8

Populacja liczy $N = 6$ elementów, a próba $n = 2$ elementy. Obserwacje wartości faktur są następujące: $x_1 = 6$, $x_2 = 4$, $x_3 = 120$, $x_4 = 54$, $x_5 = 10$, $x_6 = 6$. Korzystając ze wzorów (1.18)-(1.20), mamy: $\pi_3^{(0)} = \frac{240}{200} > 1$. Przyjmujemy, że $\pi_3 = 1$, co oznacza, że trzecią fakturę dobieramy do próby celowo. Następnie wyliczamy, że $\pi_1 = \pi_6 = 0.075$, $\pi_2 = 0.05$, $\pi_4 = 0.675$ i $\pi_5 = 0.125$. Można sprawdzić, że $\sum_{k=1}^{6} \pi_k = 2$.

Uczestnicy Panelu (1989) stwierdzają, że każdy schemat losowania bezzwrotnego próby o ustalonej liczebności z różnymi prawdopodobieństwami wyboru do niej elementów populacji zależnych od zmiennej pomocniczej określonej w jednostkach pieniężnych jest nazywany *schematem losowania dolarowego, walutowego lub monetarnego*. W związku z tym każdy schemat losowania próby z określonymi wcześniej wzorami (1.18)-(1.20) prawdopodobieństwami inkluzji rzędu pierwszego może być traktowany jako schemat monetarny losowania próby. W literaturze statystycznej, a dokładniej dotyczącej metody reprezentacyjnej, można znaleźć wiele takich schematów losowania prób. Wartymi polecenia pracami, w których znajdziemy mnogość opisów schematów losowania prób, są np. pozycje Brewera i Hanifa (1983) oraz Tillé (2006).

Losowanie odrzucające

Zostanie opisany schemat zwrotnego losowania odrzucającego próby z powtarzającymi się w niej elememtami populacji, który można nazwać *zwrotnym schematem losowania odrzucającym próby z powtarzającymi się w nich elementami*. W skrócie ten sposób wyboru próby można nazwać losowaniem odrzucającym, co jest tłumaczeniem angielskiego terminu *rejective sampling* (por. np. Hájek, 1964). Opis tego schematu losowania przy ustalonych wartościach prawdopodobieństw inkluzji rzędu pierwszego przytaczamy na podstawie prac Bergera (1998) oraz Hájka (1964, 1981).

Niech

$$q_k = \lambda \frac{\pi_k}{1 - \pi_k} \left(1 + \frac{\hat{\pi} - \pi_k}{d_\pi} \right), \quad k = 1, \ldots, N, \tag{1.21}$$

gdzie

$$d_\pi = \sum_{k=1}^{N} \pi_k (1 - \pi_k), \tag{1.22}$$

$$\hat{\pi} = \frac{1}{d_\pi} \sum_{k=1}^{N} \pi_k^2 (1 - \pi_k),$$

przy czym λ jest tak dobierane, aby

$$\sum_{k=1}^{N} q_k = 1.$$

Wówczas, algorytm losowania próby o liczebności n-elementów jest następujący. Losujemy zwrotnie n-elementową próbę, przy czym w każdym ciągnieniu k-ty element populacji jest losowany z prawdopodobieństwem q_k, $k = 1, \ldots, N$. Jeśli w skład tak wylosowanej próby wchodzą różne elementy populacji, to algorytm losowania zatrzymujemy. W drugiej możliwej sytuacji, gdy w wylosowanej próbie przynajmniej dwa elementy populacji powtarzają się, taką próbę odrzucamy (nie akceptujemy) i losujemy zwrotnie ponownie próbę z prawdopodobieństwami q_k, $k = 1, \ldots, N$ dostania się kolejnych elementów populacji do próby. Ten algorytm powtarzamy aż do skutku polegającego na wylosowaniu próby, w której nie występują powtórzenia elementów.

Dodajmy, że można nieco usprawnić (przyspieszyć) właśnie opisany algorytm. Polega to na tym, że na bieżąco podczas losowania kolejnych elementów do próby sprawdzamy, czy występują powtórzenia wśród dotychczas wylosowanych elementów populacji do tworzonej próby. Gdy natrafimy na powtórzenie elementu już wcześniej wybranego do próby, to procedurę przerywamy i rozpoczynamy losowanie próby od nowa.

Plan losowania próby za pomocą opisanego schematu losowania określa wzór:

$$P(s) = c \prod_{k \in s} q_k, \qquad s \in \mathbf{S},$$

przy czym c jest tak dobierane, aby $\sum_{s \in \mathbf{S}} P(s) = 1$. Prawdopodobieństwa inkluzji tego planu losowania są, co należy podkreślić, tylko w przybliżeniu równe postulowanym wartościom π_k, $k = 1, \ldots, N$.

Schemat losowania Rao-Sampforda

Plan losowania określa wzór (por. Hájek, 1981, s. 87; Rao, 1965 i Stampford, 1967):

$$P_{RS}(s) = c_1 \left(\sum_{k \in U - s} \pi_k \right) \prod_{k \in s} \frac{\pi_k}{1 - \pi_k}, \qquad s \in \mathbf{S},$$

przy czym c_1 jest tak dobierane, aby $\sum_{s \in \mathbf{S}} P_{RS}(s) = 1$. Prawdopodobieństwa inkluzji rzędu pierwszego zapisanego planu losowania są dokładnie równe postulowanym wartościom π_k, $k = 1, \ldots, N$.

Schemat losowania próby polega na wylosowaniu zwrotnym pierwszego elementu z prawdopodobieństwem π_k/n, $k = 1, \ldots, N$, a pozostałe elementy w ilości $n - 1$ są losowane również zwrotnie z prawdopodobieństwami $\frac{\pi_k}{1-\pi_k}$, $k = 1, \ldots, N$. Jeśli w tak wylosowanej próbie wystąpią przynajmniej dwa takie same elementy populacji, to losowanie próby ponawiamy od początku, aż do sytuacji, gdy nie wystąpi powtórzenie elementu populacji w próbie.

Próba systematyczna

W większości podręczników opis schematów losowania próby monetarnej (dolarowej) jest ograniczany jedynie do schematu losowania próby systematycznej z zadanymi prawdopodobieństwami inkluzji rzędu pierwszego (por. np. Hartley i Rao, 1962 oraz Rao i in., 1962). Oznaczmy próbę jako ciąg kolejno losowanych do niej elementów postaci: $s = (i_1, i_2, \ldots, i_n)$. Najpierw wyliczamy sumę:

$$V_e = \sum_{k=1}^{e} \pi_k, \quad e = 1, \ldots, N,$$

przy czym $V_0 = 0$ i $V_N = n$. Następnie jest generowana liczba losowa o rozkładzie jednostajnym z przedziału $(0; 1)$. Element populacji z kolei o numerze i_1 należy dobrać do próby, czyli $i_1 \in s$, gdy zachodzi następująca nierówność:

$$V_{i_1-1} < u \leq V_{i_1}.$$

Drugim elementem próby jest element populacji indeksowany przez i_2, przy czym musi być spełniona nierówność:

$$V_{i_2-1} < u + 1 \leq V_{i_2}.$$

W ogólności, element $i_z \in s$, $z = 1, \ldots, n$ wtedy i tylko wtedy, gdy:

$$V_{i_z-1} < u + z - 1 \leq V_{i_z}. \tag{1.23}$$

Przykład 1.9

Dokumenty księgowe są losowane do próby z następującymi prawdopodobieństwami inkluzji rzędu pierwszego: $\pi_1 = 0.2$, $\pi_2 = 0.8$ $\pi_3 = 0.05$, $\pi_4 = 0.3$, $\pi_5 = 0.5$, $\pi_6 = 0.15$, $\pi_7 = 0.1$, $\pi_8 = 0.3$, $\pi_9 = 0.25$, $\pi_{10} = 0.35$. Suma tych prawdopodobieństw wynosi trzy, a więc rozmiar próby losowanej bezzwrotnie wynosi $n = 3$. Wygenerowana wartość zmiennej losowej o rozkładzie jednostajnym na przedziale $(0;\ 1)$ wynosi 0.81. Wówczas korzystając ze wzoru (1.23), mamy:

$$\begin{aligned}
&i_1 = 2, \quad \text{bo} \quad V_1 = 0.2 < u = 0.81 < 1.0 = V_2,\\
&i_2 = 5, \quad \text{bo} \quad V_4 = 1.35 < u + 1 = 1.81 < 1.85 = V_5,\\
&i_3 = 10, \quad \text{bo} \quad V_9 = 2.65 < u + 2 = 2.81 < 3.0 = V_{10}.
\end{aligned}$$

Stąd wynika, że próba ma postać: $s = \{2, 5, 10\}$.

Teraz zostanie przedstawiony algorytm doboru próby nieznacznie różniącego się od opisanego wyżej schematu losowania. Definiujemy następujące sumy (skumulowane):

$$x_{c,e} = \sum_{j=1}^{e} x_j, \qquad e = 0, 1, \ldots, N, \tag{1.24}$$

przy czym $x_{c,e} = 0$, $x_{c,N} = x_U$. Następnie wyznaczamy iloraz

$$I_n = z\left(\frac{x_U}{n}\right), \tag{1.25}$$

przy czym $z(a)$ oznacza zaokrąglenie argumentu a do najbliższej liczby naturalnej. Parametr I_n jest nazywany interwałem losowania systematycznego z różnymi prawdopodobieństwami (inkluzji rzędu pierwszego) doboru elementów do próby. W rozważanym kontekście zagadnienia audytu finansowego interwał jest liczbą określającą kwotę pieniędzy, którą zawsze można wyrazić jako liczbę całkowitą, np. wyrażoną w złotówkach lub groszach. To pozwala wylosować jedną spośród liczb nauralnych tworzących ciąg $(1, 2, \ldots, I_n)$, przy czym prawdopodobieństwo wylosowania każdego wyrazu tego ciągu jest stałe i równe $\frac{1}{I_n}$. Przypuśćmy, że k-ty element tego ciągu został wylosowany. To przesądza o tym, że do próby systematycznej jako pierwszy jest losowany element populacji z przypisanym mu indeksem i_1, dla którego jest spełniona nierówność:

$$x_{c,i_1-1} < k \leq x_{c,i_1}$$

Indeks i_2 następnego elementu populacji dołączanego do próby wynika ze wzoru:

$$x_{c,i_2-1} < k + I_n \leq x_{c,i_2}$$

Kontynuując ten algorytm, za z-tym razem do próby jest dobierany element populacji, którego numer (indeks) spełnia nierówność:

$$x_{c,i_z-1} < k + (z-1)I_n \leq x_{c,i_z}, \qquad z = 1, \ldots, n. \tag{1.26}$$

Opisany teraz schemat losowania próby jest nazywany dolarowym (monetarnym), por. np. Panel (1989). Dodajmy, że szczególnym przypadkiem tego algorytmu wyznaczania elementów tworzących próbę jest schemat losowania prostej próby systematycznej, opisany w poprzednim punkcie.

Zauważmy, że jeśli I_n jest liczbą naturalną oraz $\pi_k = \frac{nx_k}{x_U} \leq 1$, dla $k = 1, \ldots, N$, to dzieląc obie nierówności układu określonego wyrażeniem (1.26) przez interwał losowania I_n, otrzymujemy układ nierówności określony wzorem (1.23). Zatem w rozważanym przypadku obydwa schematy losowania są sobie równoważne, czyli schemat losowania systematycznego próby z nierównymi prawdopodobieństwami jest równoważny schematowi dolarowemu losowania.

Który z nich preferować? Autor uważa, że pierwszy z nich, ponieważ określony wzorami (1.24)-(1.26) schemat losowania w pewnych szczególnych przypadkach (por. np. Karliński, 2005) może eliminować możliwość losowania niektórych elementów populacji, albo wskazywać na ponowny dobór do próby już wcześniej wylosowanego elementu populacji. Te problemy są m.in. wynikiem zbyt dużej wartości nominalnej przypisanej danemu elementowi lub zaokrąglania ilorazu $\frac{x_U}{n}$, zdefiniowanego wzorem (1.25).

Przykład 1.10

Nominalne (księgowe) wartości faktur (w zł) są następujące, $x_1 = 100$, $x_2 = 31, x_3 = 59, x_4 = 50, x_5 = 24, x_6 = 36, x_7 = 78, x_8 = 22, x_9 = 20$. Zadanie polega na wylosowaniu próby trójelementowej za pomocą walutowego schematu losowania próby określonego wyrażeniami (1.24) i (1.25). W tym celu wyznaczamy sumę wartości wszystkich faktur, czyli $x_U = 420$, potem interwał losowania, który wynosi $I_3 = \frac{420}{3} = 140$. Następnie spośród liczb naturalnych tworzących zbiór $\{1, 2, \ldots, 140\}$ losowo dobieramy jedną. Zatem każda liczba z tego zbioru ma szansę być wylosowana z prawdopodobieństwem $\frac{1}{140}$. Przypuśćmy, że 53 jest tą wylosowaną liczbą. Wówczas do próby dobieramy element nr 1, ponieważ

$$x_{c,0} = 0 < 53 < x_{c,1} = 100,$$

potem element nr 4, bo

$$x_{c,3} = 190 < 53 + 140 = 193 < x_{c,4} = 240$$

i w końcu element nr 7, bowiem

$$x_{c,6} = 300 < 53 + 2 \cdot 140 = 333 < x_{c,7} = 378.$$

Zatem próba jest utworzona przez zbiór elementów $s = \{1, 4, 7\}$.

Przykład 1.11

Nominalne wartości faktur są następujące: $2, 5, 4, 6, 7, 11, 8, 9, 3, 5$. Należy wylosować próbę o rozmiarze $n = 5$ za pomocą odrzucającego schematu losowania zaproponowanego przez Rao-Sampforda.

Zadanie to realizujemy z wykorzystaniem następującej procedury losowania, zaprogramowanej w języku R.

$\#liczebność\ \ populacji:$
$N < -10$
$\#liczebność\ \ próby:$
$n < -5$
$x < -c(2, 5, 4, 6, 7, 11, 8, 9, 3, 5)$
$p < -matrix(0, N, 1)$
$q < -matrix(0, N, 1)$
$s < -matrix(0, n, 1)$
$sor < -matrix(0, n, 1)$
$xx < -sum(x); p < -n * x/xx$
$for(iin1 : N)q[i] < -p[i]/(1 - p[i])$
$a < -1/sum(q); q < -a * q$
$lp < -0; id < -1$
$while(id == 1)\{$
$id < -0; lp < -lp + 1$

$s < -sample(1:N, n, replace = TRUE, q)$
$sor < -sort(s)$
$i < -1$
$while((id == 0)\&(i < n))\{$
$if(sor[i+1] == sor[i])id < -1$
$i < -i+1\}\}; s$

Realizacja tej aplikacji doprowadziła do wylosowania elementów populacji o nr 8, 3, 5, 7, przy czym wektory prawdopodobieństw p i q są następujące:

$$p = [0.133\ 0.333\ 0.267\ 0.400\ 0.467\ 0.733\ 0.533\ 0.600\ 0.200\ 0.333]$$

oraz

$$q = [0.018\ 0.058\ 0.042\ 0.077\ 0.101\ 0.316\ 0.131\ 0.172\ 0.029\ 0.058].$$

1.3.3. Próba warstwowa

Prezentowany teraz schemat losowania próby jest jednym z naczęściej używanych w praktyce. Zakładamy, że populacja U jest podzielona na niepuste podzbiory U_h, $h = 1, \ldots, H$, zwane warstwami oraz $U = \bigcup_{h=1}^{H} U_h$. Liczbę elementów populacji tworzących $h-$tą warstwę oznaczymy przez $0 < N_h < N$, przy czym $N = \sum_{h=1}^{H} N_h$. Niech $w_h = N_h N^{-1}$. Przez S_h oznaczymy próbę prostą o liczebności $0 < n_h \leq N_h$ losowaną bezzwrotnie z $h-$tej warstwy. Próbę składającą się z prób prostych losowanych bezzwrotnie z poszczególnych warstw oznaczamy przez $s = \{s_1, s_2 \ldots s_H\}$. Tak tworzona próba z uwagi na jej powszechne stosowanie w praktyce jest nazywana próbą warstwową losowaną bezzwrotnie lub krótko – próbą warstwową. Jej plan losowania ma postać:

$$P_{war}(\mathbf{s}) = \prod_{h=1}^{H} P_h(s_h), \tag{1.27}$$

gdzie:

$$P_h(s_h) = \binom{N_h}{n_h}^{-1}$$

jest planem losowania bezzwrotnego próby prostej s_h o liczebności n_h z warstwy o liczebności N_h.

Prawdopodobieństwa inkluzji określa wyrażenie:

$$\begin{cases} \pi_k = \frac{n_h}{N_h} & \text{dla} \quad k \in U_h, \\ \pi_{k,l} = \frac{n_h}{N_h}\frac{n_t}{N_t} & \text{dla} \quad k \in U_h, \quad l \in U_t, \\ \pi_{k,l} = \frac{n_h(n_h-1)}{N_h(N_h-1)} & \text{dla} \quad k \in U_h, \quad l \in U_h, \end{cases} \tag{1.28}$$

przy czym $k = 1, \ldots, N$, $l = 1, \ldots, N$, $l \neq k$, $h = 1, \ldots, H$, $t = 1, \ldots, H$, $h \neq t$. Stąd wynika, że prawdopodobieństawa inkluzji rzędu pierwszego nie

zawsze są takie same, a więc próba warstwowa nie jest próbą prostą, mimo że jej składowe podpróby losowane z poszczególnych warstw są próbami prostymi.

Przykład 1.12

Możliwe wartości faktur tworzą ciąg $\{y_i\}$, $i = 1, \ldots, N$. Zatem populację tworzą faktury, z których każda ma swój specyficzny i niepowtarzający się numer identyfikacyjny (nazywany etykietą). W związku z tym możliwe jest wzajemnie jednoznaczne odwzorowanie zbioru identyfikatorów faktur ze zbiorem kolejnych liczb naturalnych $\{1, 2, \ldots, N\}$ reprezentujących naszą populację faktur, czyli $U = \{1, 2, \ldots, N\}$. Dla ustalenia uwagi przyjmijmy, co nie umniejsza ogólności dalszych analiz, że rosnące numery faktur są zgodne z niemalejącym uporządkowaniem wartości tych faktur, czyli $y_i \leq y_{i+1}$ dla $i \in U$. Z kolei wartość y_i przypisana i-tej fakturze, jest obserwacją pewnej zmiennej y. Przedział zmienności możliwych wartości zmiennej y oznaczamy przez $I = [0; c]$. Jedną z możliwych partycji tego przedziału tworzą składniki sumy: $I = \bigcup_{h=1}^{H} I_h$, gdzie $I_h = [c_{h-1}; c_h]$, $c_{h-1} < c_h$, $h = 1, \ldots, H$, przy czym $c_0 = 0$. Zakładamy, iż te przedziały są rozłączne, czyli że ich wspólną częścią jest zbiór pusty, co symbolicznie zapisujemy wyrażeniem $I_h \cap I_t = \varnothing$, $t \neq h$, $h = 1, \ldots, H$, $t = 1, \ldots, H$. Przyjmijmy, że każdemu przedziałowi I_h jest przypisany niepusty podzbiór faktur U_h, czyli $i \in U_h$ wtedy i tylko wtedy, gdy $y_i \in I_h$. Zatem U_h, $h = 1, \ldots, H$ są warstwami populacji U.

Załóżmy, że z populacji wylosujemy n elementów. Wtedy mamy problem, jak ustalić liczebności prób losowanych z poszczególnych warstw. Ich rozmiary można ustalać w różny sposób (por. np. Bracha, 1996, 1998; Steczkowski, 1995; Pawłowski, 1972 i Wywiał, 1992, 2010). Najpopularniejszy sposób polega na proporcjonalnym wyznaczaniu liczebności próby do rozmiaru warstwy, co określa wyrażenie:

$$n_h = nw_h, \qquad w_h = \frac{N_h}{N}, \qquad h = 1, \ldots, H, \tag{1.29}$$

przy czym

$$\sum_{h=1}^{H} n_h = n. \tag{1.30}$$

Tak wyznaczone liczebności prób prowadzą do następującego uproszczenia określonych we wzorze (1.28) prawdopodobieństw inkluzji rzędu pierwszego:

$$\pi_k = \frac{n}{N} \quad \text{dla} \quad k \in U.$$

Czyli wariant proporcjonalny wyznaczania liczebności prób daje stałą wartość prawdopodobieństwa wylosowania elementu z każdej warstwy, tak samo jak w przypadku próby prostej losowanej bezzwrotnie.

Wzór (1.29) daje liczebności n_{*h}, które niekoniecznie są liczbami naturalnymi. Wówczas należy tak je pozaokrąglać (raz z nadmiarem, innym razem niedomiarem), aby równanie (1.30) było spełnione. Dodajmy, że ostatnio pewien optymalny sposób tego typu zaokrąglania zaproponował Muraszew (2011).

Przykład 1.13

Populacja liczy $N = 80000$ dokumentów księgowych. Podzielono ją na trzy warstwy, kolejno o liczebnościach: $N_1 = 50000$, $N_2 = 20000$ i $N_3 = 10000$. Zatem frakcje (udziały) rozmiarów poszczególnych warstw w liczebności populacji wynoszą odpowiednio $w_1 = 0.625$, $w_2 = 0.25$ i $w_3 = 0.125$. Planuje się wylosować ze wszystkich warstw próbę o rozmiarze $n = 100$. Wówczas, stosując proporcjonalny wariant ustalania liczebności na podstawie wzoru (1.29), wyliczamy, że $n_{*1} = 62.5$, $n_{*2} = 25$ i $n_{*3} = 12.5$. Ze względu na to, iż te liczby nie są całkowite, arbitralnie je zaokrąglamy, tak by ich suma była równa 100. Zatem, w szczególności z kolejnych warstw należałoby wylosować odpowiednio 63, 25 i 12 elementów.

1.4. Estymacja

1.4.1. Definicje podstawowe

Zwykle nie jest możliwa obserwacja wszystkich wartości cechy w całej populacji, ponieważ jest to rzecz pracochłonna i kosztowna. Dlatego m.in. trzeba zadowolić się obserwacjami poczynionymi w próbie. Celem estymacji są parametry zmiennych, czyli np. wartość globalna zmiennej lub wskaźnik struktury. Do tego celu używa się tzw. estymatorów.

Zaniedbując nieco precyzję definicji, najpierw określamy *statystykę* jako każdą rzeczywistą funkcję obserwacji zmiennej w losowanej próbie. Statystykę oznaczymy przez t_s, gdzie symbolem s oznaczono już wcześniej próbę.

Niech celem będzie ocena parametru $\theta \in \Theta$, gdzie Θ jest zbiorem możliwych wartości θ. W szczególności parametrem zmiennej może być frakcja p. Wówczas jej zbiór możliwych wartości jest przedziałem $\Theta = [0; 1]$. Statystyka t_s użyta do oceny wartości parametru zmiennej i przyjmująca wartości ze zbioru Θ jest nazywana *estymatorem* tego parametru. Dokładność estymacji określa jej błąd wyliczany jako różnica:

$$u_s = t_s - \theta$$

Ten błąd jest zmienną losową, której wartość oczekiwaną określa wzór:

$$E(u_s) = E(t_s) - \theta = \delta(t_s),$$

gdzie:

$$E(t_s) = \sum_{s \in \mathbf{S}} t_s P(s)$$

Wartość oczekiwana (inaczej nadzieja matematyczna lub wartość średnia) estymatora t_s jest zatem ważoną wartością średnią możliwych wartości estymatora, wyznaczanych na podstwie obserwacji zmiennej we wszystkich możliwych próbach, które tworzą zbiór $\mathbf{S}$, zwany przestrzenią prób. Wagami są odpowiednie prawdopodobieństwa wylosowania prób. Krócej rzecz ujmując, wartość oczekiwana $E(t_s)$ jest średnią wszystkich możliwych wartości estymatora t_s ważonych prawdopodobieństwami $P(s)$ losowania prób. Parametr δ jest nazywany *obciążeniem estymatora* t_s i jest średnią wszystkich możliwych błędów estymacji . Gdy $\delta = 0$, to mówimy, że estymator jest nieobciążony, czyli używając go nie popełniamy błędów systematycznych przy ocenie ocenie parametru θ. Różne od zera wartości obciążenia δ świadczą o występowaniu błędów systematycznych. W szczególności, jeśli $\delta = \delta_0 > 0$, to estymator t_s średnio rzecz biorąc przeszacowywuje parametr θ o wartość δ. W przeciwnym przypadku, gdy $\delta_0 < 0$, to estymator t_s nie doszacowuje parametr Θ o wartość δ_0.

Dodajmy jeszcze, iż rozważa się również estymatory w przybliżeniu nieobciążone. W szczególności chodzi tutaj o klasę tzw. *granicznie (asymptotycznie) nieobciążonych* estymatorów. Z takim estymatorem mamy do czynienia wówczas, gdy jego obciążenie zmierza do zera wraz z nieograniczonym wzrostem liczebności próby przy jednoczesnym nieograniczonym wzroście liczebności populacji. Formalniej, estymator t_s jest granicznie nieobciążony dla parametru θ, gdy $\delta(t_s) \longrightarrow 0$ dla $n \longrightarrow \infty$, $N \longrightarrow \infty$ i $N - n \longrightarrow \infty$.

Drugim parametrem estymatora jest jego wariancja wyznaczana przez wzór:

$$V(t_s) = \sum_{s \in \mathbf{S}} (t_s - E(t_s))^2 P(s).$$

Wariancja jest średnią kwadratów odchyleń wszystkich możliwych wartości estymatora t_s od jego nadziei matematycznej $E(t_s)$ ważonych prawdopodobieństwami wystąpienia prób, na podstawie których wartości estymatora są wyznaczane. Pierwiastek z wariancji nazywany *odchyleniem standardowym estymatora* ocenia przeciętne odchylenia losowych wartości estymatora od jego wartości oczekiwanej.

Błąd średniokwadratowy estymatora określa wzór:

$$MSE(t_s) = \sum_{s \in \mathbf{S}} (t_s - \theta)^2 P(s). \tag{1.31}$$

Ten błąd jest średnią kwadratów odchyleń wartości estymatora t_s od szacowanej wartości parametru θ ważonych prawdopodobieństwami wystąpienia prób, na podstawie których wartości estymatora są wyznaczane. Krótko ujmując, $MSE(t_s)$ jest średnią wszystkich możliwych kwadratów błędów estymacji ważonych prawdopodobieństwami ich wystąpienia. Pierwiastek z błędu średniokwadratowego jest nazywany *błędem średnim szacunku parametru θ za pomocą estymatora t_s* lub krótko *błędem średnim szacunku*, a oznaczać go będziemy symbolem $d(t_s) = \sqrt{MSE(t_s)}$. Parametr $d(t_s)$ jest interpretowany jako średnie odchylenie wartości estymatora od wartości szacowanego parametru lub krócej jako średni błąd estymacji. Wykazuje się, że

$$MSE(t_s) = V(t_s) + \delta^2(t_s).$$

Błąd średniokwadratowy dekomponuje się na część losową reprezentowaną wariancją estymatora oraz część systematyczną reprezentowaną kwadratem obciążenia estymacji. Zatem, jeśli dany estymator jest nieobciążony, to błąd średniokwadratowy redukuje się do wariancji estymatora, a błąd średni szacunku staje się odchyleniem standardowym estymatora.

Zwykle poszukuje się estymatorów nieobciążonych danych parametrów, nie tylko dlatego, by uchronić się przed występowaniem błędu systematycznego estymacji, lecz dlatego, że ocena tego błędu, a co za tym idzie również ocena $MSE(t_s)$, jest w praktyce trudna. Zazwyczaj trzeba się zadowolić, o czym będzie jeszcze mowa, oceną wariancji estymatora, którą również prowadzi się na podstawie próby.

Dodajmy, że błąd średniokwadratowy $MSE(t_s)$ jest proporcjonalny do wskaźnika ryzyka estymacji. Formalnie rzecz biorąc, uważa się, co wynika ze wzoru (1.31), że $MSE(t_s)$ jest wartością oczekiwaną kwadratowej funkcji straty estymacji, czyli kwadaru błędu estymacji parametru θ za pomocą statystyki t_s. Zauważmy, że można np. uwzględnić pewien współczynnik a określający jednostkową stratę, czyli przyrost finansowej straty spowodowanej przyrostem błędu. Wtedy wartość funkcji ryzyka wyniesie $a^2 MSE(t_s)$. W szczególności gdy $a = 1$, błąd średniokwadratowy estymacji jest jednocześnie kwadratową funkcją ryzyka estymacji.

Popularnym wskaźnikiem dokładności estymacji jest również względny średni błąd szacunku, który definiuje wzór:

$$\nu(t_s) = \frac{d(t_s)}{|\theta|}.$$

Określa on udział procentowy błędu średniego szacunku w wartości szacowanego parametru. Innymi słowy $\nu(t_s)100\%$ informuje o tym, jaki procent wartości szacowanego parametru stanowi błąd średni szacunku jego oceny. Dodajmy jeszcze, że w szczególności gdy brany pod uwagę estymator jest nieobciążony, to $\nu(t_s) = \frac{\sqrt{V(t_s)}}{|\theta|}$.

Rezultatem przedstawionej procedury estymacji są dwie liczby. Pierwszą z nich jest ocena nieznanego parameru, którą jest wartość wybranego estymatora, a drugą ocena błędu średniego szacunku. Dlatego ta ogólna metoda estymacji jest nazywana punktową, bowiem jej rezultatem jest pojedyncza (punktowa) ocena wartości parametru uzupełniona o ocenę średniego odchylenia możliwych wartości estymatora od szacowanego parametru.

Mamy jeszcze do czynienia z drugą ogólną procedurą estymacji zwaną przedziałową. Rezultatem tego typu procedury wnioskowania jest przedział liczbowy, który ma ze z góry przyjętym prawdopodobieństwem obejmować wartość szacowanego parametru. Popularną wersję przedziału ufności można opisać następująco. Wymaga się, aby t_s był nieobciążonym lub przynajmniej asymptotycznie nieobciążonym estymatorem θ.Wprowadzamy symbol $V_s(t_s)$ jako oznaczenie zgodnej oceny wariancji estymatora. Wówczas przedział ufności dla parametru θ, oznaczany przez $I_s = [t_{1s}; t_{1s}]$, spełnia wyrażenia

$$P\left(t_{1s} \leq \theta \leq t_{2s}\right) = \gamma,$$

przy czym

$$t_{1s} = t_s - z_\gamma \sqrt{V_s(t_s)}, \qquad t_{2s} = t_s + z_\gamma \sqrt{V_s(t_s)},$$

$$Z_s = \frac{t_s - \theta}{\sqrt{V_s(t_s)}}, \qquad P\left(|Z_s| \geq z_\gamma\right) = \gamma.$$

Przez γ oznaczono prawdopodobieństwo tego, iż przedział ufności obejmie wartość parametru θ. Dodajmy, że przyjmuje się, że dokładny rozkład prawdopodobieństwa statystyki Z_s jest znany lub jest on przynajmniej granicznie standardowym rozkładem normalnym. Warto dodać, iż dokładność estymacji przedziałowej określa się połową długości przedziału ufności.

1.4.2. Estymacja na podstawie próby prostej

Wprowadzone wcześniej definicje związane z estymacją zilustrujemy na przykładzie oceny wartości średniej zmiennej y (np. wartości przypisanych fakturom) w populacji. Naszym celem jest ocena wartości średniej $\bar{y}$. Do tego celu jest zgodnie z definicją (1.2) losowana bezzwrotnie próba prosta s o liczebności n elementów. Okazuje się, że nieobciążonym estymatorem parametru $\bar{y}$ jest następująca średnia z próby:

$$\bar{y}_s = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{n}\sum_{i \in s} y_i. \tag{1.32}$$

Jej wariancja jest postaci:

$$V(\bar{y}_s) = \frac{N-n}{Nn} v_{*U,y}. \tag{1.33}$$

Wariancja $V(\bar{y}_s)$ określa przeciętny poziom kwadratów odchyleń możliwych wartości średniej z próby $\bar{y}_s$ od średniej w populacji $\bar{y}$. Wynika to stąd, że wzór (1.33) da się przedstawić jako średnią arytmetyczną kwadratów błędów estymacji $u_s^2 = (\bar{y}_s - \bar{y})^2$ wyliczoną po wszystkich możliwych do wylosowania próbach tworzących zbiór (przestrzeń prób) $\mathbf{S}$, składającą się z $\binom{N}{n}$ prób, czyli

$$V(\bar{y}_s) = \frac{1}{\binom{N}{n}} \sum_{s \in \mathbf{S}} (\bar{y}_s - \bar{y})^2.$$

Odpowiednie operacje algebraiczne zapisanego wzoru prowadzą do wyrażenia (1.33), por. np. Wywiał (1992). Błąd średni szacunku przyjmuje postać:

$$d(\bar{y}_s) = \sqrt{V(\bar{y}_s)} = \sqrt{\frac{N-n}{Nn} v_{*U,y}}, \tag{1.34}$$

Stąd wynika, że dokładność estymacji rośnie wraz ze wzrostem liczebności próby. Zatem zwiększanie rozmiaru próby przyczynia się do zmniejszenia ryzyka wystąpienia zbyt dużych błędów estymacji. Parametr $d(\bar{y}_s)$ interpretuje się jako przeciętne (oczekiwane) odchylenie możliwych wartości średniej z próby $\bar{y}_s$ od średniej z populacji $\bar{y}$.

W praktyce wartość błędu średniego szacunku nie jest znana, ponieważ wariancja $v_{*U,y}$ nie jest znana, dlatego w celu oceny wartości tego parametru używa się następującego estymatora:

$$d_s(\bar{y}_s) = \sqrt{V_s(\bar{y}_s)} = \sqrt{\frac{N-n}{Nn} v_{y,s}}, \tag{1.35}$$

gdzie:

$$v_{y,s} = \frac{1}{n-1} \sum_{i \in s} (y_i - \bar{y}_s)^2. \tag{1.36}$$

W końcu zaznaczmy, iż względny średni błąd szacunku oraz jego estymator określają odpowiednio wzory:

$$\gamma(\bar{y}_s) = \sqrt{\frac{N-n}{Nn\bar{y}} v_{*U,y}} 100\%, \qquad \gamma_s(\bar{y}_s) = \sqrt{\frac{N-n}{Nn\bar{y}_s} v_{y,s}} 100\%. \tag{1.37}$$

Powszechnie używany nieobciążony estymator wartości globalnej ma postać:

$$y_{U,s} = N\bar{y}_s = \frac{N}{n} \sum_{i=1}^{n} y_i. \tag{1.38}$$

Jego wariancja i jej ocena są odpowiednio następujące:

$$V(y_{U,s}) = \frac{N-n}{n} v_{*U,y}, \qquad V_s(y_{U,s}) = \frac{N-n}{n} v_{y,s}. \tag{1.39}$$

Przykład 1.14

Urząd skarbowy dysponuje deklarowanymi kwotami zapłaconych podatków od 30000 jego płatników. Wartość średnia i odchylenie standardowe zaobserwowanych kwot uiszczonych podatków wynoszą odpowiednio 1211.23 zł i 221.21 zł. Z tej populacji wylosowano bezzwrotnie 100-elementową próbę prostą płatników i przeprowadzono w niej audyt kwot zapłaconych podatków, w wyniku czego otrzymano skorygowane wartości podatków, które miały być zapłacone. Okazało się, że średnia z próby i odchylenie standardowe z próby skorygowanych kwot podatków wynoszą odpowiednio 1298.39 zł i 189.54 zł.

Korzystając ze wzoru (1.35), mamy $d_s(\bar{y}_s) = \sqrt{\frac{30000-100}{100 \cdot 30000} 189.54^2} = 18.92$. Stąd wynika, że ocena skorygowana przeciętnej podatków wynosi 1298.39, a błąd średni szacunku jest równy 18.92. Zatem, możliwe oceny wartości średniej skorygowanych podatków odchylają się od nieznanej prawdziwej wartości przeciętnej skorygowanych podatków średnio o 18.92 zł. Ocena względnego średniego błędu szacunku wynosi $d_s(\bar{y}_s) = 100\%18.92/1298.39 = 1.5\%$.

Na podstawie wzorów (1.38) i (1.39) oceniamy wartość globalną podatków, która powinna była wpłynąć do urzędu skarbowego. W wyniku obliczeń uzyskano, że ocena wartości globalnej wynosi $38951700 = 1298.39 \cdot 30000$, natomiast oszacowanie błędu średniego szacunku jest równe $d_s(y_{U,s}) = \sqrt{V_s(y_{U,s})} = 30000 \cdot 18.92 = 567600$ zł.

Z praktycznego punktu widzenia jednym z podstawowych problemów jest ustalanie liczebności próby, która zagwarantuje nam postulowaną dokładność estymacji. Dążymy do takiego wyznaczenia liczebności próby, aby błąd średni szacunku przeciętej w populacji określony wzorem (1.34) nie przekroczył postulowanego poziomu d_0. Te liczebności otrzymujemy z rozwiązania nierówności $V(\bar{y}_s) \leq d_0^2$. Daje to następujący wynik:

$$n \geq n_0 = \left[\frac{1}{\frac{d_0^2}{v_{*U,y}} + \frac{1}{N}} \right] + 1, \tag{1.40}$$

przy czym symbolem $[a]$ oznaczono znaną funkcję definiowaną jako *część całkowitą z liczby a*. Zauważmy, że jeśli liczebność populacji jest bardzo znaczna, to powyższy wzór upraszcza się do postaci:

$$n \geq n_0 = \left[\frac{v_{*U,y}}{d_0^2} \right] + 1.$$

Stąd natychmiast wynika, że potrzebna liczebność próby jest tym większa, im wariancja $v_{*U,y}$ analizowanej zmiennej jest większa lub gdy postulowany dopuszczalny błąd średni estymacji d_0 jest mniejszy. W praktyce zazwyczaj

nie jest znana wartość wariancji $v_{*U,y}$. Dlatego w miejsce wariancji podstawia się jej estymator określony wzorem (1.36). Drugim źródłem oszacowania wariancji $v_{*U,y}$ są jej oceny uzyskane we wcześniejszych badaniach.

W przypadku estymacji wartości globalnej zmiennej, gdy teraz d_0 jest dopuszczalnym błędem średnim szacunku, wzór określający potrzebną liczebność próby ma postać:

$$n \geq n_0 = \left[\frac{N^2 v_{*U,y}}{d_0^2 + N v_{*U,y}} \right] + 1, \tag{1.41}$$

Inny sposób formułowania zadania wyznaczania niezbędnej liczebności próby polega na takim jej wyliczeniu, aby względny średni błąd szacunku przeciętnej w populacji nie przekroczył żądanego poziomu dopuszczalnego, a oznaczanego dalej przez ν_0. Zatem z nierówności $\nu(\bar{y}_s) \leq \nu_0$ wynika, że

$$n \geq n_0 = \left[\frac{1}{\frac{\nu_0^2 \bar{y}_s^2}{v_{*U,y}} + \frac{1}{N}} \right] + 1. \tag{1.42}$$

Dodajmy, że zapisany wzór również dotyczy przypadku estymacji wartości globalnej przy postulowanym względnym średnim błędzie szacunku.

Przykład 1.15

Pozostajemy przy danych z przykładu 1.14, którego rezultatem były oceny wartości średniej i wartości globalnej prawidłowo ustalonego podatku. Przypuśćmy, że dokładność przeprowadzonej estymacji nie zadowala audytora. Wówczas, aby podnieść precyzję szacunków, m.in. zwiększa się liczebność losowanej próby. Przypuśćmy, że żąda się, aby wartość globalna była szacowana z błędem średnim szacunku nieprzekraczającym poziomu 200 000 zł. Wówczas, korzystając ze wzoru (1.41), mamy, iż wylosowanie próby prostej o liczebności $n_0 = 1063$ elementów pozwoli nam oczekiwać, że błąd średni szacunku wartości globalnej wpływów z prawidłowo ustalonych podatków nie przekroczy poziomu 200 000 zł.

Teraz wyznaczamy niezbędną liczebność próby potrzebną do estymacji wartości globalnej ze średnim błędem względnym szacunku nieprzekraczającym poziomu 1%. Korzystając tym razem ze wzoru (1.42), wyliczamy, że potrzebna liczebność próby powinna być co najmniej równa 212 elementom.

Przykład 1.16

Podobnie jak w poprzednim przykładzie, korzystamy z danych określonych w przykładzie 1.14. Najpierw załóżmy, że próba prosta nie została jeszcze wylosowana. Wówczas możemy wykorzystać dane o już zadeklarowanych i wpłaconych kwotach podatków. Zgodnie z addytywnym modelem generowania się błędów obserwacje zapłaconego podatku można zapisać

jako sumę: $x_k = y_k + e_k$, gdzie y_k jest prawidłowo określonym poziomem podatku, a e_k poziomem błędu, $k = 1, \ldots, N$. Wówczas, przy założeniu, że obserwacje błędów nie są ujemnie skorelowane z prawidłowo określonymi poziomami podatku, mamy, iż $v_{*U,x} \geq v_{*U,y}$ (por. wzór (1.6)). Ten fakt pozwoli wyznaczyć niezbędną liczebność próby gwarantującą to, iż błąd średni szacunku będzie mniejszy od postulowanego. Zakładając, że dopuszczalny błąd średni szacunku wynosi $d_0 = 1000$ zł i korzystając ze wzoru (1.41), w którym nieznaną wariancję $v_{*U,y}$ należy zastąpić przez $v_{*U,x}$, wyliczamy, iż $n_0 = 1400$. Zatem taka liczebność próby prowadzi do oceny wartości globalnej zmiennej y z błędem średnim mniejszym od 1000 zł.

Przedział ufności dla wartoći średniej ma postać:

$$[\bar{y}_s - z_\gamma d_s(\bar{y}_s); \bar{y}_s + z_\gamma d_s(\bar{y}_s)], \tag{1.43}$$

przy czym poziom ufności oznaczono przez γ, natomiast $d_s(\bar{y}_s)$ określają wzory (1.35) i (1.36). Ponadto, z_γ jest tak wyznaczone, aby $F(z_\gamma) = \frac{1+\gamma}{2}$, gdzie $F(.)$ jest dystrybuantą rozkładu normalnego standardowego. Określony przedział jest stosowany wtedy, gdy liczebność próby jest dostatecznie duża, ponieważ wówczas można ten fakt dowodzić za pomocą tzw. twierdzeń granicznych rozważanych m.in. przez Fullera (2009) i Hajka (1981).

Dodajmy, że przy tych samych założeniach przedział ufności dla wartości globalnej y_U ma postać:

$$[N\bar{y}_s - z_\gamma N d_s(\bar{y}_s).N\bar{y}_s + z_\gamma N d_s(\bar{y}_s)]. \tag{1.44}$$

Przykład 1.17

Przypuśćmy, że na podstawie zaobserwowanych wartości faktur sprzedaży w 1600-elementowej próbie prostej losowanej bezzwrotnie z populacji 80000--elementowej tych dokumentów, wyliczono oceny średniej i wariancji: $\bar{y}_s = 345.03$ zł, $v_s = 3685.55$. Na podstawie wzoru (1.35) wyliczamy ocenę błędu średniego szacunku, która wynosi

$$d_s(\bar{y}_s) = \sqrt{\frac{80000 - 1600}{1600 \cdot 80000} 3685.55} = 1.50 zł.$$

Zatem, przeciętnie rzecz biorąc, możliwe realizacje średniej z próby odchylają się od nieznanej wartości średniej faktury o 1.50 zł.

Przyjmując poziom ufności $\gamma = 0.95$, mamy, że $z_\gamma = 1.96$, ponieważ $F(1.96) = 0.975$. Te wyniki i wzór (1.43) prowadzą do przedziału ufności: $[342.09; 347.98]$, który z prawdopodobieństwem 0.95 obejmuje nieznaną przeciętną wartość faktury. Dokładność tej estymacji przedziałowej wynosi $\Delta = (347.98 - 342.09)/2 = 2.95$ zł.

Ocena punktowa wartości globalnej sprzedaży wynosi 27602400 zł z błędem średnim szacunku 120197.2 zł. Zatem oceny wartości globalnej sprzedaży z możliwych prób odchylają się od nieznanej wartości globalnej w populacji o 120197.2 zł. Z kolei wzór (1.43) prowadzi do oceny przedziałowej

wartości globalnej $I = [27366814; 27837986]$ z dokładnością do $\Delta = 235586.4$. Zatem z prawdopodobieństwem 0.95 przedział I obejmuje wartość globalną sprzedaży.

Wcześniej sygnalizowano, że wskaźnik struktury jest szczególnym przypadkiem średniej arytmetycznej, gdy zmienna y jest zero-jedynkowa, czyli $p = p_U = \frac{m_U}{N}$.

W związku z tym do oceny punktowej tego parametru używa się częstości względnej z próby, którą oznaczamy przez

$$p_s = \frac{m_s}{n}, \tag{1.45}$$

przy czym m_s jest liczbą elementów z cechą wyróżnioną w próbie (np. liczby dokumentów z nieprawidłowościami). Statystyka p_s jest nieobciążonym estymatorem wskaźnika struktury w populacji oznaczonego wcześniej przez p, a jej wariancja jest postaci:

$$V(p_s) = \frac{N-n}{N-1}\frac{p(1-p)}{n}. \tag{1.46}$$

Zatem błąd średni szacunku oraz jego estymator określają odpowiednio wzory

$$d(p_s) = \sqrt{\frac{N-n}{N-1}\frac{p(1-p)}{n}}, \qquad d_s(p_s) = \sqrt{\frac{N-n}{N-1}\frac{p_s(1-p_s)}{n}}. \tag{1.47}$$

Stąd natychmiastowo wynika, że błąd średni szacunku oraz jego estymator są z góry ograniczone przez

$$d_p = \frac{1}{2}\sqrt{\frac{N-n}{(N-1)n}}. \tag{1.48}$$

Przypuśćmy, że naszym zadaniem jest wyznaczenie takiej liczebności próby, która zapewni, iż błąd średni szacunku wskaźnika struktury nie przekroczy poziomu dopuszczalnego d_0. Nierówność $d_p \leq d_0$ daje szukaną liczebność:

$$n \geq n_0 = \frac{N}{1 + 4(N-1)d_0^2} + 1. \tag{1.49}$$

W sytuacji, gdy znamy oszacowanie wskaźnika struktury z wcześniejszych badań audytowych lub oszacowanego na podstawie próby wstępnej, to niezbędną liczebność próby proponuje się ustalić na poziomie $n_1 \leq n_0$, gdzie:

$$n_1 = \frac{N}{1 + \frac{(N-1)d_0^2}{p_s(1-p_s)}} + 1. \tag{1.50}$$

Gdy liczebność populacji jest duża, to $n_0 = \left[\frac{1}{4d_0^2}\right] + 1$, natomiast $n_1 = \left[\frac{p_s(1-p_s)}{d_0^2}\right] + 1$.

Aby wyznaczyć przedział ufności dla wskaźnika struktury $p = p_U = \frac{m_U}{N}$, trzeba znać rozkład prawdopodobieństwa statystyki p_s. Jak wiadomo, jej rozkład prawdopodobieństwa jest wyznaczany przez rozkład prawdopodobieństwa liczby elementów z cechą wyróżnioną obserwowaną w próbie, a którą oznaczaliśmy przez m_s, a ta ma tzw. rozkład hipergeometryczny, którego funkcja prawdopodobieństwa ma postać:

$$P(m_s = h) = \frac{\binom{m_U}{h}\binom{N-m_U}{n-h}}{\binom{N}{n}} = P\left(p_s = \frac{h}{n}\right). \tag{1.51}$$

Gdy liczebność próby jest dostatecznie duża (formalnie, gdy $N \to \infty$), to ten rozkład można dobrze przybliżyć rozkładem dwumianowym (Bernoulliego) postaci:

$$P(m_s = h) = P\left(p_s = \frac{h}{n}\right) = \binom{n}{h} p^h (1-p)^{n-h}. \tag{1.52}$$

Z kolei ten rozkład można przybliżyć rozkładem normalnym. Ze znanych tzw. twierdzeń lokalnego i integralnego de Moivre'a-Laplace'a (por. np. Gerstenkorn i Śródka, 1974), wynika, że jeśli $n \to \infty$, $N \to \infty$ oraz $N-n \to \infty$, to

$$P(m_s = h) = P\left(p_s = \frac{h}{n}\right) = \frac{1}{\sqrt{np(1-p)2\pi}} \exp\left(-\frac{(h-np)^2}{2np(1-p)}\right), \tag{1.53}$$

$$P(m_s < h) = P\left(p_s < \frac{h}{n}\right) = \varphi\left(\frac{h-np}{\sqrt{np(1-p)}}\right), \tag{1.54}$$

gdzie $\varphi(.)$ jest dystrybuantą standardowego rozkładu normalnego, czyli $\varphi(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} \exp\left(-\frac{u^2}{2}\right) dt$.

Te własności pozwalają już wyznaczyć przedział ufności postaci:

$$\left[p_s - u_\gamma d_s(p_s) < p < p_s + u_\gamma d_s(p_s)\right], \tag{1.55}$$

gdzie: $\varphi(u_\gamma) = \frac{1+\gamma}{2}$, natomiast $d_s(p_s)$ określa wyrażenie (1.47). Otrzymany przedział w przybliżeniu z prawdopodobieństwem γ obejmuje wartość częstości względnej z populacji p. Dokładność tego estymatora przedziałowego, będąca połową długości przedziału ufności, wynosi $\Delta = u_\gamma d_s(p_s) \leq u_\gamma d_p = \Delta_p$, gdzie d_p określa wzór (1.48).

Określony przedział ufności zaleca się stosować dla liczebności próby $n \geq 100$ i dla $p \geq 0.04$.

Gdy postulowana dokładność estymacji ma nie przekroczyć wartości Δ_0, natomiast poziom ufności ma wynosić γ, to z nierówności $\Delta_p \leq \Delta_0$ wynika, że potrzebny rozmiar próby wynosi:

$$n \geq n_2 = \left[\frac{1}{\frac{1}{N} + 4\left(\frac{d_0}{u_\gamma}\right)^2} \right] + 1. \tag{1.56}$$

Gdy mamy oceny wstępne p_s wartości parametru p, to $n_2 \geq n_3$, gdzie

$$n_3 = \left[\frac{1}{\frac{1}{N} + \frac{d_0^2}{u_\gamma^2 p_s(1-p_s)}} \right] + 1. \tag{1.57}$$

W końcu, gdy rozmiar populacji jest duży, to $n_2 = \left[\left(\frac{u_\gamma}{2d_0}\right)^2 \right] + 1$, natomiast $n_3 = \left[\frac{u_\gamma^2 p_s(1-p_s)}{d_0^2} \right] + 1$.

Przykład 1.18

Ważnym wskaźnikiem jakości dokumentów księgowych jest udział dokumentów z błędami wśród wszystkich dokumentów. Dotychczasowo prowadzone badania kontrolne wykazywały, że udział faktur z błędami wśród kontrolowanych dokumentów wynosił 0.08. Zachodzi potrzeba wylosowania bezzwrotnego próby prostej spośród 10 000 faktur o liczebności zapewniającej precyzję oceny punktowej tego wskaźnika z dokładnością do poziomu 0.01.

Na podstawie wzoru (1.50) mamy:

$$n_1 = \left[\frac{10000}{1 + \frac{1}{0.08(1-0.08)}} \right] + 1 = 686.$$

Gdy dodatkowo mamy zamiar oszacować przedziałowo częstość występowania błędnych faktur na poziomie ufności 0.95 z dokładnością do 0.01, to na podstawie wzoru (1.57) wyliczamy, iż należy wylosować próbę o liczebności

$$n_3 = \left[\frac{1}{0.0001 + \frac{0.01^2}{0.08(1-0.08)1.96}} \right] + 1 = 1261,$$

przy czym $P(U < 1.96) = 0.975$.

1.5. Elementy weryfikacji hipotez

Występowanie błędów w zapisach księgowych jest zwykle w praktyce rzeczą nie do uniknięcia, o czym już pisano wcześniej. Rozmiar tych błędów

charakteryzowany odpowiednim parametrem (wskaźnikiem) jakości jednak nie jest dowolny. Przyjmując, że wskaźnik jakości rośnie wraz ze wzrostem rozmiaru błędu, jego górną dopuszczalną wielkość określa audytor, co już wcześniej omawiano. Do rutynowych zadań audytora należy ocena, czy obserwowany rozmiar wskaźnika jakości przekracza dopuszczalny poziom. Ta rzecz staje się trudniejsza, gdy mamy do czynienia z problemem kontroli jakości dużej ilości dokumentów. Wówczas na podstawie próby można ocenić wartość występującego w niej wskaźnika jakości, czym zajmowaliśmy się wcześniej. Występuje tu jednak pewien problem; oszacowany na podstawie próby wskaźnik jakości może przekroczyć wartość dopuszczalną albo jej nie przekroczyć. Jednak nieznana pozostaje wartość tego wskaźnika w całej populacji dokumentów. W szczególności teraz trzeba zdecydować, czy w sytuacji, gdy wskaźnik jakości obliczony na podstawie próby przekracza dopuszczalną wartość, należy również przyjąć, że wskaźnik wyliczony na podstawie całej populacji, również przekroczy wartość dopuszczalną. Tak nie musi być, bowiem fakt, że wartość z próby wskaźnika przekracza wartość dopuszczalną jest rezultatem losowego doboru próby.

Precyzyjniej ten problem opiszemy, przyjmując następujące założenia. Niech θ będzie parametrem charakteryzującym jakość dokumentów (np. częstość względną występowania w nich błędów). Zbiór możliwych wartości tego parametru oznaczymy przez Θ, czyli $\theta \in \Theta$. W zbiorze Θ wyróżnia się dwa wykluczające się podzbiory konkurujących ze sobą wartości, które oznaczamy przez Θ_0 i Θ_1. W szczególności te podzbiory mogą być pojedynczymi wartościami, które oznaczymy przez θ_0 i θ_1. Chodzi o to, by rozstrzygnąć, do którego z tych pozdbiorów należy wartość parametru θ. Zatem zasadne jest formułowanie następujących hipotez. Hipotezę sprawdzaną oznaczymy przez

$$H_0 : \ \theta \in \Theta_0,$$

a do niej hipotezę alternatywną (konkurencyjną), przez

$$H_1 : \ \theta \in \Theta_1.$$

W celu rozstrzygnięcia, która z określonych hipotez jest prawdziwa, stosuje się procedurę wnioskowania statystycznego, zwaną *testem statystycznym*. Test na podstawie wyników obserwacji zmiennej w próbie pozwala w sposób racjonalny wskazać, którą z hipotez – H_0 czy H_1 – można uznać za prawdziwą. Test jednak może doprowadzić do błędnych decyzji co do uznania prawdziwości formułowanych hipotez. Jeśli test odrzuca prawdziwą hipotezę sprawdzaną H_0, to popełniamy tzw. *błąd pierwszego rodzaju*. Z kolei *błąd drugiego rodzaju* polega na tym, że przyjmujemy hipotezę H_0, gdy jest fałszywa, czyli gdy hipoteza H_1 jest prawdziwa (Tablica 5.1).

Tablica 5.1

Rodzaje błędów

	prawdziwa H_0	prawdziwa H_1
przyjęta H_0	słuszna decyzja	błąd II rodzaju
przyjęta H_1	błąd I rodzaju	słuszna decyzja

Prawdopodobieństwo popełnienia błędu I rodzaju jest nazywane *istotnością (znamiennością) testu* i jest oznaczane przez α, przy czym

$$\alpha = P(\text{odrzucamy } H_0, \text{ gdy } H_0 \text{ jest prawdziwa}).$$

Moc testu, oznaczana przez β, to prawdopodobieństwo niepopełnienia błędu II rodzaju, czyli

$$\beta = P(\text{odrzucamy } H_0, \text{ gdy } H_1 \text{ jest prawdziwa}).$$

Prawdopodobieństwo popełnienia błędu II rodzaju jest oznaczane przez ν, przy czym

$$\nu = P(\text{przyjmujemy } H_0, \text{ gdy } H_1 \text{ jest prawdziwa}) = 1 - \beta.$$

Po to, aby przedstawić ,,działanie"testu statystycznego, rozważmy następujący akademicki problem. Zakładamy, że będziemy testowali wartość oczekiwaną (średnią) zmiennej losowej X, która ma rozkład normalny ze średnią równą właśnie μ i wariancją wynoszącą σ^2, co krótko zapisujemy wyrażeniem $Y \sim N(\mu, \sigma^2)$. Zadanie polega na weryfikacji hipotezy sprawdzanej głoszącej, że wartość przeciętna zmiennej losowej o rozkładzie normalnym jest równa ustalonej wartości μ_0, co zapisujemy symbolicznie jak następuje:

$$H_0 : \ \mu = \mu_0, \qquad N(\mu, \sigma^2).$$

Hipotezę alternatywną określa wyrażenie

$$H_1 : \ \mu = \mu_1 > \mu_0, \qquad N(\mu, \sigma^2),$$

co oznacza, że konkurencyjną do wartości średniej μ_0 jest większa od niej wartość μ_1. Ponadto, zakładamy, że zarówno w przypadku hipotezy H_0, jak i H_1 wariancja przyjmuje tę samą wartość.

W celu podjęcia decyzji o prawdziwości hipotezy sprawdzanej wyznacza się na podstawie próby prostej wartość tzw. *sprawdzianu testu (statystyki testowej)*, którą tradycyjnie oznaczamy symbolem $\bar{Y}_s$. Rozsądek wskazuje, że wartość średniej z próby $\bar{Y}_s$ istotnie większa od hipotetycznej średniej μ_0 świadczy na korzyść hipotezy alternatywnej H_1 kosztem hipotezy sprawdzanej H_0. Wartość krytyczną y_α, której przekroczenie skutkuje odrzuceniem

hipotezy H_0 na korzyść H_1 wyznacza się tak, aby było równe poziomowi istotności α prawdopodobieństwo jej przekroczenia przez sprawdzian testu $\bar{Y}_s$ przy założeniu, że prawdziwą jest hipoteza H_0, czyli

$$\alpha = P(\bar{Y}_s \geq y_\alpha | H_0) = P(Z \geq z_\alpha),$$

gdzie:

$$y_\alpha = \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}},$$

przy czym zmienna losowa Z ma tzw. rozkład normalny standardowy, co zapisujemy symbolicznie przez $Z \sim N(0,1)$. Moc testu jest liczona na podstawie wyrażenia

$$\beta = P(\bar{Y}_s \geq y_\alpha | H_1) = P\left(Z \geq z_\beta\right),$$

gdzie:

$$z_\beta = z_\alpha + \frac{\mu_0 - \mu_1}{\sigma}\sqrt{n}.$$

Gdy $\bar{y}_s \geq y_\alpha$, to odrzucamy H_0 na korzyść H_1 z prawdopodobieństwem popełnienia błędnej decyzji równym α. W przeciwnym przypadku, jeśli $\bar{y}_s < y_\alpha$, to przyjmujemy hipotezę H_0 za prawdziwą z prawdopodobieństwem popełnienia błędnej decyzji równym $\nu = 1 - \beta$.

Zwykle w praktyce wygodnie jest podejmować decyzje na podstawie sprawdzianu

$$Z_s = \frac{\bar{Y}_s - \mu_0}{\sigma}\sqrt{n},$$

który jest standardową postacią średniej z próby. Wtedy

$$\alpha = P(Z_s \geq z_\alpha | H_0) = P(Z \geq z_\alpha),$$

$$\beta = P(Z_s \geq z_\beta | H_1) = P\left(Z \geq z_\beta\right).$$

Wówczas, podobnie jak w przypadku sprawdzianu $\bar{y}_s$, hipotezę H_0 odrzucamy na korzyść H_1, gdy zaobserwowana wartość z_s sprawdzianu Z_s jest nie mniejsza od wartości z_α, która teraz jest wartością krytyczną testu. W tym przypadku prawdopodobieństwo popełnienia błędnej decyzji jest równe α. W przeciwnym przypadku, jeśli $z_s < z_\alpha$, to przyjmujemy hipotezę H_0 z prawdopodobieństwem popełnienia błędnej decyzji równym $\nu = 1 - \beta$.

Wygodnie jest również podejmować decyzję na podstawie tzw. p-wartości, zwanej czasem obserwowanym poziomem istotności (znamienności) testu, którą wyznaczamy na podstawie wyrażenia:

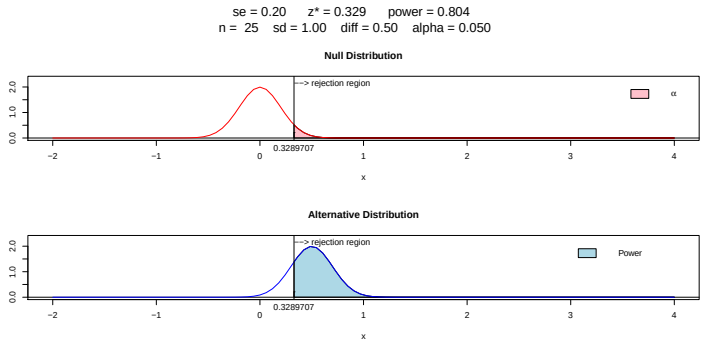
$$p = P(Z_s \geq z_s | H_0) = P(Z \geq z_s).$$

Wtedy hipotezę H_0 odrzucamy na korzyść H_1, gdy zaobserwowana p-wartość jest nie większa od poziomu istotności α. W tym przypadku prawdopodobieństwo popełnienia błędnej decyzji jest również równe α. W przeciwnym przypadku, jeśli $p > \alpha$, to przyjmujemy hipotezę H_0 za prawdziwą z prawdopodobieństwem popełnienia błędnej decyzji równym $\nu = 1 - \beta$. Wnioskowanie na podstwie p-wartości jest powszechne w przypadku posiłkowania się komputerowymi pakietami statystycznymi.

Na zakończenie dodajmy, że w praktyce nie zawsze jest możliwe wyliczenie mocy testu β, a co za tym idzie również prawdopodobieństwa błędu II rodzaju, oznaczanego przez ν. Wtedy, w sytuacji gdy $p > \alpha$, co w naszym przypadku jest równoważne nierównościom $z_s < z_\alpha$ lub $\bar{y}_s < y_\alpha$, to wartość ν prawdopodobieństwa popełnienia błędu II rodzaju nie jest wyliczana. Zatem nie możemy w tym przypadku ocenić, czy błąd II rodzaju jest na tyle mały, że możemy zaniedbując go przyjąć hipotezę sprawdzaną H_0 za prawdziwą. Wówczas zwykle pisze się, iż nie ma podstaw do odrzucenia hipotezy sprawdzanej.

Przykład 1.19

Dydaktyczną prezentację poziomu istotności i mocy testu w pakiecie R realizujemy następująco. Najpierw należy wybrać okno 'Packages', potem 'Install packages' i stąd instalujemy program 'Teachingdemos'. Następnie, ponownie z okna 'Packages' wybieramy opcję 'Load package', 'Teachingdemos'. W końcu wpisujemy na konsolę:
$power.examp(n = \ldots, s = \ldots, diff = \ldots, alpha = \ldots)$,
gdzie:
n – liczebność próby,
s – wariancja,
$diff$ – różnica między przeciętnymi hipotetycznymi, czyli $\mu_1 - \mu_0$, specyfikowanymi przez hipotezy H_0 i H_1,
α – poziom istotności.
Wówczas wynik realizacji np. instrukcji $power.examp(n = 25, s = 1, diff = 0.5, alpha = 0.05)$ prezentuje rys. 2.1.

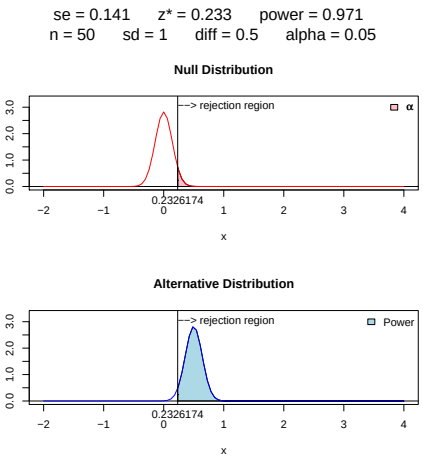


Rys. 2.1. Moc testu dla $n = 25$, $\sigma = 1$, $\Delta = 0.5$ i $\alpha = 0.05$

Rysunek 2.1 dotyczy weryfikacji hipotezy H_0 o wartości średniej zmiennej losowej o rozkładzie normalnym z nadzieją matematyczną równą μ i odchyleniem standardowym równym $\sigma = sd = 1.00$. Hipoteza sprawdzana głosi, że $\mu = \mu_0 = 0$, natomiast hipoteza alternatywna, że $\mu = \mu_1 = \Delta = diff = 0.5$, gdzie $\Delta = diff = \mu_1 - \mu_0$. Zatem równoważnie nasze hipotezy sprawdzaną i alternatywną można zapisać odpowiednio wzorami: $H_0 : \Delta = 0$ i $H_1 : \Delta = 0.5$. Do weryfikacji tych hipotez użyto próby prostej składającej się z $n = 25$ obserwacji. Zatem błąd średni szacunku przeciętnej w populacji μ, estymowanej za pomocą średniej z próby, wynosi $d = \frac{\sigma}{\sqrt{n}} = 0.2$. Wyznaczona przy poziomie istotności $\alpha = alpha = 0.05$ wartość krytyczna testu wynosi $y_\alpha = 0.329$. Obszar krytyczny testu *rejection region* jest prawostronny. Przy tych założeniach moc testu wynosi $\beta = power = 0.804$. Stąd więc wynika, że prawdopodobieństwo popełnienia błędu II rodzaju wynosi $\nu = 1 - \beta = 0.196$. W końcu proces podjęcia decyzji jest następujący. Jeśli wyliczona na podstwie próby średnia jest nie mniejsza od wartości krytycznej y_α, to odrzucamy hipotezę H_0 na korzyść H_1 z prawdopodobieństwem popełnienia błędnej decyzji równym $\alpha = 0.05$. W przeciwnym przypadku, gdy średnia z próby jest mniejsza od wartości krytycznej y_α, to przyjmujemy hipotezę H_0 z prawdopodobieństwem popełnienia niewłaściwej decyzji równym $\nu = 0.196$.

Zmiana wartości jednego z wyjściowych parametrów, przy których jest prowadzona weryfikacja hipotez skutkuje zmianą pozostałych wskaźników, co można sprawdzić, uruchamiając wyżej opisany program komputerowy. Pewne szczególne wyniki w tym zakresie prezentują rys. 2.2-2.5.

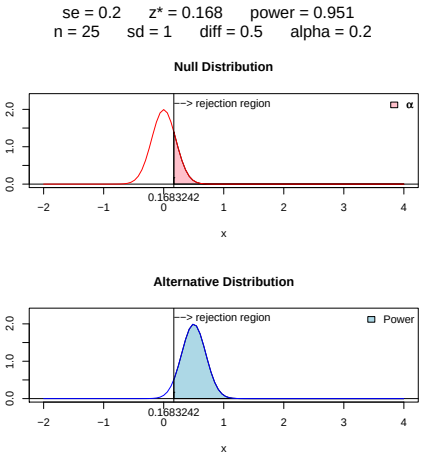
Przykładowo z rys. 2.2 wynika, że dwukrotny wzrost liczebności próby do rozmiaru $n = 50$ prowadzi do zwiększenia mocy testu z liczby $\beta = 0.804$ do poziomu $\beta = 0.971$, co jest równoważne zmniejszeniu prawdopodobieństwa popełnienia błędu drugiego rodzaju z poziomu $\nu = 0.196$ do liczby $\nu = 0.029$.



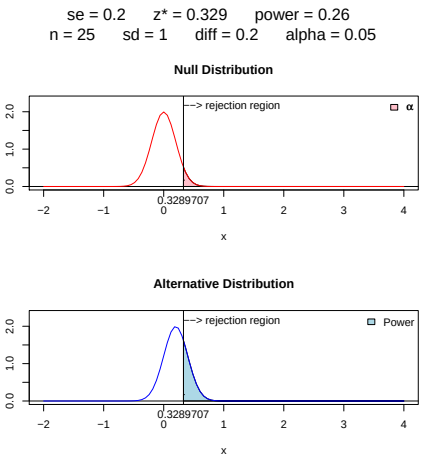
Rys. 2.2. Moc testu dla $n = 50$, $\sigma = 1$, $\Delta = 0.5$ i $\alpha = 0.05$

Podobnie, zwiększając poziom istotności testu (czyli prawdopodobieństwa popełnienia błędu I rodzaju) z liczby $\alpha = 0.05$ do poziomu $\alpha = 0.2$, doprowadzamy do wzrostu jego mocy z poziomu $\beta = 0.804$ do liczby $\beta = 0.951$, na co wskazują dane z rys. 2.3. W tym przypadku prawdopodobieństwo popełnienia błędu II rodzaju spada do poziomu $\nu = 0.041$, lecz dzieje się to kosztem zwiększenia prawdopodobieństwa popełnienia błędu I rodzaju.

Wpływ zmiany parametru $\Delta = \mu_1 - \mu_0$ (czyli różnicy między wartościami oczekiwanymi specyfikowanymi przez hipotezy odpowiednio H_1 i H_0) na moc testu prezentują dane z rys. 2.4. W tym przypadku zmniejszenie tej różnicy z poziomu $\Delta = 0.5$ do liczby $\Delta = 0.2$ prowadzi do dramatycznego spadku mocy testu z liczby $\beta = 0.804$ do poziomu $\beta = 0.26$, co pociąga za sobą wzrost prawdopodobieństwa popełnienia błędu II rodzaju z liczby $\nu = 0.196$ aż do poziomu $\nu = 0.74$.

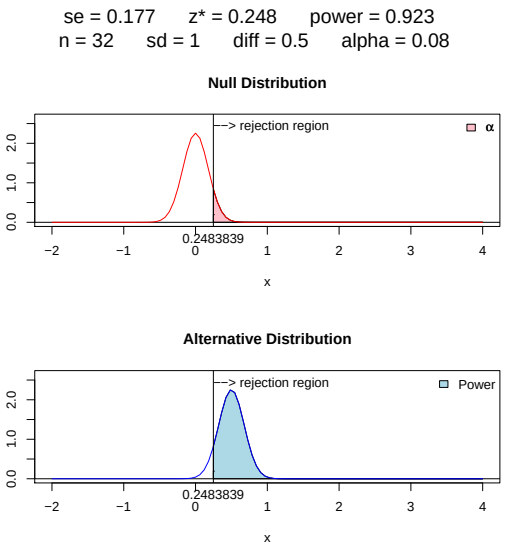


Rys. 2.3. Moc testu dla $n = 25$, $\sigma = 1$, $\Delta = 0.5$ i $\alpha = 0.2$



Rys. 2.4. Moc testu dla $n = 25$, $\sigma = 1$, $\Delta = 0.5$ i $\alpha = 0.2$

W praktyce statystyk może jedynie zmieniać w pewnych rozsądnych granicach rozmiar próby oraz poziom istotności testu, co skutkuje zmianą prawdopodobieństwa popełnienia błędu II rodzaju. Przykładem tego są dane z rys. 2.5, z których wynika zmiana rozmiaru próby z liczby $n = 25$ do poziomu $n = 32$ oraz poziomu istotności z liczby $\alpha = 0.05$ do poziomu $\alpha = 0.08$ prowadzi do wzrostu mocy testu z poziomu $\beta = 0.804$ do poziomu $\beta = 0.923 \approx 0.92$. Tak ustawione parametry dają więc w przybliżeniu równość poziomów prawdopodobieństw popełnienia błędów I i II rodzaju, bowiem $\nu = 0.077 \approx 0.08 = \alpha$.



Rys. 2.5. Moc testu dla $n = 32$, $\sigma = 1$, $\Delta = 0.5$ i $\alpha = 0.08$

Rozdział 2

BADANIE ZGODNOŚCI

2.1. Wprowadzenie

Niniejszy rozdział odnosi się do problemów zastosowania wnioskowania statystycznego w zagadnieniach dotyczących audytu. Zajmiemy się wnioskowaniem o proporcji p_U (udziału, częstości względnej) nieprawidłowości występujących w populacji U dokumentów. Występowanie tych nieprawidłowości ujawnia się poprzez występowanie błędów formalnych, jak np. brak odpowiednich podpisów, lub błędów numerycznych, jak np. nieprawidłowo dodane kwoty pieniędzy czy też źle przepisane. Do tych nieprawidłowości można jeszcze zaliczyć niekompetencję osób sygnujących dane dokumenty lub nieprawidłową klasyfikację dokumentów bądź ich obieg w systemie biurokratycznym firmy. Tego typu nieprawidłowości powinny być wykrywane przez system kontroli wewnętrznej, której istnienie ma stymulować prawidłowe funkcjonowanie systemu księgowego. W związku z tym zachodzi potrzeba określenia dopuszczalnego poziomu udziału nieprawidłowości, który będzie oznaczany przez p_0. Oznacza to, że proporcja nieprawidłowości nieprzekraczająca poziomu p_0 jest akceptowalna. Przez $p_1 > p_0$ oznaczymy poziom nieprawidłowości dyskwalifikujący, którego przekroczenie w żaden sposób nie może być akceptowane. Zatem, jeśli proporcja nieprawidłowości nie przekroczy poziomu p_0, to system kontroli wewnętrznej jest uznawany za działający skutecznie (poprawnie). W przypadku gdy udział wykrytych nieprawidłowości przekroczy poziom dyskwalifikujący, system kontroli wewnętrznej uznaje się za niewydolny (nieskuteczny).

Spotyka się jeszcze następującą interpretację parametru p. Mianowicie traktuje się go jako proporcję dopuszczalnej nieprawidłowości, lecz w sensie przeciętnego poziomu występowania nieprawidłowości spotykanych w audycie określonego systemu kontroli wewnętrznej. Zatem można spodziewać się również większych wartości tej częstości względnej występowania nieprawidłowości, ale na pewno nieprzekraczających wartości niedopuszczalnej p_1, która dyskwalifikuje system działania kontroli wewnętrznej.

Zadanie audytora polega na autorytatywnym stwierdzeniu, czy system kontroli wewnętrznej jest skuteczny, czy nieskuteczny. Jego decyzje mogą

jednak być nieprawidłowe, ponieważ parametr p jest oceniany na podstawie próby losowej. Po pierwsze może stwierdzić prawidłowość działania systemu kontroli wewnętrznej, gdy tak nie jest. Prawdopodobieństwo wystąpienia tej sytuacji jest nazywane *ryzykiem aprobaty nieskutecznego systemu kontroli* i oznaczane przez κ, a przez η będzie oznaczane *ryzyko dezaprobaty skutecznego systemu kontroli wewnętrznej*, czyli prawdopodobieństwo tego, że zostanie stwierdzona nieskuteczność systemu kontroli, gdy nie jest to prawdą (por. np. Karliński, 2005). Dodajmy, że Hołda i Pociecha (2009) parametr κ nazywają ryzykiem nieprawidłowej akceptacji, natomiast η – ryzykiem nieprawidłowego odrzucenia.

2.2. Analiza skuteczności systemu kontroli przy ustalonych obu rodzajach ryzyka

Tłumacząc wyżej postawione audytorowi zadanie na język testowania hipotez statystycznych, mamy do czynienia z problemem weryfikacji hipotezy sprawdzanej H_0, głoszącej, że proporcja nieprawidłowości nie przekracza dopuszczalnego poziomu p_0 przeciwko hipotezie alternatywnej głoszącej, iż ta proporcja jest nie miejsza niż poziom dyskwalifikujący p_1, co krótko zapisujemy wyrażeniem:

$$H_0 : \ p_U \leq p_0, \qquad H_1 : \ p_U \geq p_1 > p_0 \tag{2.1}$$

Zakładamy, że poziom istotności α testu jest równy ryzyku η dezaprobaty skutecznego system kontroli, czyli $\alpha = \eta$. Z kolei prawdopodobieństwo popełnienia błędu II rodzaju ν jest równe ryzyku κ aprobaty nieskutecznego systemu kontroli, czyli $\kappa = \nu$.

Określone teraz hipotezy zastępuje się następującymi do nich równoważnymi hipotezami.

$$H_0 : \ M_U \leq M_0, \qquad H_1 : \ M_U \geq M_1 > M_0, \tag{2.2}$$

przy czym przez $M_U = Np$, $M_0 = Np_0$ i $M_1 = Np_1$, gdzie N jest rozmiarem populacji. Przez M_U jest oznaczana np. liczba dokumentów księgowych, w których występują nieprawidłowości. Z kolei hipotetyczne liczby nieprawidłowych dokumentów oznacza się przez M_0 i M_1, przy czym pierwszy symbol jest dopuszczalnym poziomem tych nieprawidłowości, a drugi dyskwalifikującym poziomem nieprawidłowości.

Problem weryfikacji postawionej hipotezy rozpoczniemy od przypadku, gdy statystykę testową można przybliżyć rozkładem normalnym prawdopodobieństwa. W tym przypadku jest możliwe analityczne wyznaczenie wartości krytycznej testu wraz z niezbędną liczebnością próby prostej losowanej bezzwrotnie zapewniającą podejmowanie decyzji o przyjęciu albo

odrzuceniu hipotezy H_0 z zadanymi ryzykami η i κ. Dopiero potem zajmiemy się pozostałymi przypadkami, począwszy od sytuacji, gdy rozkład sprawdzianu testu jest rozkładem dokładnym.

2.2.1. Weryfikacja hipotezy z wykorzystaniem aproksymacji sprawdzianu testu rozkładem normalnym prawdopodobieństwa

Weryfikacją sformułowanej wyrażeniem (2.1) hipotezy zajmiemy się najpierw w najprostszym przypadku, gdy dopuszczalna proporcja nieprawidłowości p_0 jest co najmniej równa 0.04. Wtedy (por. np. Gerstenkorn i Śródka, 1974) dla co najmniej 100-elementowej próby prostej losowanej bezzwrotnie z populacji rozkład prawdopodobieństwa częstości względnej z próby p_s (por. wzór (1.45)), jest dostatecznie dobrze przybliżany rozkładem normalnym. To już pozwala na stosowanie następującej procedury.

Zdefiniowane ryzyka podejmowania błędnych decyzji określają wzory:

$$\eta = \alpha = P(p_s \geq p_\eta | H_0) = P(Z_s \geq z_\eta | H_0) = P(Z \geq z_\eta) = 1 - \phi(z_\eta),$$

$$\kappa = \nu = P(p_s < p_\eta | H_1) = P(Z_s < z_\eta | H_1) = P(Z < z_\kappa) = \phi(z_\kappa),$$

gdzie:

$$p_\eta = p_0 + z_\eta \sqrt{\frac{N-n}{N-1} \frac{p_0(1-p_0)}{n}}, \tag{2.3}$$

$$z_\kappa = \frac{p_0 - p_1}{\sqrt{\frac{N-n}{N-1} \frac{p_1(1-p_1)}{n}}} + z_\eta \sqrt{\frac{p_0(1-p_0)}{p_1(1-p_1)}}, \tag{2.4}$$

$$Z_s = \frac{p_s - p_0}{\sqrt{\frac{N-n}{(N-1)n} p_0(1-p_0)}}, \tag{2.5}$$

przy czym $\phi(z) = P(Z < z)$ jest dystrybuantą rozkładu $Z \sim N(0;1)$.

Rozwiązując równanie (2.4) względem n, łatwo wykazać, że po to by nie przekroczyć rodzajów ryzyka: η i κ, należy wylosować próbę o liczebności co najmniej: $[n_*] + 1$, gdzie:

$$n_* = N \left(1 + \frac{(N-1)(p_1 - p_0)^2}{\left(z_\eta \sqrt{p_0(1-p_0)} - z_\kappa \sqrt{p_1(1-p_1)} \right)^2} \right)^{-1}, \tag{2.6}$$

przy czym $P(Z < z_\kappa) = \kappa = \nu$, $Z \sim N(0;1)$, natomiast przez $[n_*]$ oznaczono część całkowitą liczby n_*. Gdy liczebność populacji jest duża, to powyższy wzór upraszcza się do następującej postaci, por. np. Ryan (2013):

$$n_{**} = \frac{\left(z_\eta \sqrt{p_0(1-p_0)} - z_\kappa \sqrt{p_1(1-p_1)}\right)^2}{(p_1 - p_0)^2}. \tag{2.7}$$

Załóżmy, że zarówno ryzyko akceptacji nieskutecznego systemu kontroli, jak i ryzyko dezaprobaty skutecznego systemu kontroli są mniejsze od 0.5, czyli $\kappa < 0.5$ oraz $\eta < 0.5$. Wówczas, gdy optymalną liczebność próby określoną wzorem (2.6) podstawimy w miejsce liczebności n do wzoru (2.3), to otrzymujemy wartość krytyczną testu, którą oznaczamy przez $p_{*\eta}$. W szczególności gdy liczebność populacji jest duża, to ze wzoru (2.7) i prawej strony równania (2.3) uproszczonego do postaci $z_\eta \sqrt{\frac{p_0(1-p_0)}{n}}$ wynika, iż

$$p_{*\eta} = \frac{p_1 z_\eta \sqrt{p_0(1-p_0)} - p_0 z_\kappa \sqrt{p_1(1-p_1)}}{z_\eta \sqrt{p_0(1-p_0)} - z_\kappa \sqrt{p_1(1-p_1)}}. \tag{2.8}$$

Reguły postępowania są następujące. Gdy zaobserwowana w próbie częstość p_s jest nie mniejsza niż wartość krytyczna $p_{*\eta}$, czyli $p_s \geq p_{*\eta}$, co jest równoważne nierówności $z_s \geq z_\eta$, to odrzucamy H_0 na korzyść H_1 z prawdopodobieństwem popełnienia błędnej decyzji (ryzykiem) η. Oznacza to, iż z prawdopodobieństwem pomyłki η stwierdzamy, że system kontroli wewnętrznej jest nieskuteczny.

Gdy $p_s < p_{*\eta}$, co jest równoważne nierówności $z_s < z_\eta$, to przyjmujemy H_0 z prawdopodobieństwem popełnienia błędnej decyzji (ryzykiem) κ. Oznacza to, iż wnioskujemy, że system kontroli wewnętrznej jest skuteczny z prawdopodobieństwem pomyłki κ.

Dodajmy jeszcze, że p-wartość testu określa wzór:

$$p = P(Z_s \geq z_s | H_0) = P(Z \geq z_s). \tag{2.9}$$

Zatem, gdy $p \leq \eta = \alpha$, to odrzucamy H_0 na korzyść H_1 z prawdopodobieństwem popełnienia błędnej decyzji (ryzykiem) η. Gdy $p > \eta = \alpha$, to przyjmujemy H_0 z prawdopodobieństwem (ryzykiem) popełnienia błędnej decyzji κ.

Przykład 2.1

System księgowy pewnego przedsiębiorstwa wytworzył w ciągu roku $N = 10543$ dokumenty, w których mogą występować nieprawidłowości. Prowadzący kontrolę audytor ustalił dopuszczalną proporcję nieprawidłowości na poziomie 0.04, natomiast dyskwalifikującą proporcję nieprawidłowości na poziomie 0.06, czyli $p_0 = 0.04$ i $p_1 = 0.06$. Skuteczność systemu kontroli wewnętrznej jest weryfikowana poprzez testowanie hipotezy określonej wyrażeniem (2.1). Audytor ustala ryzyko dezaprobaty skutecznego systemu kontroli wewnętrznej na poziomie $\eta = 0.05$, ryzyko aprobaty nieskutecznego systemu kontroli wewnętrznej na poziomie $\kappa = 0.05$. Po to, by ta

weryfikacja dała podstawy do podjęcia decyzji z oboma rodzajami ryzyka nieprzekraczającymi ustalonych poziomów, trzeba wylosować bezzwrotnie próbę prostą o rozmiarze określonym wzorem (2.6). W wyniku przeprowadzonych obliczeń otrzymujemy, że $n \geq 1135$. Ta liczebność próby jest jednak zbyt duża, gdyż nie ma ani czasu, ani pieniędzy na kontrolowanie tylu dokumentów. W związku z tym audytor godzi się na zwiększenie ryzyka dezaprobaty skutecznego systemu kontroli wewnętrznej do poziomu $\eta = 0.2$. Ta zmiana prowadzi do niezbędnej liczebności próby $n \geq 720 = n_*$.

Teraz czas na wylosowanie próby i wyznaczenie wartości sprawdzianu testu p_s, określonego wzorem (2.5). Następnie trzeba go porównać z wartością krytyczną określoną wzorem (2.3), do którego należy w miejsce liczebności n podstawić właśnie wyliczony rozmiar optymalny próby $n_* = 720$. Wówczas otrzymujemy $p_\eta = 0.0627$. Wtedy, jeśli zajdzie nierówność $p_s \geq p_\eta$, to dezaprobujemy system kontroli wewnętrzenej z prawdopodobieństwem popełnienia błędnej decyzji równym 0.2. Gdy zajdzie odwrotna sytuacja, czyli prawdziwa byłaby nierówność $p_s < p_\eta$, to aprobujemy system kontroli wewnętrznej z prawdopodobieństwem popełnienia błędnej decyzji równym 0.05. Przypuśćmy, że wartość statystyki testowej wyniosłaby $p_s = 0.0423$, czyli $p_s = 0.0423 < 0.0627 = p_\eta$. Zatem akceptujemy system kontroli z prawdopodobieństwem popełnienia błędnej decyzji równym 0.05. W sposób równoważny, do tej samej decyzji dochodzimy na podstawie p-wartości określonej wzorem (2.9). Wyliczamy ją następująco:

$$p = P\left(Z \geq \frac{p_s - p_0}{\sqrt{\frac{N-n}{(N-1)n} p_0(1-p_0)}} = 0.0852\right) = 0.4661 > \eta.$$

Ta nierówność prowadzi do tej samej decyzji, co wcześniej wyznaczona nierówność $p_s = 0.0423 < 0.0627 = p_\eta$.

Problem podejmowania decyzji o skuteczności systemu kontroli wewnętrznej można również rozwiązać poprzez weryfikację następujących hipotez:

$$H_0 : \ p_U \geq p_1, \qquad H_1 : \ p_U \leq p_0 < p_1. \tag{2.10}$$

W pewnym sensie te hipotezy są przeciwstawne w stosunku do tych określonych wyrażeniem (2.1). W tym przypadku hipoteza sprawdzana H_0 głosi, że proporcja nieprawidłowości w populacji p_U jest nie mniejsza od dyskwalifikującego poziomu równego p_1 przeciwko hipotezie alternatywnej H_1 głoszącej, że ta proporcja jest nie większa od dopuszczalnej proporcji nieprawidłowości.

Sprawdzianem testu sformułowanych hipotez jest również częstość względna występowania nieprawidłowości w próbie oznaczanej przez p_s. Tym razem poziom istotności α testu, czyli prawdopodobieństwo popełnienia błędu I rodzaju jest równe ryzyku κ aprobaty nieskutecznego system kontroli, czyli

$\alpha = \kappa$. Zatem prawdopodobieństwo popełnienia błędu II rodzaju jest równe ryzyku η dezaprobaty skutecznego systemu kontroli, czyli $\nu = \eta$. Te ryzyka określają wzory:

$$\kappa = \alpha = P(p_s \leq p_\kappa | H_0) = P(Z_s \leq z_\kappa | H_0) = \phi(z_\kappa),$$
$$\eta = \nu = P(p_s > p_\kappa | H_1),$$

gdzie:

$$p_\kappa = p_1 + z_\kappa \sqrt{\frac{N-n}{N-1} \frac{p_1(1-p_1)}{n}}, \tag{2.11}$$

$$Z_s = \frac{p_s - p_1}{\sqrt{\frac{N-n}{(N-1)n} p_1(1-p_1)}}, \tag{2.12}$$

przy czym $\phi(z_\kappa) = P(Z < z_\kappa) = \kappa = \alpha$, $Z \sim N(0;1)$.

Można wykazać, że w rozważanym przypadku optymalną liczebność próby określają również wzory (2.6) lub (2.7). Wartość krytyczną $p_{*\kappa}$ otrzymujemy ze wzoru (2.11), podstawiając do niego w miejsce liczebności n prawą stronę wzoru (2.6). W szczególności, jeśli liczebność populacji jest duża, to podobnie jak w przypadku wyprowadzenia wzoru (2.8), otrzymujemy:

$$p_{*\kappa} = \frac{p_1 z_\eta \sqrt{p_0(1-p_0)} - p_0 z_\kappa \sqrt{p_1(1-p_1)}}{z_\eta \sqrt{p_0(1-p_0)} - z_\kappa \sqrt{p_1(1-p_1)}}. \tag{2.13}$$

Na podstawie równości $P(Z < z) = P(Z > -z)$ można wykazać, że $p_{*\kappa} = p_{*\eta}$, gdzie $p_{*\eta}$ określa wzór (2.8). Z tego wynika, że rezultaty testowania hipotez określonych wyrażeniami (2.1) i (2.10), są sobie równoważne. Zatem wystarczy wybrać i weryfikować jeden z tych dwóch sposobów formułowania hipotez.

Gdybyśmy sformułowali hipotezy tak jak to określa wyrażenie (2.10), to istotnie małe wartości sprawdzianu świadczą przeciwko hipotezie H_0. Gdy zaobserwowana w próbie częstość p_s jest nie większa niż wartość krytyczna $p_{*\kappa}$, czyli $p_s \leq p_{*\kappa}$, co jest równoważne nierówności $z_s \leq z_\kappa$, to odrzucamy H_0 na korzyść H_1 z prawdopodobieństwem popełnienia błędnej decyzji (ryzykiem) κ. Oznacza to, iż z prawdopodobieństwem pomyłki κ stwierdzamy, że system kontroli wewnętrznej jest skuteczny.

Gdy $p_s > p_{*\kappa}$, co jest równoważne nierówności $z_s > z_\kappa$, to przyjmujemy H_0 z prawdopodobieństwem popełnienia błędnej decyzji (ryzykiem) η. Oznacza to, iż system kontroli wewnętrznej jest traktowany jako nieskuteczny z prawdopodobieństwem pomyłki η.

2.2.2. Weryfikacja hipotezy, w sytuacji gdy sprawdzian testu ma rozkład dokładny

Hipotezę H_0 określoną wyrażeniem (2.2) weryfikuje się za pomocą obserwowanej w próbie liczby wykrytych nieprawidłowości, oznaczanej dalej

przez M_s. Próba prosta o liczebności n jest losowana bezzwrotnie. Wiadomo, że statystyka M_s ma rozkład hipergeometryczny, którego funkcję prawdopodobieństwa w przypadku prawdziwości hipotezy H_0 określa wzór:

$$P(M_s = m) = \frac{\binom{M_0}{m}\binom{N-M_0}{n-m}}{\binom{N}{n}}. \tag{2.14}$$

Duże wartości statystyki testowej M_s świadczą przeciwko hipotezie sprawdzanej H_0, a na na korzyść hipotezy H_1. Poziom istotności testu, który nie może być większy od ryzyka dezaprobaty skutecznego systemu kontroli η, określa wyrażenie:

$$\eta \geq \alpha = P(M_s \geq m_\eta | H_0), \tag{2.15}$$

gdzie m_η jest wartością krytyczną testu. Podkreślmy, że liczba m_η jest najmniejszą liczbą naturalną, spełniającą powyższą nierówność.

Gdy $m_s \geq m_\eta$, to system kontroli wewnętrznej nie jest akceptowany jako skuteczny z ryzykiem nie większym niż η popełnienia błędnej decyzji. Wynika to stąd, że spełnienie nierówności $m_s \geq m_\eta$ prowadzi do odrzucenia hipotezy H_0 na korzyść H_1 z prawdopodobieństwem popełnienia błędu I rodzaju nieprzekraczającym poziomu istotności $\alpha = \eta$.

Gdy $m_s < m_\eta$, to nie ma podstaw do odrzucenia sądu o skuteczności systemu kontroli wewnętrznej, ponieważ nie ma podstaw do odrzucenia hipotezy H_0. W tym przypadku hipotezę H_0 przyjmujemy, jeśli prawdopodobieństwo ν popełnienia błędu II rodzaju jest dostecznie małe i nie większe od ryzyka akceptacji nieskutecznego systemu kontroli. W rozważanym przypadku prawdopodobieństwo ν określa wyrażenie:

$$\kappa \geq \nu = P(M_s \leq m_\eta | H_1). \tag{2.16}$$

Oczywiście, żąda się, aby ν było nie większe od założonego ryzyka κ akceptacji nieskutecznego systemu kontroli, czyli aby $\kappa \geq \nu$. Przy ustalonej liczebności próby ta nierówność nie musi być spełniona, dlatego pozostaje wyznaczyć minimalny rozmiar próby, przy którym nierówność (2.16) będzie spełniona. Szukany rozmiar próby, dalej oznaczany przez $n_{\eta\kappa}$ można w paktyce wyznaczyć za pomocą iteracyjnej procedury, rozwiązania następującego układu nierówności:

$$\begin{cases} \eta \geq P(M_s \geq m | H_0) = 1 - \sum_{k=0}^{m-1} \frac{\binom{M_0}{k}\binom{N-M_0}{n-k}}{\binom{N}{n}}, \\ \\ \kappa \geq P(M_s < m | H_1) = \sum_{k=0}^{m-1} \frac{\binom{M_1}{k}\binom{N-M_1}{n-k}}{\binom{N}{n}}, \end{cases} \tag{2.17}$$

ze względu na niewiadome m i n. Rozwiązanie $m_{\eta,\kappa}$ i $n_{\eta\kappa}$ tego układu nierówności otrzymujemy iteracyjnie; rozpoczynając od liczebności próby $n = 1$ wyznacza się wartość krytyczną $m_{\eta,\kappa}$ jako największą liczbę naturalną

spełniającą pierwszą nierówność układu. Następnie, jeśli druga nierówność jest spełniona dla $m_{\eta,\kappa}$, to $n_{\eta\kappa} = 1$, a w przeciwnym razie powiększa się o jeden liczebność próby do poziomu $n = 2$. Procedurę powtarza się aż do uzyskania takiej liczebności $n_{\eta\kappa}$ oraz wartości krytycznej $m_{\eta,\kappa}$, dla których układ nierówności jest spełniony.

Przykład 2.2

W urnie jest 200 pytań, które są zadawane studentowi podczas egzaminu magisterskiego. Zwykle, dobrze przygotowany do egzaminu student potrafi odpowiedzieć na 80% pytań. Z kolei słabo przygotowany student zna prawidłowe odpowiedzi na 50% pytań. Zatem badana populacja składa się z $N = 200$ pytań. Dopuszczalna liczba pytań, na które student może nie znać odpowiedzi wynosi $M_0 = 60$, natomiast już dyskwalifikującą liczbą pytań (na które student nie potrafi odpowiedzieć) jest poziom $M_1 = 100$. Student odpowiada na n wylosowanych bezzwrotnie pytań. Osoba prowadząca egzamin gra rolę swoistego audytora poziomu wiedzy studenta. Na podstawie jego poprawnych odpowiedzi może uznać, że jego wiedza jest na dobrym albo niedostatecznym poziomie. Innymi słowy, niewiedza studenta może być na dopuszczalnym poziomie, gdy nie potrafi odpowiedzieć na nie więcej niż M_0 pytań, albo na poziomie dyskwalifikującym, jeśli nie potrafi odpowiedzieć na co najmniej M_1 pytań. Egzaminator na podstawie odpowiedzi studenta może uznać jego wiedzę za niewystarczającą (czyli nie zalicza egzaminu), gdy student był dobrze przygotowany z ryzykiem η. Z drugiej strony może ten egzamin zaliczyć, mimo że student reprezentuje dyskwalifikujące go umiejętności odpowiedzi z ryzykiem κ. Egzaminator postuluje, aby te ryzyka były równe $\eta = 0.1$ i $\kappa = 0.2$. Mamy więc problem dotyczący weryfikacji hipotez zdefiniowanych wyrażeniem (2.2) z określonymi właśnie dopuszczalnymi prawdopodobieństwami popełnienia błędów I i II rodzaju. Problem polega na ustaleniu liczby pytań, które powinien wylosować student podczas egzaminu, aby te postulaty były spełnione. Zatem należy rozwiązać układ nierówności określony wyrażeniem (2.17). W tym celu korzystamy z następującej procedury napisanej w języku R programowania komputerów.

#Wyznaczanie liczebności próby przy ustalonym η i κ
#Liczebność populacji:
$N < -200$
#Hipoteza H0:
$M0 < -60$
#Hipoteza H1:
$M1 < -100$
#Ryzyka:
$eta < -0.1;\ \ kappa < -0.2$
$pk < -1;\ \ n < -1$

while $((pk > kappa)\&(n < N - M0))$
$\{wk < -qhyper(1 - eta, M0, N - M0, n)$
$pk < -phyper(wk, M1, N - M1, n)$
$n < -n + 1\}$
#niezbędna liczebność próby:
$n < -n - 1;\ n$
#wartość krytyczna testu:
wk

Wynikiem realizacji tej aplikacji są wartość krytyczna $m_{\eta,\kappa} = 10$ oraz liczebność $n_{\eta\kappa} = 25$. Zatem, przy określonych postulatach student powinien odpowiadać na 25 wylosowanych bezzwrotnie pytań. Jeśli nie odpowie na co najmniej 10 spośród nich, to egzaminator nie daje zaliczenia studentowi z prawdopodobieństwem błędnej decyzji (polegającej na niedocenieniu wiedzy studenta) nie większym od 0.1. W sytuacji gdy student nie odpowiada poprawnie na mniej niż 10 pytań, to egzaminator daje mu zaliczenie, jakkolwiek jest świadom, że mógł się pomylić (czyli przecenić umiejętności studenta) z prawdopodobieństwem nie większym od 0.2.

Odpowiadanie na egzaminie ustnym na 25 pytań jest równoważne ,,zamęczeniu" studenta, dlatego egzaminator wspaniałomyślnie zwiększa ryzyko zaliczenia egzaminu studentowi, nawet gdy na to nie zasługuje, do poziomu $\kappa = 0.33$. Wtedy $m_{\eta,\kappa} = 7$ oraz liczebność $n_{\eta\kappa} = 17$. W tej sytuacji egzaminator już nie decyduje się na dalsze zwiększanie prawdopodobieństwa popełnienia błędu κ, ponieważ obniżyłoby to wiarygodność egzaminu, bo zwiększyłoby się ryzyko zdania egzaminu przez niedostatecznie przygotowanych studentów.

2.2.3. Przybliżanie rozkładu prawdopodobieństwa sprawdzianu testu rozkładami dwumianowym albo Poissona

W literaturze znane są twierdzenia graniczne (por. np. Gerstenkorn, Śródka, 1974 lub Krzyśko, 2000), pozwalające aproksymować rozkład hipereometryczny w sytuacji, gdy liczebność próby jest bardzo duża lub jeśli prawdopodobieństwo $p_U = \frac{M_u}{N}$ jest bliskie zeru. Z jednego z nich wynika, że funkcję prawdopodobieństwa rozkładu hipergeometrycznego można przybliżyć funkcją prawdopodobieństwa rozkładu dwumianowego, gdy liczebność populacji jest duża. W praktyce to przybliżenie zaleca się, gdy $N \geq 1000$ i $p_u \geq 0.04$. Wówczas wyrażenie (2.17) prowadzące do wyznaczenia wartości krytycznej $m_{\eta,\kappa}$ i rozmiaru $n_{\eta\kappa}$ próby spełniających postulaty nieprzekroczenia ryzyka κ błędnej akceptacji systemu kontroli i ryzyka η błędnej jego dezaprobaty przyjmuje postać:

$$\begin{cases} \eta \geq P(M_s \geq m|H_0) = 1 - \sum_{k=0}^{m-1} \binom{n}{k} p_0^k (1-p_0)^{n-k}, \\ \\ \kappa \geq P(M_s < m|H_1) = \sum_{k=0}^{m-1} \binom{n}{k} p_1^k (1-p_1)^{n-k}. \end{cases} \tag{2.18}$$

Kolejne przybliżenie rozkładu hipergeometrycznego rozkładem prawdopodobieństwa Poissona stosuje się, gdy liczebność populacji jest bardzo duża i wraz ze wzrostem liczebności próby n wartość prawdopodobieństwa p_U maleje w taki sposób, że np_U zmierza do granicznej wartości λ. W praktyce zwykle stosuje się to przybliżenie, jeśli $n \geq 1000$, natomiast $p_U < 0.04$. Wtedy wyrażenie (2.17) przyjmuje postać:

$$\begin{cases} \eta \geq P(M_s \geq m|H_0) = 1 - \sum_{k=0}^{m-1} \frac{(np_0)^k}{k!} e^{-np_0}, \\ \\ \kappa \geq P(M_s < m|H_1) = \sum_{k=0}^{m-1} \frac{(np_0)^k}{k!} e^{-np_0}. \end{cases} \tag{2.19}$$

Prawdopodobieństwa rozkładu hipergeometrycznego, dwumianowego lub Poissona można wyliczyć za pomocą odpowiednich instrukcji języka pakietu statystycznego R, które dalej będziemy wykorzystywać. Warto tutaj zwrócić uwagę, że dla dużych rozmiarów populacji zarówno rozkład hipergeometryczny, jak i dwumianowy jest przybliżany rozkładem beta, co wynika z dostępnego online elektronicznego podręcznika pakietu statystycznego R.

Przykład 2.3

System komputerowy przeprowadzania egzaminów składa się m.in. z bazy pytań. Odpowiedź na każde z nich może być tylko twierdząca albo przecząca. Baza obejmuje $N = 2000$ pytań, które dotyczą treści wyczerpujących pewną dziedzinę wiedzy. Uważa się, że student posiada wystarczający zasób wiedzy o przedmiocie, jeśli potrafi odpowiedzieć na 80% pytań, a gdy ta umiejętność dotyczy tylko 60% pytań, jego wiedza o przedmiocie nie jest wystarczająca. Innymi słowy, dopuszcza się, by student nie potrafił udzielić prawidłowych odpowiedzi na 20% pytań. W przypadku gdy nie zna odpowiedzi prawidłowych na 40% pytań, taki poziom wiedzy uważa się za niedopuszczalny. Można więc sformułować hipotezy określone równoważnie wyrażeniami (2.1) lub (2.2), gdzie przez $p_0 = 0.2$, $p_1 = 0.4$ lub $M_0 = Np_0 = 400$, $M_1 = Np_1 = 800$. Aby sprawdzić wiedzę studenta, losuje się n pytań. Liczbę pytań, na które udzielił nieprawidłowej odpowiedzi oznaczymy przez m. Wartość krytyczną testu oznaczamy przez $m_{\eta,\kappa}$. Zatem, jeśli student nie odpowie na co najmniej $m_{\eta,\kappa}$ pytań, czyli $m \geq m_{\eta,\kappa}$, to nie zalicza egzaminu z ryzykiem podjęcia krzywdzącej go decyzji równym lub mniejszym η. Innymi słowy, w tym przypadku może zdarzyć się, że student akurat trafił w próbie na te pytania, na które nie znał prawidłowej odpowiedzi. Ta sytuacja występuje jednak z małym prawdopodobieństwem

równym właśnie η. W sytuacji gdy $m < m_{\eta,\kappa}$, zaliczamy studentowi egzamin, mimo że taka decyzja może okazać się nietrafna z ryzykiem (prawdopodobieństwem) nie większym od κ. Student może nie znać prawidłowych odpowiedzi na część pytań, lecz ma szczęście i w wylosowanej próbie trafia na pytania, na które właściwe odpowiedzi są mu znane. Tak może zdarzyć się właśnie z prawdopodobieństwem κ. Zatem objaśniane obydwa ryzyka powinny być bliskie zeru. To jednak prowadzi do zwiększenia liczebności próby, co wydłuża egzamin i stresuje studenta. W związku z tym szuka się poziomów kompromisowych tych rodzajów ryzyka i co więcej, działa się na korzyść studenta, przyjmując, że $\eta < \kappa$. W naszym przypadku przyjmijmy, że $\kappa = 0.1$, natomiast $\eta = 0.05$. Wówczas potrzebną do spełnienia tych postulatów liczebność próby wyznaczamy na podstawie układu nierówności określonego wyrażeniem (2.18), co uzasadniamy faktem, że prawdopodobieństwa p_0 i p_1 przyjmują stosunkowo duże wartości, a liczebność populacji jest dość duża. Potrzebne obliczenia przeprowadzamy na podstawie następującej procedury przygotowanej języku R.

#Wyznaczanie liczebności próby przy ustalonym η i κ
#Przybliżenie dwumianowe
#Liczebność populacji:
$N < -2000$
#Hipoteza H0:
$p0 < -0.2$
#Hipoteza H1:
$p1 < -0.4$
#Ryzyka:
$eta < -0.05;\ kappa < -0.1$
$pk < -1;\ n < -1$
while $((pk > kappa)\&(n < N))$
$\{wk < -qbinom(1 - eta, n, p0)$
$pk < -pbinom(wk, n, p1)$
$n < -n + 1\}$
#niezbędna liczebność próby:
$n < -n - 1;\ n$
#wartość krytyczna testu:
wk

W wyniku przeprowadzonych obliczeń otrzymujemy, że $m_{\eta,\kappa} = 14$ i $n_{\eta,\kappa} = 47$. Student losuje więc 47 pytań. Jeśli student nie odpowie na co najmniej 14 pytań, to nie zaliczamy mu egzaminu z ryzykiem popełnienia niesprawiedliwej decyzji mniejszym lub równym 0.05. Gdy student nie odpowie na mniej niż 14 pytań, to zaliczamy mu egzamin z ryzykiem popełnienia błędnej decyzji nie większym od 0.1.

Przykład 2.4

Przeprowazono audyt 19 800 operacji bankowych celem oceny systemu kontroli wewnętrznej, który powinien uniemożliwić występowanie ewentualnych błędów w tych operacjach. Dopuszczalny udział mogących wystąpić nieprawidłowości ustala się na poziomie $p_0 = 0.005$, a poziom dyskwalifikujący wynosi $p_1 = 0.02$. Ryzyko aprobaty nieskutecznego systemu kontroli ustalono na poziomie $\kappa = 0.05$, a ryzyko dezaprobaty skutecznego systemu kontroli na poziomie $\eta = 0.1$. Zaznaczmy, że w tym przypadku w przeciwieństwie do problemu egzaminowania rozważanego w przykładzie 2.3 mamy, iż $\kappa < \eta$, czyli żąda się, aby ryzyko akceptacji nieskutecznego systemu kontroli było mniejsze od ryzyka dezaprobaty prawidłowo działającego systemu kontroli.

W związku z tym, że mamy do czynienia z dużą populacją i małym prawdopodobieństwem wystąpienia błędu, które można traktować jako tzw. zjawisko rzadkie. W tej sytuacji do wyznaczenia wartości krytycznej testu i rozmiaru próby spełniających postulaty ograniczające poziom ryzyk, użyjemy układu nierówności określonej wyrażeniem (2.19). Aby tę nierówność rozwiązać, wykorzystujemy następującą iteracyjną aplikację napisaną w języku R.

#Wyznaczanie liczebności próby przy ustalonym η i κ
#Przybliżenie Poissona
#Liczebność populacji:
$N < -19800$
#Hipoteza H0:
$p0 < -0.005$
#Hipoteza H1:
$p1 < -0.02$
#Ryzyka:
$eta < -0.1;\ kappa < -0.05$
$pk < -1;\ n < -1$
while $((pk > kappa)\&(n < N))$
$\{wk < -qpois(1 - eta, n * p0)$
$pk < -ppois(wk, n * p1)$
$n < -n + 1\}$
#niezbędna liczebność próby:
$n < -n - 1;\ n$
#wartość krytyczna testu:
wk

Ta procedura wylicza, iż wartość krytyczna testu wynosi $m_{\eta,\kappa} = 4$, niezbędny rozmiar próby $n \geq 458$. Stąd więc wynika, że jeśli w 458 elementowej próbie prostej losowanej zwrotnie zostanie stwierdzonych mniej niż 4 wadliwie przeprowadzone operacje bankowe, to akceptujemy system

kontroli wewnętrznej, przy czym będzie to błędna decyzja z prawdopodobieństwem nie większym od 0.05. W sytuacji gdy ilość nieprawidłowych operacji osiągnie poziom krytyczny 4, bądź będzie od niego większa, to dezaprobujemy system kontroli wewnętrznej, z ryzykiem popełnienia błędnej decyzji nieprzekraczającym poziomu 0.1.

2.3. Wnioskowanie przy ustalonym ryzyku aprobaty nieskutecznie działającego systemu kontroli

2.3.1. Przypadek próby prostej

Rozważymy teraz sytuację, gdy tzw. dyskwalifikująca system kontroli wewnętrznej wartość proporcji nieprawidłowości nie jest precyzyjnie określona. Wyznaczona jest jedynie wartość dopuszczalna tej proporcji oznaczana przez p_0. Wówczas wiemy, iż $p_1 > p_0$ i wartość parametru p_1 może być dowolnie bliska wartości dopuszczalnej p_0. W tej sytuacji hipotezy formułowane w podrozdziale 2.1 wyrażeniem (2.1) upraszczają się do postaci

$$H_0 : \ p_U > p_0, \qquad H_1 : \ p_U \leq p_0. \tag{2.20}$$

Zatem H_0 głosi, że system kontroli jest nieskuteczny, bo proporcja nieprawidłowości jest większa od dopuszczalnej p_0, czyli równa najmniejszej proporcji dyskwalifikujacej $p_1 = p_0 + \delta$, gdzie $\delta > 0$, lecz δ nie jest ustalone. Z kolei H_1 głosi, że system kontroli jest skuteczny, bo nie przekracza dopuszczalnej wartości p_0. Ryzyko κ aprobaty nieskutecznego systemu kontroli jest teraz poziomem istotności testu statystycznego.

Najpierw rozważymy raczej akademicki przypadek wnioskowania o skuteczności systemu kontroli wewnętrznej na podstawie danych obserwowanych w próbie prostej losowanej bezzwrotnie. Z racji tego, że proporcja nieprawidłowości $p_U = \frac{M_U}{N}$, wyżej określone hipotezy można zastąpić następującymi równoważnymi:

$$H_0 : \ M_U \geq M_0 + 1 > M_0, \qquad H_1 : \ M_U \leq M_0, \tag{2.21}$$

gdzie przez M_0 oznaczono ilość elementów nieprawidłowych w populacji o rozmiarze N. Hipoteza sprawdzana H_0 głosi, że liczba nieprawidłowości przekracza dopuszczalny poziom M_0, a hipoteza alternatywna H_1 głosi, że liczba nieprawidłowości nie przekracza dopuszczalnego poziomu M_0.

Do weryfikacji hipotezy H_0 użyjemy liczby wykrytych nieprawidłowości, oznaczanej dalej przez M_s, obserwowanych w próbie prostej losowanej bezzwrotnie o liczebności n. Wiadomo, że statystyka M_s ma rozkład hipergeometryczny, którego funkcję prawdopodobieństwa w przypadku prawdziwości hipotezy H_0 określa wzór (2.14). Małe wartości statystyki testowej

M_s świadczą przeciwko hipotezie sprawdzanej H_0, a na na korzyść hipotezy H_1. Poziom istotności testu, będący jednocześnie ryzykiem aprobaty nieskutecznego systemu kontroli κ, określa wyrażenie:

$$\alpha = \kappa \geq P(M_s \leq m_\kappa | H_0), \tag{2.22}$$

gdzie m_κ jest wartością krytyczną testu. Liczba m_κ jest największą liczbą naturalną spełniającą nierówność (2.22).

Gdy $m_s \leq m_\kappa$, to akceptujemy skuteczność systemu kontroli wewnętrznej z ryzykiem κ popełnienia błędnej decyzji. Wynika to stąd, że spełnienie nierówności $m_s \leq m_\kappa$ prowadzi do odrzucenia hipotezy H_0 na korzyść H_1 z prawdopodobieństwem popełnienia błędu I rodzaju równemu poziomowi istotności $\alpha = \kappa$.

Gdy $m_s > m_\kappa$, to nie ma podstaw do akceptacji skuteczności systemu kontroli wewnętrznej, ponieważ nie ma podstaw do odrzucenia hipotezy H_0.

Wyznaczanie wartości krytycznej testu przy ustalonej liczebności próby

W sytuacji gdy liczebność próby jest ustalona i równa n oraz ryzyko κ jest również określone, to wartość krytyczną m_κ wyznaczamy jako rozwiązanie nierówności:

$$\kappa \geq P(M_s \leq m_\kappa) = \sum_{k=0}^{m_\kappa} \frac{\binom{M_0}{k}\binom{N-M_0}{n-k}}{\binom{N}{n}}, \tag{2.23}$$

przy czym m_κ jest maksymalną liczbą spełniającą powyższe wyrażenie.

Przykład 2.5

Populacja dokumentów składa się z $N = 300$ sztuk. Audytor ustalił, że dopuszczalna liczba błędnych dokumentów może wynosić $M_0 = 25$. Ponadto, stwierdził, że ryzyko aprobaty nieskutecznego systemu kontroli $\kappa = 0.1$. W celu sprawdzenia funkcjonowania systemu kontroli wewnętrznej jakości tworzonych dokumentów wylosowano z analizowanej ich zbiorowości $n = 45$ dokumentów. W wyniku ich kontroli stwierdzono, że w dwóch spośród nich wykryto nieprawidłowości.

Z formalnego punktu widzenia mamy do czynienia z weryfikacją hipotezy H_0 określonej wyrażeniem (2.21). W celu wyznaczenia wartości krytycznej m_κ testu, spełniającej nierówność określoną wzorem (2.23), korzystamy z następującej instrukcji napisanej w języku komputerowym pakietu R:

$$qhyper(\kappa, M_0, N - M_0, n) - 1.$$

Realizacja instrukcji *hyper(0.1,25,475,45)-1* jest właśnie szukaną wartością krytyczną, która w naszym przypadku wynosi $m_\kappa = 1$. Wyliczona na pod-

stawie próby wartość sprawdzianu testu wynosi $m_s = 2$. Zatem, z nierówności $m_s = 2 > 1 = m_\kappa$ wynika, że nie ma podstaw do odrzucenia hipotezy H_0 głoszącej, iż liczba nieprawidłowości w badanym zbiorze dokumentów jest większa od dopuszczalnej. To również oznacza, że nie ma podstaw do akceptacji jakości funkcjonowania systemu kontroli.

Rozwiązanie nierówności (2.23) uprości się, gdy liczebność populacji N będzie duża (rzędu co najmniej 1000). Wówczas z teorii rachunku prawdopodobieństwa wynika, że nierówność (2.23) można zastąpić następującą:

$$\kappa \geq \sum_{k=0}^{m_\kappa} \binom{n}{k} p_0^k (1-p_0)^{n-k}, \qquad p_0 = \frac{M_0}{N}. \tag{2.24}$$

Przykład 2.6

Zbiór dokumentów badanych przez audytora liczy 8500 faktur. Dopuszczalna liczba faktur z błędami wynosi 5% ich ilości. Z badanej zbiorowości wylosowano bezzwrotnie 80 faktur. Okazało się, że w wyniku przeprowadzonej kontroli wśród tych faktur znalazły się dwie faktury z błędami. Czy z ryzykiem błędnej decyzji równym 0.09 można twierdzić, że system kontroli wewnętrznej jest skuteczny? Mamy więc do czynienia z weryfikacją określonej wzorem (2.21) hipotezy sprawdzanej dla parametru $M_0 = p_0 N = 425$, przy czym $p_0 = 0.05$, $N = 8500$, $n = 80$, $\kappa = 0.09$, $m_s = 1$. Wartość krytyczną testu wyznaczamy na podstawie wzoru (2.24). W tym celu używamy następującej instrukcji pakietu (systemu) komputerowego R:

$$qbinom(\kappa, n, p_0) - 1$$

W naszym przypadku *qbinom(0.09,80,0.05)-1=1*, czyli $m_\kappa = 1$. Zatem $m_s = 1 = m_\kappa$, co prowadzi do wniosku, że hipotezę H_0 należy odrzucić z prawdopodobieństwem błędnej decyzji równym 0.09, co w kontekście badania audytowego oznacza, iż akceptujemy system kontroli wewnętrznej z ryzykiem (popełnienia błędnej decyzji) równym 0.09.

Dodajmy jeszcze, że jeśli wartość krytyczną wyznaczalibyśmy na podstawie dokładnego rozkładu hipergeometrycznego, to wówczas określona w przykładzie 2.5 instrukcja programu R: *qhyper(0.09,425,8075,80)-1* prowadzi również do wartości krytycznej testu $m_\kappa = 1$.

Następny przybliżony sposób rozwiązywania nierówności (2.23) wynika ze znanego twierdzenia granicznego, pozwalającego uzasadnić przybliżenie rozkładu dwumianowego rozkładem Poissona (por. np. Gerstenkorn, Śródka, 1974; Jakubowski i Sztencel, 2004 i Krzyśko, 2000). Praktyczny wniosek wynikający z tego twierdzenia pozwala, przy założeniu, że $p < 0.05$, $N \geq 1000$ i $\lambda = np$, nierówność (2.23) zastąpić przez następującą:

$$\kappa \geq e^{-np_0} \sum_{k=0}^{m_\kappa} \frac{(np_0)^k}{k!}. \tag{2.25}$$

Przykład 2.7

Spośród 30 000 dokumentów wylosowano bezzwrotnie próbę prostą o rozmiarze 900. W wyniku przeprowadzonej kontroli zaobserwowano wśród nich 8 dokumentów z występującymi w nich błędami. Dopuszcza się, że jedynie 1% liczby tych dokumetów może być nieprawidłowo sporządzonych. Czy z prawdopodobieństwem 0.04 można zaryzykować, iż w świetle tego rezultatu system kontroli można zaakceptować jako skuteczny? Reasumując, $N = 30000$, $n = 900$, $p_0 = 0.01$, $\kappa = 0.04$ oraz $m_s = 8$. Wartość krytyczną testu hipotezy H_0, danej wyrażeniem (2.21) wyznaczamy na podstawie wzoru (2.25), korzystając z następującej instrukcji programu R:

$$qpois(\kappa, np_0) - 1.$$

W naszym przypadku $m_\kappa = qpois(0.04,900*0.01)\text{-}1=3$. Zatem nierówność $m_s = 8 > m_\kappa = 3$ prowadzi audytora do wniosku, że nie należy odrzucać hipotezy o tym, że badany system kontroli wewnętrznej jest nieskuteczny.

Znane centralne twierdzenie graniczne de Moivre'a-Laplace'a pozwala na przybliżone rozwiązanie nierówności (2.23) dla $N \geq 1000$, $p \geq 0.05$, $n \geq 100$ wyznaczyć jako rozwiązanie następującego równania:

$$\kappa = P\left(Z < \frac{p_\kappa - p_0}{\sqrt{\frac{(N-n)}{(N-1)n} p_0(1-p_0)}} = z_\kappa\right), \tag{2.26}$$

z którego wyliczamy, że $m_\kappa = np_\kappa$, gdzie:

$$p_\kappa = p_0 + z_\kappa \sqrt{\frac{N-n}{N-1} \frac{p_0(1-p_0)}{n}}, \tag{2.27}$$

przy czym $\kappa = P(Z < z_\kappa)$ i $Z \sim N(0;1)$.

Przykład 2.8

Spośród 6500 sprawozdań finansowych wylosowano bezzwrotnie próbkę prostą 120-elementową. Po kontroli wylosowanych sprawozdań okazało się, że 8 spośród nich jest błędnie przygotowanych. Dopuszcza się, że aż 0.06% źle przygotowanych sprawozdań może być niezauważonych podczas kontroli wewnętrznej. Na poziomie ryzyka błędnej akceptacji systemu kontroli wewnętrznej równym 0.08 sprawdzamy, czy można uznać, że system kontroli jest wystarczająco wydolny, co sprowadza się do weryfikacji hipotezy określonej wzorem (2.21). Z racji tego, że wyjściowe warunki spełniają kryteria, przy których stosuje się centralne twierdzenie de Moivre'a–Laplace'a, wykorzystujemy wzory (2.26) i (2.27) do wyliczenia wartości krytycznej testu. Korzystając z intrukcji programu R:

$$qnorm\left(\kappa, np_0, \sqrt{\frac{N-n}{(N-1)} np_0(1-p_0)}\right).$$

W naszym przypadku $N = 6500$, $n = 120$, $m_s = 8$, $p_0 = 0.06$, $\kappa = 0.08$. To już pozwala wyliczyć wartość *qnorm(0.08,7.2,2.5776)=3.5783* $= m_\kappa$. Zatem $m_s = 8 > m_\kappa$. W związku z tym nie ma podstaw do aprobaty skuteczności badanego systemu kontroli wewnętrznej.

Wyznaczanie liczebności próby prostej losowanej bezzwrotnie przy ustalonej wartości krytycznej

W przypadku gdy audytor ustala poziom ryzyka κ oraz wartość krytyczną testu m_κ, to pojawia się możliwość wyznaczania niezbędnej liczebności próby, zapewniającej podjęcie decyzji o prawdziwości hipotezy H_0. Czynimy to również na podstwie nierówności (2.23)-(2.25), lecz tym razem rozwiązujemy je względem niewiadomej, którą jest właśnie szukana liczebność próby. To doprowadziło do stworzenia specjalnych tablic (por. np. Karliński, 2005; Hołda i Pociecha, 2004; Guy, Carmichael i Whittinton, 2002) pozwalających przy ustalonym p_0, κ i p_κ wyznaczyć niezbędną liczebność próby.

Przykład 2.9

Populacja składa się z $N = 400$ dokumentów. Dopuszcza się, że $M_0 = 20$ spośród nich może być nieprawidłowo sporządzonych. Audytor w celu przekonania się, czy ta dopuszczalna ilość nieprawidłowości jest zachowana, weryfikuje hipotezę określoną wyrażeniem (2.21), ustalając ryzyko akceptacji nieskutecznego systemu kontroli wewnętrznej na poziomie $\kappa = 0.1$. W celu weryfikacji postawionej hipotezy sprawdzanej będzie wylosowana bezzwrotnie próba prosta. Ponadto, audytor ustalił, że hipotezę H_0 odrzuca z ryzykiem κ, gdy zaobserwowana liczba dokumentów z błędami nie przekroczy krytycznej liczby $m_\kappa = 2$. Pozostaje więc wyznaczenie niezbędnej liczebności próby spełniającej określone postulaty, co czynimy na podstawie wzoru (2.23), wspomagając się następującą procedurą napisaną w języku pakietu R.

#Liczebność populacji:
$N < -400$
#Hipoteza H0:
$M0 < -20$
#liczebność krytyczna:
$mk < -2$
#ryzyko:
$kappa < -0.1$
$pk < -1$
$n < -mk$
$while\ ((pk > kappa)\&(n < N - M0))$
$\{pk < -phyper(mk, M0, N - M0, n)$
$n < -n + 1\}$

#niezbędna liczebność próby:
n-1

W wyniku realizacji zapisanej aplikacji otrzymujemy, że liczebność próby $n \geq 97$ wystarczy po to, aby test z ryzykiem nie większym od $\kappa = 0.1$ zaakceptuje nieskuteczny system kontroli, gdy liczba nieprawidłowych dokumentów będzie nie większa od $m_\kappa = 2$.

Przykład 2.10

Ze zbiorowości rolników pewnej gminy liczącej 2500 osób będzie losowana bezzwrotnie próba prosta w celu weryfikacji określonej wyrażeniem (2.21) hipotezy sprawdzanej, głoszącej, że co najmniej 0.2 wniosków o dofinansowanie działalności na rzecz ochrony środowiska jest nieprawidłowo sporządzonych i zostało to niezauważone przez przyjmujących te dokumenty. Ta hipoteza będzie odrzucana, jeśli liczba stwierdzonych w próbie błędnych sprawozdań nie przykroczy poziomu 5. Zatem oznacza to, że system kontroli jest nieskuteczny. Ryzyko przyjęcia nieskutecznego systemu kontroli ustalono na poziomie 0.1. Teraz pozostaje wyznaczyć niezbędną liczebność próby, spełniającą określone postulaty. Reasumując, rozmiar próby powinien zapewniać, że dla wartości krytycznej testu $m_\kappa = 5$ weryfikujemy hipotezę $H_0 : M_U > p_0 N = 500$ z ryzykiem jej błędnego odrzucenia równym $\kappa = 0.1$, przy czym $p_0 = 0.2$ i $N = 2500$. W związku z tym, że rozmiar populacji dokumentów jest duży (co zapewnia zbieżność rozkładu hipergeometrycznego do rozkładu dwumianowego), wykorzystujemy nierówność (2.24) do wyznaczenia potrzebnej liczebności próby, która będzie rozwiązaniem tej nierówności. Do wyliczenia tej liczebności doprowadzi nas następująca rekurencyjna procedura napisana w języku R.

#wyznaczanie liczebności próby przy ustalonym kappa
#przybliżenie dwumianowe
#liczebność populacji:
$N < -2500$
#Hipoteza
$H0 : p0 < -0.2$
$M0 < -N * p0$
#wartość krytyczna:
$mk < -5$
#ryzyko:
$kappa < -0.1$
$pk < -1$
$n < -mk$
$while((pk > kappa)\&(n < N - M0))$
$\{pk < -pbinom(mk, n, p0)$
$n < -n + 1\}$

#niezbędna liczebność próby:
n-1

W wyniku realizacji zapisanej aplikacji otrznymujemy szukany rozmiar próby: $n \geq 45$.

Przykład 2.11

Ze zbiorowości podatników pewnej gminy liczącej 9500 osób dorosłych będzie losowana bezzwrotnie próba prosta w celu weryfikacji określonej wyrażeniem (2.21) hipotezy sprawdzanej, głoszącej, że co najmniej 0.01 sprawozdań podatkowych jest nieprawidłowo sporządzona i zostało to niezauważone przez przyjmujących te dokumenty. Ta hipoteza będzie odrzucana, jeśli liczba stwierdzonych w próbie błędnych sprawozdań nie przykroczy poziomu 5. Ryzyko przyjęcia nieskutecznego systemu kontroli ustalono na poziomie 0.1. Teraz pozostaje wyznaczyć niezbędną liczebność próby spełniającą określone postulaty. Reasumując, rozmiar próby powinien zapewniać, że dla wartości krytycznej testu $m_\kappa = 5$ weryfikujemy hipotezę $H_0 : M_U > p_0 N = 95$ z ryzykiem jej błędnego odrzucenia równym $\kappa = 0.1$, przy czym $p_0 = 0.01$ i $N = 9500$. W związku z tym, że rozmiar populacji sprawozdań jest duży, parametr p_0 mały (co umożliwia zbieżność rozkładu hipergeometrycznego do rozkładu Poissona), wykorzystujemy nierówność (2.25) do wyznaczenia potrzebnej liczebności próby. Zatem, będzie ona rozwiązaniem tej nierówności, do czego doprowadzi nas następująca rekurencyjna aplikacja napisana w języku R.

#wyznaczanie liczebności próby przy ustalonym kappa
#przybliżenie Poissona
#liczebność populacji:
$N < -9500$
#Hipoteza
$H0 : p0 < -0.01$
$M0 < -N * p0$
#wartość krytyczna:
$mk < -5$
#ryzyko:
$kappa < -0.1$
$pk < -1$
$n < -mk$
$while((pk > kappa)\&(n < N - M0))$
$\{pk < -ppois(mk, n * p0)$
$n < -n + 1\}$
#niezbędna liczebność próby:
n-1

W wyniku przeprowadzonych obliczeń otrzymujemy, że należy wylosować co najmniej 928 sprawozdań do ich skontrolowania.

Ze wspomnianego już twierdzenia de Moivre'a–Laplace'a przy założeniu, że $N \geq 1000$, $p \geq 0.05$ oraz ustalonej częstości krytycznej p_κ z równania (2.27) wynika, iż

$$n \geq \frac{1}{\frac{1}{N} + \frac{N-1}{N}\frac{(p_\kappa - p_0)^2}{z_\kappa^2 p_0(1-p_0)}} \approx \frac{z_\kappa^2 p_0(1-p_0)}{(p_\kappa - p_0)^2}, \tag{2.28}$$

przy czym $\kappa = P(Z < z_\kappa)$, $Z \sim N(0;1)$, a wartość krytyczną $p_\kappa = \frac{m_\kappa}{n}$ audytorzy zwykle określają na poziomie nie większym niż $0.5p_0$, czyli $p_\kappa = 0.5p_0$ (por. np. Karliński, 2005).

Przykład 2.12

Zwykle z powodu nieuwagi, czyli braku samokontroli, 15% nauczycieli oddaje protokoły ocen z egzaminów z błędami. Przeprowadzono dodatkową kontrolę protokołów wśród wylosowanych nauczycieli z populacji $N = 2000$ pracowników dydaktycznych uniwersytetu. Weryfikujemy hipotezę sprawdzaną, określoną wyrażeniem (2.21), czyli $H_0 : \ M_U > M_0 = p_0 N = 300$, gdzie $p_0 = 0.15$, przy ryzyku jej błędnego odrzucenia na korzyść hipotezy konkurencyjnej $H_1 : M_U \leq 300$ wynoszącym $\kappa = 0.01$. Wartość krytyczną testu ustalamy na poziomie $p_\kappa = 0.07$. W związku z tym, że spełnione są założenia centralnego twierdzenia granicznego, korzystamy z wyrażenia (2.28) w celu wyznaczenia niezbędnej liczebności próby. Wówczas liczebność próby wyliczamy za pomocą następującej procedury zapisanej w języku R:

```
#wyznaczanie liczebności próby przy ustalonym kappa
#przybliżenie normalne
#liczebność populacji:
N < -2000
#Hipoteza
p0 < -0.15
M0 < -N * p0
#wartość krytyczna:
mpk < -0.07
#ryzyko:
kappa < -0.01
n < -1/(1/N + (N-1) * (wpk - p0)^2/(N * p0 * (1-p0) * qnorm(kappa)^2))
#niezbędna liczebność próby:
n-1
```

Zapisana procedura wylicza, że niezbędny rozmiar próby wynosi 103.

2.3.2. Przypadek próby warstwowej

Rozważamy dalej problem weryfikacji hipotez określonych równoważnymi wyrażeniami (2.20) i (2.21), lecz tym razem użyjemy statystyki testowej, będącej funkcją estymatora wskaźnika struktury z próby warstwowej. W pod-

rozdziale 1.3.3 wyjaśniono konstrukcję schematu losowania próby warstwowej. Niech populacja dokumentów będzie podzielona na H rozłącznych i niepustych warstw. Sumę liczebności prób prostych losowanych bezzwrotnie z warstw oznaczamy przez n. Do weryfikacji hipotezy sprawdzanej H_0 używa się statystyki:

$$Z_s = \frac{\bar{p}_s - p_0}{\sqrt{V_s(\bar{p}_s)}}, \tag{2.29}$$

gdzie:

$$\bar{p}_s = \sum_{h=1}^{H} w_h p_{s_h}, \qquad w_h = \frac{N_h}{N}, \tag{2.30}$$

$$V_s(\bar{p}_s) = \sum_{h=1}^{H} \frac{N_h - n_h}{(N_h - 1)n_h} w_h^2 p_{s_h}(1 - p_{s_h}), \tag{2.31}$$

przy czym przez p_{s_h} oznaczono częstość występowania nieprawidłowości w próbie prostej s_h o liczebności n_h, losowanej bezzwrotnie z h-tej warstwy o liczebności N_h.

Zastosujmy twierdzenie de Moivre'a–Laplace'a dla każdej warstwy z osobna. Wtedy przy założeniu, że rozmiar każdej warstwy osiąga co najmniej 1000 elementów, a rozmiar próby z niej losowanej ma liczebność równą co najmniej 100 elementów oraz prawdopodobieństwa wystąpienia nieprawidłowości w każdej warstwie nie jest mniejsze od 0.04, określona wzorem (2.29) statystyka Z_s ma rozkład normalny standardowy, gdy prawdziwą jest hipoteza sprawdzana H_0 określona wyrażeniem (2.20). Ponadto, zakłada się, iż jeśli liczebności prób rosną, to rosną również odpowiednie rozmiary warstw. Wówczas, przy ustalonej liczebności próby warstwowej wartość krytyczną p_κ testu hipotezy H_0 wyznacza wzór:

$$P\left(\bar{p}_s \leq p_\kappa | H_0\right) = \kappa, \tag{2.32}$$

gdzie:

$$p_\kappa = p_0 + z_\kappa \sqrt{V_s(\bar{p}_s)}, \tag{2.33}$$

przy czym $\kappa = P(Z < z_\kappa)$, $Z \sim N(0,1)$. Gdy zaobserwujemy, że $\bar{p}_s \leq p_\kappa$, to akceptujemy system kontroli wewnętrznej z prawdopodobieństwem (ryzykiem) zawyżonego do niego zaufania na poziome κ.

Przykład 2.13

Planuje się przeprowadzenie kontroli poprawności wypisywanych recept na leki, wystawionych w lutym br. Populacja recept jest podzielona na trzy warstwy ze względu na źródło ich pochodzenia. W pierwszej warstwie liczącej 55 000 recept są recepty wystawione w podstawowej opiece zdrowotnej. Recepty pochodzące ze specjalistycznej opieki zdrowotnej są w drugiej

warstwie, zawierającej 20 000 tych dokumentów. Trzecia warstwa liczy 5 000 recept wystawionych w lecznictwie zamkniętym. Z kolejnych warstw planuje się wylosować 300, 200 i 100 recept.

Prawidłowość wystawianych recept jest w sposób naturalny sprawdzana przez aptekarzy. W związku z tym apteki można traktować jako system kontrolujący jakość recept. Przyjmuje się, że dopuszczalny udział recept nieprawidłowo wystawionych ustalono na poziomie 0.05. Problem polega na sprawdzeniu sądu głoszącego, że nie jest skutecznym działanie tego swoistego systemu kontroli jakości recept przeciwko stwierdzeniu, że ten system działa skutecznie. Formalnie rzecz biorąc, prowadzi do tego weryfikacja hipotezy, że udział nieprawidłowych recept przekracza dopuszczalny poziom $p_0 = 0.05$, przy konkurencyjnej hipotezie, że ten poziom dopuszczalny nie jest przekroczony, czyli te hipotezy formalnie określa wyrażenie (2.20). Przyjmuje się, że ryzyko przyjęcia nieskutecznego systemu kontroli nie przekracza poziomu $\kappa = 0.1$.

Na podstawie informacji z wybranych aptek oszacowano, że udziały nieprawidłowo wypisanych aptek w kolejnych warstwach wynoszą 0.06, 0.02 i 0.1. Zgodnie z tymi ocenami najmniej nieprawidłowych recept pochodzi od specjalistów, a najwięcej od pracujących w szpitalach, czyli lecznictwie zamkniętym.

Przy określonych postulatach wartość krytyczną testu wyznaczamy na podstawie wzorów (2.33) i (2.31). Przyjmujemy, że $p_{s_1} = 0.06$, $p_{s_2} = 0.02$ i $p_{s_3} = 0.1$. Korzystając np. z arkusza kalkulacyjnego Excel lub programu R można wyliczyć, że szukana wartość krytyczna wynosi: $p_\kappa = 0.037$. Zatem, jeśli po kontroli recept wylosowanych do próby warstwowej o rozmiarze 600 stwierdzamy, że udział recept nieprawidłowych nie przekracza poziomu krytycznego p_κ, to odrzucamy hipotezę sprawdzaną, że system kontroli recept prowadzony w aptekach nie jest skuteczny, przy czym ta decyzja może być błędna z prawdopodobieństwem 0.1. Innymi słowy, wnioskujemy, że kontrola apteczna jakości recept jest skuteczna, z ryzykiem wydania błędnej decyzji wynoszącym 0.1. Z kolei w przeciwnym wypadku, gdy udział recept nieprawidłowych przekroczy poziom krytyczny p_κ, to nie ma podstaw do odrzucenia hipotezy, iż apteczna kontrola jakości recept jest nieskuteczna.

Zakładamy, że liczebności prób losowanych z poszczególnych warstw są proporcjonalne do rozmiarów odpowiednich warstw. Wtedy ocenę wariancji estymatora $\bar{p}_s$ określa wzór:

$$V_s(\bar{p}_s) = \frac{N-n}{n} \sum_{h=1}^{H} \frac{w_h^2}{(Nw_h - 1)} p_{s_h}(1 - p_{s_h}) \approx \frac{N-n}{Nn} \sum_{h=1}^{H} w_h p_{s_h}(1 - p_{s_h}). \tag{2.34}$$

Wówczas ze wzoru (2.33) wynika, że niezbędną liczebność n próby dla ustalonych κ i p_κ wyznacza wyrażenie:

$$n \geq \left[\frac{z_\kappa^2 \delta_0}{\frac{z_\kappa^2 \delta_0}{N} + (p_\kappa - p_0)^2} \right] + 1 \approx \left[\frac{z_\kappa^2 \delta_s}{(p_\kappa - p_0)^2} \right] + 1, \tag{2.35}$$

gdzie:

$$\delta_0 = \sum_{h=1}^{H} w_h p_h (1 - p_h) \leq \frac{1}{4}. \tag{2.36}$$

Założenie, że $p_h = \frac{1}{2}$ dla $h = 1, \ldots, H$, a co za tym idzie, iż $\delta_0 = \frac{1}{4}$ byłoby zbyt pesymistyczne, ponieważ doprowadzi to do znacznego zawyżenia liczebności próby ponad potrzebę wynikającą z postulowanej wartości ryzyka κ. Dlatego postuluje się, aby wartości p_h, $h = 1, \ldots, H$ ocenić na podstawie wstępnych prób prostych losowanych bezzwrotnie z warstw populacji. W zagadnieniach audytowych oczekuje się, że te wartości są bliskie zeru, a w każdym radzie nieprzekraczające poziomu $\frac{1}{5}$. Zatem, przyjmując, że $p_h \leq \frac{1}{5}$ dla $h = 1, \ldots, H$ otrzymujemy ze wzoru (2.36), że $\delta_0 \leq 0.16$.

Przykład 2.14

Populację faktur liczącą 20 000 podzielono na trzy warstwy na podstawie ich nominalnych wartości księgowych. Dopiero po dokonaniu kontroli wylosowanych faktur może okazać się, że te nominalne wartości trzeba będzie poprawić na prawdziwe. Pierwszą warstwę stanowiły faktury, których wartość nie przekraczała poziomu 200 zł. W drugiej warstwie są faktury o wartości większej od 200 zł i nie większej od 1000 zł, a w trzeciej warstwie są dokumenty kwocie większej od 1000 zł. W kolejnych warstwach znajduje się odpowiednio 40%, 30% i 30% liczby wszystkich fakur, czyli $w_1 = 0.4$, $w_2 = w_3 = 0.3$.

Weryfikujemy hipotezę sprawdzaną H_0 (por. wyrażenia (2.20) i (2.21)), głoszącą, że proporcja faktur z błędami nie przekracza wartości dopuszczalnej, $H_0 : \ p_U > p_0 = 0.05$, przy hipotezie alternatywnej $H_0 : \ p_U \leq p_0 = 0.05$. Ryzyko akceptacji nieprawidłowego systemu kontroli ma nie przekraczać poziomu $\kappa = 0.01$. Ponadto, ustalamy, że sformułowana hipoteza sprawdzana będzie odrzucana, gdy częstość występowania błędnych faktur nie przekroczy poziomu $p_\kappa = 0.01$.

Decydujemy się na losowanie próby warstwowej z liczebnościami prób proporcjonalnymi do wartości rozmiarów warstw. Przyjmujemy, że udział faktur z błędami w każdej warstwie nie przekracza poziomu 0.2, co daje wartość $\delta_0 = 0.16$. Wartość $z_\kappa = -2.3264$, ponieważ z tablic dystrybuanty rozkładu normalnego standardowego wynika, że $P(Z < z_\kappa) = 0.01 = \kappa$,

gdzie $Z \sim N(0,1)$. Te rezultaty i wzór (2.35) prowadzą do następującego wyniku:

$$n \geq \left[\frac{(-2.3264)^2 0.16}{\frac{(-2.3264)^2 0.16}{20000} + (0.01 - 0.05)^2} \right] + 1 = 527.$$

Teraz na podstawie wzoru (1.29) wyliczamy, że rozmiary prób, które mają być losowane z warstw wynoszą $n_1 = 527 \cdot 0.4 = 210.8 \approx 211$, $n_2 = n_3 = 527 \cdot 0.3 = 158.1 \approx 158$. Zatem tak wyznaczona próba spełnia ustalone postulaty, czyli hipoteza H_0 będzie odrzucana z ryzykiem popełnienia błędnej decyzji wynoszącym $\kappa = 0.01$, gdy zaobserwowana w próbie częstość występowania błędnych faktur $\bar{p}_s$ nie przekroczy wartości krytycznej $p_\kappa = 0.01$. W przeciwnym przypadku, jeśli $\bar{p}_s > 0.01$, to nie ma podstaw do odrzucenia hipotezy H_0.

2.4. Wnioskowanie przy ustalonym ryzyku dezaprobaty skutecznie funkcjonującego systemu kontroli

Niniejszy podrozdział jest niejako komplementarny w stosunku do treści poprzedniego. Tym razem proces weryfikacji skuteczności systemu kontroli jest prowadzony tak, aby ewentualnie podjąć możliwie małe ryzyko odrzucenia skutecznie działającego systemu kontroli wewnętrznej. Dodajmy jeszcze, iż prezentowane tutaj podejście do analizy skuteczności kontroli jest rzadko brane pod uwagę w praktyce, ponieważ zwykle większą wagę przywiązuje się do dbałości o nieprzekroczenie postulowanego ryzyka przyjęcia źle działającego systemu kontroli.

2.4.1. Przypadek próby prostej

Podobnie jak wcześniej, wartość dopuszczalną proporcji nieprawidłowości oznaczamy przez p_0. O wartości dyskwalifikującej p_1 wiemy jedynie tyle, że jest ona większa od dopuszczalnej, czyli $p_1 > p_0$. W związku z tym hipotezy formułowane w podrozdziale 2.2 wyrażeniem (2.1) upraszczają się do postaci

$$H_0 : \ p_U \leq p_0, \qquad H_1 : \ p_U > p_0. \tag{2.37}$$

Zatem H_0 głosi, że system kontroli jest skuteczny, bo proporcja nieprawidłowości jest nie wieksza od dopuszczalnej p_0. Z kolei H_1 głosi, że system kontroli jest nieskuteczny, bo proporcja nieprawidłowości przekracza dopuszczalną wartość p_0. Ryzyko η dezaprobaty skutecznego systemu kontroli jest teraz poziomem istotności testu statystycznego.

Zakładamy, że próba potrzebna do wnioskowania o skuteczności systemu kontroli wewnętrznej jest prostą losowaną bezzwrotnie. Proporcja

nieprawidłowości $p_U = \frac{M_U}{N}$, co pozwala wyżej określone hipotezy zastąpić następującymi równoważnymi:

$$H_0 : \ M_U \leq M_0, \qquad H_1 : \ M_U > M_0, \tag{2.38}$$

gdzie przez M_0 oznaczono ilość elementów nieprawidłowych w populacji o rozmiarze N. Hipoteza sprawdzana H_0 głosi, że liczba nieprawidłowości nie przekracza dopuszczalnego poziomu M_0, a hipoteza alternatywna H_1 głosi, że liczba nieprawidłowości jest większa od dopuszczalnego poziomu M_0.

Podobnie jak wcześniej do weryfikacji hipotezy H_0 użyjemy liczby wykrytych nieprawidłowości, oznaczanej dalej przez M_s. Ocena wartości M_s jest wyznaczana na podstwie próby prostej losowanej bezzwrotnie o liczebności n. Duże wartości statystyki testowej M_s świadczą przeciwko hipotezie sprawdzanej H_0, a na na korzyść hipotezy H_1. Poziom istotności testu, będący jednocześnie ryzykiem η dezaprobaty skutecznego systemu kontroli, określa wyrażenie:

$$\alpha = \eta \geq P(M_s \geq m_\eta | H_0), \tag{2.39}$$

gdzie m_η jest wartością krytyczną testu. Liczba m_η jest najmniejszą liczbą naturalną spełniającą nierówność (2.39).

Gdy $m_s \geq m_\eta$, to dezaprobujemy skuteczność systemu kontroli wewnętrznej z ryzykiem η popełnienia błędnej decyzji. Wynika to stąd, że spełnienie nierówności $m_s \geq m_\eta$ prowadzi do odrzucenia hipotezy H_0 na korzyść H_1 z prawdopodobieństwem popełnienia błędu I rodzaju równemu poziomowi istotności $\alpha = \eta$.

Gdy $m_s < m_\kappa$, to nie ma podstaw do dezaprobaty skuteczności systemu kontroli, ponieważ nie ma podstaw do odrzucenia hipotezy H_0.

W sytuacji gdy liczebność próby jest ustalona i równa n oraz ryzyko η jest również określone, to wartość krytyczną m_η wyznaczamy jako rozwiązanie nierówności:

$$\eta \geq P(M_s \geq m_\eta) = \sum_{k=m_\eta}^{min(n,M_0)} \frac{\binom{M_0}{k}\binom{N-M_0}{n-k}}{\binom{N}{n}}, \tag{2.40}$$

przy czym m_η jest najmniejszą liczbą spośród spełniających powyższe wyrażenie.

Przykład 2.15

Zbiór pytań testowych użytych do egzaminowania składa się z $N = 500$ sztuk. Egzaminator ustalił, że dopuszczalna liczba pytań, na które student może nie znać poprawnej odpowiedzi wynosi $M_0 = 100$. Podczas egzaminu student losuje $n = 50$ pytań. Gdy nie odpowie na co najmniej m_η pytań, to nie zdaje egzaminu. Ustalono, że ryzyko niezaliczenia egzaminu,

gdy student zna odpowiedzi na co najmniej 400 pytań, wynosi $\eta = 0.05$. Problem polega na ustaleniu wartości krytycznej liczby pytań m_η.

Z formalnego punktu widzenia mamy do czynienia z weryfikacją hipotezy $H_0: M_0 \leq 100$, przeciwko hipotezie $H_0: M_0 > 100$ (por. wyrażenie (2.38)). Po to aby wyznaczyć wartość krytyczną m_η testu spełniającą nierówność określoną wzorem (2.40), korzystamy z następującej instrukcji napisanej w języku komputerowym pakietu R:

$$qhyper(\eta, M_0, N - M_0, n, lower.tail = FALSE) + 1.$$

W naszym przypadku ta instrukcja konkretyzuje się do postaci:

$$qhyper(0.05, 100, 400, 50, lower.tail = FALSE) + 1.$$

Jej realizacja daje szukaną wartością krytyczną, która wynosi $m_\kappa = 16$. Jeśli student nie odpowie na 16 lub więcej pytań spośród 50 wylosowanych, to nie zalicza egzaminu, przy czym jest to błędna ocena zasobu jego wiedzy z prawdopodobieństwem (ryzykiem) równym 0.05. W przeciwnej sytuacji, jeśli student nie odpowie na mniej niż 16 pytań, to nie ma podstaw, aby mu nie zaliczyć egzaminu.

Rozwiązanie nierówności (2.40) upraszcza się, gdy liczebność populacji N będzie duża (rzędu co najmniej 1000). Wówczas, podobnie jak w poprzednim podrozdziale, można ją zastąpić następującą:

$$\eta \geq \sum_{k=m_\eta}^{n} \binom{n}{k} p_0^k (1 - p_0)^{n-k}, \qquad p_0 = \frac{M_0}{N}, \tag{2.41}$$

przy czym m_η jest najmniejszą liczbą spośród tych spełniających powyższą nierówność.

Przykład 2.16

Populacja faktur liczy 5000 dokumentów. Dopuszczalna liczba faktur z błędami wynosi 5% ich ilości. Z badanej zbiorowości wylosowano bezzwrotnie 50 faktur. Okazało się, że w wyniku przeprowadzonej kontroli wśród nich znalazły się trzy faktury z błędami. Czy z ryzykiem błędnej decyzji równym 0.1 można twierdzić, że system kontroli wewnętrznej jest nieskuteczny? Mamy więc do czynienia z weryfikacją określonej wzorami (2.37) lub (2.38)hipotezy sprawdzanej dla parametru $M_0 = p_0 N = 250$, przy czym $p_0 = 0.05$, $N = 5000$, $n = 50$, $\eta = 0.1$, $m_s = 3$. Wartość krytyczną testu wyznaczamy na podstwie wzoru (2.41). W tym celu używamy następującej instrukcji pakietu komputerowego R:

$$qbinom(\eta, n, p_0, lower.tail = FALSE) + 1$$

W naszym przypadku $qbinom(0.1, 50, 0.05, lower.tail = FALSE) = 5$, czyli $m_\eta = 5$. Zatem, $m_s = 3 < m_\eta = 5$, co prowadzi do wniosku, że nie

ma podstaw do odrzucenia hipotezy H_0 głoszącej, iż system kontroli jest skuteczny.

Wyrażenie (2.41) dotyczy sytuacji, gdy rozmiar populacji N jest duży, a prawdopodobieństwo p_0 umiarkowanie małe, czyli $0.05 \leq p_0 \leq 0.1$. W sytuacji, gdy $p_0 < 0.05$ i $N \geq 1000$ i $\lambda = np$ nierówność (2.40) zastępuje się przez następującą, wynikającą z przybliżenia Poissona:

$$\eta \geq 1 - e^{-np_0} \sum_{k=0}^{m_\eta - 1} \frac{(np_0)^k}{k!}, \tag{2.42}$$

przy czym m_η jest najmniejszą liczbą spośród tych, które spełniają powyższą nierówność.

Przykład 2.17

Spośród 300 000 wierszy tekstu pewnej wielotomowej encyklopedii wylosowano 1000 wierszy celem wykrycia w nich błędów, które przeoczyli korektorzy. W wyniku tego zaobserowano wśród nich 30 wierszy z błędani. Przyjmuje się, że dopuszczalny udział niezauważonych wierszy z błędami wynosi 0.05%. Czy z prawdopodobieństwem (ryzykiem) 0.05 można twierdzić, że system kontroli korektorskiej tekstu nie jest skuteczny? Reasumując, $N = 300000$, $n = 1000$, $p_0 = 0.005$, $\eta = 0.05$. Wartość krytyczną testu hipotezy H_0, danej wyrażeniem (2.37) wyznaczamy na podstawie wzoru (2.42), korzystając z następującej instrukcji programu R:

$$qpois(\eta, n * p_0, lower.tail = FALSE) + 1.$$

W naszym przypadku $m_\eta = qpois(0.05, 1000{*}0.005, lower.tail = FALSE) = 9$. Zatem, jeśli zaobserwowana w próbie liczba wierszy z błędami jest co najmniej równa 9, czyli $m_s \geq 9 = m_\eta$, to system kontroli korektorskiej uznajemy za nieskuteczny z prawdopodobieństwem popełnienia niewłaściwej decyzji równym 0.05. W sytuacji odwrotnej, gdy $m_s < 9 = m_\eta$, to nie ma podstaw, by nie uznać, że system kontroli korektorskiej jest skuteczny. W związku z tym w naszym przypadku stwierdzamy, że kontrola korektorska nie jest skuteczna z prawdopodobieństwem podjęcia błędnej decyzji równym 0.05.

W sytuacji gdy $N \geq 1000$, $p \geq 0.05$, $N > n \geq 100$, na podstawie znanego centralnego twierdzenia de Moivre'a–Laplace'a zaleca się, aby rozwiązanie nierówności (2.40) zastąpić rozwiązaniem następującego równania:

$$\eta = P\left(Z \geq \frac{p_\eta - p_0}{\sqrt{\frac{(N-n)}{(N-1)n} p_0 (1 - p_0)}} = z_\eta \right), \tag{2.43}$$

z którego wyliczamy, że

$$m_\eta = \left[np_0 + z_\eta \sqrt{\frac{N-n}{N-1} np_0(1-p_0)} \right] + 1, \tag{2.44}$$

przy czym $\eta = 1 - P(Z < z_\eta)$ i $Z \sim N(0;1)$.

Przykład 2.18

Spośród 10 000 zeznań podatkowych wylosowano bezzwrotnie próbkę prostą 200 sprawozdań. Po kontroli wylosowanych sprawozdań okazało się, że w 11 spośród nich są błędy. Dopuszcza się, że aż 10% sprawozdań źle przygotowanych może wystąpić jeszcze po pobieżnej kontroli czynionej podczas ich przyjmowania. Na poziomie ryzyka błędnej dezaprobaty jakości działania systemu kontroli wewnętrznej równym 0.05, sprawdzamy, czy system kontorli nie jest skuteczny, co sprowadza się do weryfikacji hipotezy określonej wzorem (2.37). Z racji tego, że wyjściowe warunki spełniają kryteria, przy których stosuje się centralne twierdzenie de Moivre'a–Laplace'a, wykorzystujemy wzór (2.44) do wyliczenia wartości krytycznej testu. Korzystając z instrukcji programu R:

$$qnorm\left(1-\eta, np_0, \sqrt{\frac{N-n}{(N-1)} np_0(1-p_0)}\right).$$

W naszym przypadku $N = 10000$, $n = 200$, $m_s = 11$, $p_0 = 0.1$, $\eta = 0.05$. To już pozwala wyliczyć wartość *qnorm(0.05,20,2.9699)=15.12*. Zatem, ze wzoru (2.44) wynika, że $m_\eta = 16$, czyli $m_s = 11 < m_\eta = 16$. W związku z tym nie ma podstaw do dezaprobaty skuteczności badanego systemu kontroli wewnętrznej.

W przypadku gdy audytor ustala poziom ryzyka η oraz wartość krytyczną testu m_η, to pozostaje wyznaczenie niebędnej liczebności próby, zapewniającej podjęcie decyzji o prawdziwości hipotezy H_0. Prowadzą do tego celu również nierówności (2.40)- (2.44), lecz tym razem rozwiązujemy je względem niewiadomej, którą jest właśnie szukana liczebność próby.

2.4.2. Przypadek prób prostych losowanych z warstw

Kontynuujemy dalej problem weryfikacji hipotez określonych równoważnymi wyrażeniami (2.37) lub (2.38), lecz tym razem użyjemy sprawdzianu testu będącego funkcją estymatora wskaźnika struktury z próby warstwowej. Skorzystamy z ustaleń wprowadzonych w poprzednim podrozdziale, a w szczególności z definicji statystyk zapisanych wyrażeniami (2.29)-(2.31).

Powtarzamy założenia, że rozmiar każdej warstwy osiąga co najmniej 1000 elementów, a rozmiar próby prostej bezzwrotnie losowanej z niej jest równy co najmniej 100 elementów. Ponadto, niech prawdopodobieństwo

wystąpienia nieprawidłowości w każdej warstwie nie jest mniejsze od 0.04. Wówczas, przy ustalonej liczebności próby warstwowej wartość krytyczną p_η testu hipotezy H_0 wyznacza wzór (2.31) oraz wyrażenia:

$$\eta = P(\bar{p}_s \geq p_\eta | H_0) \qquad i \qquad p_\eta = p_0 + z_\eta \sqrt{V_s(\bar{p}_s)}, \tag{2.45}$$

przy czym w analizowanej teraz sytuacji $\eta = P(Z \geq z_\eta)$, $Z \sim N(0,1)$.

Załóżmy, że liczebności prób losowanych z poszczególnych warstw są proporcjonalne do rozmiarów odpowiednich warstw, co pozwala wyliczyć wzór (1.29). Wtedy ocenę wariancji estymatora $\bar{p}_s$ określa wzór (2.34). Wówczas niezbędną liczebność n próby dla ustalonego ryzyka η i wartości krytycznej p_η wyznacza wyrażenie:

$$n \geq \left[\frac{z_\eta^2 \delta_0}{\frac{z_\eta^2 \delta_0}{N} + (p_\eta - p_0)^2} \right] + 1 \approx \left[\frac{z_\eta^2 \delta_s}{(p_\eta - p_0)^2} \right] + 1, \tag{2.46}$$

przy czym

$$\delta_0 = \sum_{h=1}^{H} w_h p_h (1 - p_h) \leq p_0(1 - p_0) \leq \frac{1}{4}. \tag{2.47}$$

Wartości p_h, $h = 1, \ldots, H$ są oceniane na podstawie wcześniejszych badań lub na sprawą wstępnych prób prostych losowanych bezzwrotnie z populacji. Wiadomo, że wartość δ jest tym większa, im wartości p_h, $h = 1, \ldots, H$ są bliższe $\frac{1}{2}$. Z kolei w zagadnieniach audytowych oczekuje się, że te wartości są bliskie zeru, a w każdym razie nie przekraczają poziomu $\frac{1}{5}$. Zatem, przyjmując, że $p_h \leq \frac{1}{5}$ dla $h = 1, \ldots, H$, otrzymujemy ze wzoru (2.36), że $\delta_0 \leq 0.16$.

Przykład 2.19

Populacja pytań egzaminacyjnych składa się z dwóch warstw jednorodnych ze względu na wiedzę, której dotyczą. W pierwszej warstwie jest 1000 pytań, w drugiej 1500. Zatem udziały pytań z poszczególnych warstw w ogólnej puli pytań wynoszą odpowiednio $w_1 = \frac{1}{3}$, $w_2 = \frac{2}{3}$. Dodajmy, że zwykle w przypadku egzaminów kończących studia wszystkie pytania są udostępniane studentom.

Egzaminatorzy wymagają, aby student znał odpowiedzi na przynajmniej 80% i 60% pytań zawartych odpowiednio w pierwszej, drugiej warstwie. Czyli student może nie odpowiedzieć na co najwyżej 20% i 40% pytań odpowiednio z kolejnych warstw. Udziały dopuszczalnej liczby pytań, na które student może nie znać odpowiedzi, wynoszą więc $p_1 = 0.2$ i $p_2 = 0.4$ odpowiednio dla pierwszej i drugiej warstwy. Dopuszcza się, że student zaliczy egzamin, jeśli nie potrafi odpowiedzieć na co najwyżej

$$p_0 = w_1 p_1 + w_2 p_2 = \frac{1}{3} \cdot 0.2 + \frac{2}{3} \cdot 0.4 = \frac{1}{3}$$

ogółu pytań. Tłumacząc to na język statystyki, egzaminowanie danego studenta polega na weryfikacji hipotezy sprawdzanej H_0 (por. wyrażenia (2.37) lub (2.38)), w naszym przypadku głoszącej, że proporcja pytań, na które student nie zna prawidłowej odpowiedzi, nie przekracza wartości dopuszczalnej $p_0 \leq \frac{1}{3}$ przy hipotezie alternatywnej stanowiącej, że $p_0 > \frac{1}{3}$. Egzamin polega na odpowiedzi na pytania wylosowane z poszczególnych warstw, przy czym ich ilości są proporcjonalne (przynajmniej w przybliżeniu) do rozmiarów odpowiednich warstw. Tak wylosowane pytania stanowią próbę warstwową. Egzaminatorzy ustalają, że student nie zdaje egzaminu, jeśli proporcja odpowiedzi nieprawidłowych na pytania spośród tych, które wylosował, jest co najmniej równa proporcji krytycznej wynoszącej $p_\eta = 0.4$. W przeciwnym przypadku, gdy ta proporcja pytań, na które nie potrafił odpowiedzieć jest mniejsza od p_η, to student zalicza egzamin. Zwróćmy uwagę, iż może zdarzyć się, że student mimo iż nie znał prawidłowych odpowiedzi na wystarczającą ilość pytań z całej listy, to wylosował próbę, w której proporcja pytań, na które nie umie odpowiedzieć jest mniejsza niż krytyczna proporcja p_η, a więc zdaje egzamin. W związku z tym mamy do czynienia z błędem II rodzaju, czyli przyjęciem hipotezy H_0, gdy nie jest prawdziwa. W kontekście rozważanego problemu ten błąd polega na niezasłużonym zdaniu egzaminu. Z drugiej strony udział pytań, na które student może nie znać odpowiedzi prawidłowych jest mniejsza niż dopuszalna proporcja p_0, ale wylosował tak pechowo zbiór (próbę) pytań, że proporcja nieprawidłowych odpowiedzi wynosi co najmniej p_η. To oznacza, że nie zdaje egzaminu. W tym przypadku mamy do czynienia błędem I rodzaju polegającym na odrzuceniu hipotezy H_0, gdy jest prawdziwa. Innymi słowy, ten błąd polega na niedocenieniu wystarczającej wiedzy studenta wynikłej z losowego doboru próby. Prawdopodobieństwo (ryzyko) popełnienia tego typu błędu w przypadku egzaminowania ustala się zwykle na niskim poziomie oznaczanym dalej przez η, aby nie krzywdzić studentów. Dlatego przyjmijmy, że $\eta = 0.01$.

Decydujemy się na losowanie próby warstwowej z liczebnościami prób proporcjonalnymi do wartości rozmiarów warstw. Nasz problem polega na wyznaczeniu rozmiaru tej próby, przy postulowanym ryzyku niesłusznego niezaliczenia egzaminu studentowi, wynoszącym $\eta = 0.01$, oraz dla ustalonej wartości proporcji krytycznej $p_\eta = 0.4$. Rozmiar potrzebnej próby wyznaczamy na podstawie wzorów (2.46) i (2.47), przy czym występujący w niej parametr δ_0 ustalamy na jego górnym możliwym poziomie, czyli $\delta = 0.25$. Z kolei wartość z_η spełnia wyrażenie $P(Z \geq z_\eta) = \eta$, przy czym $Z \sim N(0;1)$. Teraz wykorzystujemy następującą procedurę wyliczającą rozmiar próby określony wzorem (2.46), napisaną w języku R:

```
#wyznaczanie liczebności próby warstwowej przy ustalonym eta
#proporcjonalny wariant ustalania liczebności prób
#przybliżenie normalne
```

#liczebność populacji:
$N < -2500$
#Hipoteza
$H0 : p0 < -1/3$
#proporcja krytyczna:
$peta < -0.4$
#ryzyko:
$eta < -0.01$
#wartość parametru delta0:
$delta0 < -0.25$
$n < -ceiling(qnorm(1-eta)^2 * delta0/(qnorm(1-eta)^2 * delta0/N + (peta - p0)^2)) + 1$

W wyniku realizacji tej procedury otrzymujemy, że $n \geq 230$. Teraz na podstawie wzoru (1.29) wyliczamy, że rozmiary prób, które mają być losowane z warstw wynoszą $n_1 = 230 \cdot \frac{1}{3} = 76.7 \approx 77$, $n_2 = 230 \cdot \frac{1}{3} = 153.3 \approx 153$. Zatem, tak wyznaczona próba spełnia ustalone postulaty. Z faktu, że $p_\eta = 0.4$ i rozmiar próby $n = 230$ wynika, że krytyczna liczba pytań wynosi $m_\eta = n \cdot p_\eta = 230 \cdot 0.4 = 92$. Wówczas, jeśli spośród wylosowanych z warstw pytań student nie odpowie na co najmniej 92 pytania, to nie zdaje egzaminu z ryzykiem podjęcia błędnej decyzji (polegającym na niedocenieniu jego wiedzy) wynoszącym $\eta = 0.01$. Z kolei, jeśli student nie odpowie na mniej niż 92, to zdaje egzamin, a precyzyjniej mówiąc, nie ma podstaw do odrzucenia hipotezy stanowiącej, że wiedza studenta jest wystarczająca.

Rozdział 3

BADANIE WIARYGODNOŚCI SPRAWOZDAŃ

3.1. Definicje podstawowe

Dotychczas zajmowaliśmy się częstością względną występowania błędnych np. zapisów księgowych. Teraz jeszcze uwzględniamy ich wartości, czyli np. kwoty, na które opiewa faktura. Chodzi o to, że w sposób niezamierzony lub świadomy mogły wystąpić błędy dotyczące tych kwot. W szczególności z powodu pomyłki zostały źle zsumowane liczby, co doprowadziło do zapisania błędnej wartości faktury. W wyniku kontroli ten błąd jest wychwytywany i poprawiany. Populację dokumentów księgowych oznaczamy tak jak dotychczas, czyli przez U, natomiast podpopulację, w której występują błędy w dokumentach – przez U_1, a podpopulację, w której nie ma błędów – przez U_0. Zatem $U = U_0 \cup U_1$. Przypomnijmy, że liczebności zbiorów U, U_0 i U_1 oznaczono już odpowiednio przez N, N_0 i N_1. Obserwowaną ewentualnie z błędami zmienną oznaczamy dalej przez x, a jej wartość przypisaną k-temu elementowi populacji przez x_k, gdzie $k = 1, \ldots, N$. Prawdziwą wartość zmiennej, np. wartość faktury, oznaczamy dalej przez y, a jej wartość przez y_k, $k = 1, \ldots, N$. Wówczas błąd obserwacji zwykle określa się przez różnicę $e_k = x_k - y_k$, którą traktujemy jako wartość zmiennej e. Oznaczenia podstawowych parametrów tych zmiennych zamieszczono w podrozdziale 1.2. Z punktu widzenia badań audytowych błąd obserwacji zapisuje się szczegółowo następująco:

$$e_k = \begin{cases} 0, & \text{gdy} \quad k \in U_0, \\ x_k - y_k, & \text{gdy} \quad k \in U_1. \end{cases} \tag{3.1}$$

Wówczas stąd i ze wzoru (1.5) mamy:

$$\bar{x}_U = w_0 \bar{y}_{U_0} + w_1 \bar{x}_{U_1} \tag{3.2}$$

lub

$$\bar{x}_U = \bar{y}_U + w_1 \bar{e}_{U_1}, \tag{3.3}$$

przy czym

$$w_0 = \frac{N_0}{N}, \quad w_1 = \frac{N_1}{N}, \quad w_0 + w_1 = 1,$$

$$\bar{y}_{U_0} = \frac{1}{N_0}\sum_{k\in U_0} y_k, \quad \bar{x}_{U_1} = \frac{1}{N_1}\sum_{k\in U_1} x_k, \quad \bar{e}_{U_1} = \frac{1}{N_1}\sum_{k\in U_1} e_k.$$

Dodajmy, że wartość globalną zmiennej x można wyrazić następująco:

$$x_U = N(\bar{y}_U + w_1\bar{e}_{U_1}). \tag{3.4}$$

Ze wzoru (3.4) mamy:

$$e_U = x_U - y_U = \sum_{k\in U_1} e_k = Nw_1\bar{e}_{U_1}, \tag{3.5}$$

$$g_U = \frac{e_U}{y_U} = w_1\frac{\bar{e}_{U_1}}{\bar{y}_U}. \tag{3.6}$$

Stąd wynika, że rozmiar błędu oceny wartości globalnej zmiennej y jest wprost proporcjonalny do częstości występowania błędów obserwacji poszczególnych wartości zmiennej y. Ponadto, błąd oceny wartości globalnej zmiennej y zależy od średniego poziomu wartości występujących błędów.

Kolejne parametry zmiennych x i y są następujące. Wariancję zmiennej x można zdekomponować jak pokazuje wyrażenie:

$$v_U(x) = v_U(y) + 2w_1\sqrt{v_{U_1}(y)v_{U_1}(e)}r_{U_1}(y,e) + w_1v_{U_1}(e) + \\ + w_0w_1\left(\bar{e}_{U_1} + 2\left(\bar{y}_{U_1} - \bar{y}_{U_0}\right)\right)\bar{e}_{U_1}, \tag{3.7}$$

przy czym współczynnik korelacji między zmienną y, a błędami obserwacji e w podpopulacji U_1 określa wzór:

$$r_{U_1}(y,e) = \frac{v_{U_1}(y,e)}{\sqrt{v_{U_1}(y)v_{U_1}(e)}},$$

gdzie:

$$v_{U_1}(y) = \frac{1}{N_1}\sum_{k\in U_1}(y_k - \bar{y}_{U_1})^2, \qquad v_{U_1}(e) = \frac{1}{N_1}\sum_{k\in U_1}(e_k - \bar{e}_{U_1})^2,$$

$$v_{U_1}(y,e) = \frac{1}{N_1}\sum_{k\in U_1}(y_k - \bar{y}_{U_1})(e_k - \bar{e}_{U_1}), \quad \bar{y}_{U_1} = \frac{1}{N_1}\sum_{k\in U_1} y_k.$$

W szczególności ze wzoru (3.7) wynika, że jeśli średnia występujących błędów jest równa zeru, czyli $\bar{e}_{U_1} = 0$, to wzór (3.7) upraszcza się do postaci:

$$v_U(x) = v_U(y) + 2w_1\sqrt{v_{U_1}(y)v_{U_1}(e)}r_{U_1}(y,e) + w_1v_{U_1}(e). \tag{3.8}$$

Z kolei, jeśli $\bar{y}_{U_1} = \bar{y}_{U_0}$, to

$$v_U(x) = v_U(y) + 2w_1\sqrt{v_{U_1}(y)v_{U_1}(e)}r_{U_1}(y,e) + w_1v_{U_1}(e) + w_0w_1\bar{e}^2_{U_1}. \tag{3.9}$$

Dekompozycja kowariancji zmiennych x i y ma postać:

$$v_U(x,y) = v_U(y) + w_1\sqrt{v_{U_1}(y)v_{U_1}(e)}r_{U_1}(y,e) + w_0w_1\left(\bar{y}_{U_1} - \bar{y}_{U_0}\right)\bar{e}_{U_1}. \quad (3.10)$$

W szczególności, jeśli $\bar{e}_{U_1} = 0$ lub $\bar{y}_{U_1} = \bar{y}_{U_0}$, to

$$v_U(x,y) = v_U(y) + w_1\sqrt{v_{U_1}(y)v_{U_1}(e)}r_{U_1}(y,e). \quad (3.11)$$

Zauważmy jeszcze, iż ze wzorów (3.7) i (3.10) wynika, że

$$v_U(x) = 2v_U(x,y) - v_U(y) + w_1v_{U_1}(e) + w_0w_1\bar{e}^2_{U_1}. \quad (3.12)$$

Teraz szczegółowo rozważmy zapis współczynnika rzetelności obserwacji zmiennych wprowadzonego wzorami (1.11) i (1.12), po to, aby głębiej wniknąć w jego własności. W tym celu należy po prawej stronie wzoru (1.11) odpowiednio podstawić prawe strony wzorów (3.10) i (3.12). Wtedy mamy:

$$r_{xy} = \frac{v_U(y) + w_1\sqrt{v_{U_1}(y)v_{U_1}(e)}r_{U_1}(y,e) + w_0w_1\left(\bar{y}_{U_1} - \bar{y}_{U_0}\right)\bar{e}_{U_1}}{\sqrt{v_U(y)\left(2v_U(x,y) - v_U(y) + w_1v_{U_1}(e) + w_0w_1\bar{e}^2_{U_1}\right)}}. \quad (3.13)$$

Otrzymamy skomplikowany zapis analizowanego wskaźnika, z którego nie jest łatwo wyciągnąć jasne wnioski. Dlatego upraszczając nieco rozważania, najpierw załóżmy, że $\bar{e}_{U_1} = 0$. Wtedy:

$$r_{xy} = \frac{v_U(y) + w_1\sqrt{v_{U_1}(y)v_{U_1}(e)}r_{U_1}(y,e)}{\sqrt{v_U(y)\left(v_U(y) + 2w_1\sqrt{v_{U_1}(y)v_{U_1}(e)}r_{U_1}(y,e) + w_1v_{U_1}(e)\right)}}. \quad (3.14)$$

Jeśli $\bar{y}_{U_1} = \bar{y}_{U_0}$, to

$$r_{xy} = \frac{v_U(y) + w_1\sqrt{v_{U_1}(y)v_{U_1}(e)}r_{U_1}(y,e)}{\sqrt{v_U(y)\left(v_U(y) + 2w_1\sqrt{v_{U_1}(y)v_{U_1}(e)}r_{U_1}(y,e) + w_1v_{U_1}(e) + w_0w_1\bar{e}^2_{U_1}\right)}}. \quad (3.15)$$

Gdy $r_{U_1}(y,e) = 0$, to

$$r_{xy} = \frac{v_U(y) + w_0w_1\left(\bar{y}_{U_1} - \bar{y}_{U_0}\right)\bar{e}_{U_1}}{\sqrt{v_U(y)\left(v_U(y) + w_1v_{U_1}(e) + w_0w_1\left(\bar{e}_{U_1} + 2\left(\bar{y}_{U_1} - \bar{y}_{U_0}\right)\right)\bar{e}_{U_1}\right)}}. \quad (3.16)$$

W końcu, jeśli $r_{U_1}(y,e) = 0$ i $\bar{e}_{U_1} = 0$, to

$$r_{xy} = \frac{v_U(y)}{\sqrt{v_U(y)\left(v_U(y) + w_1v_{U_1}(e)\right)}} = \sqrt{\frac{v_U(y)}{v_U(y) + w_1v_{U_1}(e)}}. \quad (3.17)$$

Stąd wynika, że współczynnik rzetelności obserwacji wartości zmiennej y jest malejącą funkcją wariancji błędu obserwacji $v_{U_1}(e)$ oraz udziału w_1 rozmiaru podpopulacji U_1 (w której występują błędy) w liczebności całej populacji.

3.2. Oceny podstawowych charakterystyk błędów księgowych

3.2.1. Próba prosta

Naturalny estymator wartości globalnej y_U i różnicy $e_U = x_U - y_U$ to:

$$y_S = N\bar{y}_s, \quad \text{i} \quad e_s = x_U - y_s = N(\bar{x} - \bar{y}_s). \tag{3.18}$$

Wariancja tego estymatora ma postać:

$$V(y_S) = \frac{N(N-n)}{n} v_{*U}(y) = \frac{N(N-n)}{n} v_{*U}(e) = V(e_S). \tag{3.19}$$

Nieobciążoną ocenę tego estymatora daje statystyka:

$$V_s(y_S) = V_s(e_S) = \frac{N(N-n)}{n} v_{*s}(y), \quad v_{*s}(y) = \frac{1}{n-1} \sum_{k \in s} (y_k - \bar{y}_s)^2. \tag{3.20}$$

Znane twierdzenia graniczne (por. np. Hájek, 1960; Fuller, 2009) pozwalają wnioskować, że jeśli $n \to \infty$, $N \to \infty$, $N - n \to \infty$ oraz rozkład statystyki e_s spełnia pewne ogólne warunki jednorodności, to

$$Z_{1s} = \frac{e_s - e_U}{\sqrt{V_s(e_s)}} \to Z \ \sim N(0,1). \tag{3.21}$$

Ta własność umożliwia estymację przedziałową przeciętnej błędów obserwacji, a także weryfikację hipotez o tej wartości średniej.

Przykład 3.1

Na podstawie wielokrotnych badań audytowych oceniono, że stopień zróżnicowania (mierzonego odchyleniem standardowym) wartości błędów wynosi 0.4 zł. Kolejne badanie audytowe jest przygotowywane. Zakłada się, że oceny sumy błędów księgowych mają być oszacowane z dokładnością do 500 zł. Zadanie polega na wyznaczeniu niezbędnej liczebności próby, która ma być losowana z populacji o rozmiarze 20 400 dokumentów, potrzebnej do tej estymacji z postulowaną dokładnością.

Najpierw załóżmy, że ocena przeciętnego błędu księgowego będzie wyznaczana na podstawie estymatora określonego wzorem (3.18). Postulowaną dokładność estymacji oznaczmy przez d, czyli $d = 500$. Wówczas ze wzoru na jego wariancję określonego wyrażeniem (3.19) wynika, że

$$n \geq n_0 = \left[\frac{N}{1 + \frac{d^2}{N v_{*U}(e)}} \right] + 1. \tag{3.22}$$

Ten wzór pozwala już wyliczyć szukaną liczebność:

$$n_0 = \left\lceil \frac{20400}{1 + \frac{500^2}{20400 \cdot 0.4^2}} \right\rceil + 1 = 263.$$

Zatem, jeśli postulujemy, aby precyzja estymacji sumy wartości błędów księgowych za pomocą statystyki e_s określonej wzorem (3.18) nie przekroczyła poziomu 500 zł, to należy wylosować próbę prostą o liczebności co najmniej 263 elementów.

3.2.2. Próba monetarna

Kilka planów i schematów losowania tzw. prób monetarnych (walutowych, dolarowych) opisano w podrozdziale 1.3.2. Do oceny wartości globalnej y_U użyjemy estymatora:

$$\tilde{y}_s = N\bar{y}_{HTs}, \tag{3.23}$$

przy czym $\bar{y}_{HTs}$ jest znanym estymatorem Horvitza-Thompsona (1952), por. również Bracha (1996) lub Wywiał (2010), który określa wzór:

$$\bar{y}_{HTs} = \frac{1}{N} \sum_{k \in s} \frac{y_k}{\pi_k}. \tag{3.24}$$

To prowadzi do następującego estymatora wartości globalnej błędów audytowych.

$$e_{HTs} = N\left(\bar{x} - \bar{y}_{HTs}\right). \tag{3.25}$$

Załóżmy, że prawdopodobieństwa inkluzji π_k, $k = 1, \ldots, N$, są proporcjonalne do wartości zmiennej dodatkowej (diagnostycznej) oraz są wyznaczane na podstawie wzorów (1.18)-(1.20). Ponadto załóżmy, że próba jest losowana zwrotnie za pomocą schematu losowania odrzucającego powtórzenia elementów populacji w próbie, który określają wzory (1.21) i (1.22). Hájek (1964, 1981) wykazał, że przy spełnieniu dość ogólnych warunków oraz przy założeniu, że $n \to \infty$ i $N - n \to \infty$ rozkład prawdopodobieństwa statystyki $\bar{y}_{HTs}$ zmierza do rozkładu normalnego z wartością oczekiwaną $E\left(\bar{y}_{HTs}\right) = \bar{y}$ oraz wariancją określoną wzorem:

$$V\left(\bar{y}_{HTs}\right) = \frac{1}{N^2} \sum_{k=1}^{N} \left(\frac{y_k}{\pi_k} - \frac{x_U}{n} I_{y:x}\right)^2 (1 - \pi_k)\pi_k, \tag{3.26}$$

przy czym

$$I_{y:x} = \frac{\sum_{k=1}^{N} y_k(1 - \pi_k)}{\sum_{k=1}^{N} x_k(1 - \pi_k)}, \qquad x_U = \sum_{k=1}^{N} x_k$$

lub wzorem:

$$V\left(\bar{y}_{HTs}\right)=\frac{1}{2N^2}\sum_{k=1}^{N}\sum_{j=1}^{N}\left(\frac{y_k}{\pi_k}-\frac{y_j}{\pi_j}\right)^2(\pi_k\pi_j-\pi_{kj}). \tag{3.27}$$

przy czym π_{kj} jest prawdopodobieństwem wylosowania do tej samej próby elementów populacji o numerach k oraz j do próby.

Załóżmy, że błędna wartość zapisu księgowego jest modelowana wzorem (1.5). Wówczas można wykazać, że

$$V\left(\bar{y}_{HTs}\right)=\frac{\bar{x}}{n}\left(v_{2,1}\left(h,x\right)+\bar{x}v_{2,0}(h,x)-\bar{x}^2\left(\frac{\bar{e}}{\bar{x}}-\bar{h}\right)^2\right). \tag{3.28}$$

przy czym wartości zmiennej h są wartościami błędów względnych, czyli $h_k=\frac{e_k}{x_k}$ oraz

$$v_{u,z}(h,x)=\frac{1}{N}\sum_{k=1}^{N}(h_k-\bar{h})^u(x_k-\bar{x})^z. \tag{3.29}$$

Do oceny wariancji Hájek (1964) proponuje estymator:

$$V_s\left(\bar{y}_{HTs}\right)=\frac{n}{N^2(n-1)}\sum_{k\in s}\left(\frac{y_k}{\pi_k}-\frac{x_U}{n}I_{y:x,s}\right)^2(1-\pi_k), \tag{3.30}$$

przy czym

$$I_{y:x,s}=\frac{\sum_{k\in s}\frac{y_k}{x_k}(1-\pi_k)}{\sum_{k\in s}(1-\pi_k)}\approx\frac{1}{n}\sum_{k\in s}\frac{y_k}{x_k}.$$

Drugi estymator ma postać:

$$\hat{V}_s\left(\bar{y}_{HTs}\right)=\frac{1}{2N^2}\sum_{k\in s}\sum_{j\in s}\left(\frac{y_k}{\pi_k}-\frac{y_j}{\pi_j}\right)^2(1-\pi_k)(1-\pi_j), \tag{3.31}$$

Niech:

$$Z_s=\frac{\bar{y}_{HTs}-\bar{y}}{\sqrt{V_s\left(\bar{y}_{HTs}\right)}}. \tag{3.32}$$

Przyjmując, że założenia znanego twierdzenia granicznego Hájka (1964) są spełnione, Wywiał (2012) wykazuje, że estymator $\hat{V}_s\left(\bar{y}_{HTs}\right)$ daje zgodne oceny wariancji $V\left(\bar{y}_{HTs}\right)$, a także że statystyka Z_s ma granicznie rozkład normalny standardowy, jeśli $n\to\infty$ i $N-n\to\infty$.

Ze wzorów (3.25)-(3.30) otrzymujemy następujące wzory na wartość oczekiwaną, wariancję estymatora wartości globalnej błędów oraz estymator tej wariancji:

$$E\left(e_{HTs}\right)=e_U=x_u-y_U,$$

$$V\left(e_{HTs}\right) = N^2 V\left(\bar{y}_{HTs}\right),$$
$$V_s\left(e_{HTs}\right) = N^2 V_s\left(\bar{y}_{HTs}\right).$$

Niech:

$$U_s = \frac{e_{HTs} - e_U}{\sqrt{V_s\left(e_{HTs}\right)}}. \tag{3.33}$$

Wówczas przy tych samych założeniach określających warunki zbieżności rozkładu prawdopodobieństwa zmiennej losowej Z_s do rozkładu granicznego rozkład prawdopodobieństwa statystyki U_s zmierza do rozkładu normalnego standardowego.

3.2.3. Próba warstwowa

Plan i schemat losowania próby warstwowej $s = \{s_1, s_2, \ldots, s_H\}$ opisano w podrozdziale 1.3.3.

Estymator wartości globalnej y_U ma postać:

$$y_{w,s} = N\bar{y}_{w,s}, \tag{3.34}$$

gdzie:

$$\bar{y}_{w,s} = \sum_{h=1}^{H} w_h \bar{y}_{s_h}, \quad \bar{y}_{s_h} = \frac{1}{n_h}\sum_{k\in s_h} y_k. \tag{3.35}$$

Estymator błędów księgowych e_U, czyli różnicy $e_U = y_U - x_U$ określa wzór:

$$e_{w,s} = y_{w,s} - x_U. \tag{3.36}$$

Wariancje statystyk $y_{w,s}$ i $e_{w,s}$ są takie same, a określa je wzór:

$$V(y_{w,s}) = V(e_{w,s}) = N^2 V(\bar{y}_{w,s}) = \sum_{h=1}^{H} \frac{N_h(N_h - n_h)}{n_h} v_{*U_h}(y), \tag{3.37}$$

przy czym

$$v_{*U_h}(y) = \frac{1}{N_h - 1}\sum_{k\in U_h}(y_k - \bar{y}_{U_h})^2, \qquad \bar{y}_{U_h} = \frac{1}{N_h - 1}\sum_{k\in U_h} y_k.$$

Nieobciążony estymator tej wariancji ma postać:

$$V_s(y_{w,s}) = V_s(e_{w,s}) = \sum_{h=1}^{H} \frac{N_h(N_h - n_h)}{n_h} v_{*s_h}(y), \tag{3.38}$$

przy czym

$$v_{*s_h}(y) = \frac{1}{n_h - 1}\sum_{k\in s_h}(y_k - \bar{y}_{s_h})^2. \tag{3.39}$$

Gdy liczebności podprób s_h, $h = 1, \ldots, H$ są ustalone proporcjonalnie do rozmiarów warstw, czyli $n_h = nw_h$, to wówczas

$$V(y_{w,s}) = V(e_{w,s}) = \frac{N(N-n)}{n} \sum_{h=1}^{H} w_h v_{*U_h}(y), \tag{3.40}$$

$$V_s(y_{w,s}) = V_s(e_{w,s}) = \frac{N(N-n)}{n} \sum_{h=1}^{H} w_h v_{*s_h}(y). \tag{3.41}$$

Na podstawie odpowiednich twierdzeń granicznych (por. Fuller, 2009) wnioskujemy, że jeśli dla $h = 1, \ldots, H$ $n_h \to \infty$, $N_h \to \infty$, $N_h - n_h \to \infty$ oraz rozkłady średnich błędów w warstwach spełniają pewne ogólne warunki jednorodności, to

$$Z_{w,s} = \frac{e_{w,s} - e_U}{\sqrt{V_s(e_{w,s})}} \to Z \ \sim N(0,1). \tag{3.42}$$

W praktyce zwykle postuluje się, aby rozmiary każdej z warstw lub zawartej w niej domeny były równe co najmniej 1000, a liczebność każdej próby losowanej z odpowiedniej warstwy była równa co najmniej 100.

3.3. Podejmowanie decyzji o wiarygodności sprawozdania

Podstawą rozważań prowadzonych w niniejszej części jest założenie, że obserwowane wartości zapisów księgowych są generowane zgodnie z równaniem (1.5). Zatem błąd zapisu danej pozycji księgowej jest wielkością nielosową. Wówczas, losowość dalej rozważanych estymatorów lub sprawdzianów testów ma swoje źródło tylko w stosowanym schemacie wyboru próby.

3.3.1. Ustalone ryzyko odrzucenia wiarygodnego sprawozdania i ryzyko akceptacji niewiarygodnego sprawozdania

Zadanie polega w szczególności na sprawdzeniu, czy suma zaksięgowanych wartości faktur dotyczących sprzedaży towarów przez przedsiębiorstwo jest rzeczywiście równa wartości sprzedanych towarów. Oznaczając sumę ksiegową przez x_U, a sumę prawdziwą przez y_U, weryfikujemy następującą hipotezę sprawdzaną H_0 względem alternatywnej H_1:

$$H_0 : \ |x_U - y_U| = |e_U| \leq e_0, \qquad H_1 : \ |e_U| \geq e_1 > e_0, \tag{3.43}$$

przy czym e_0 jest dopuszczalnym poziomem niezgodności sumy wartości księgowej dokumentów z rzeczywistą sumą wartości dokumentów, natomiast

e_1 jest niedopuszczalnym (dyskwalifikującym) poziomem niezgodności. Innymi słowy, hipoteza H_0 głosi, że kontrolowane sprawozdanie księgowe bilansuje się z błędem dopuszczalnym.

Na podstawie niżej prezentowanego testu podejmujemy decyzje o przyjęciu jednej z dwóch określonych hipotez. W związku z tym, że decyzje podejmujemy na podstawie próby losowej, to możemy odrzucić hipotezę H_0, gdy jest prawdziwa, czyli popełniamy błąd pierwszego rodzaju. W tym przypadku to oznacza odrzucenie sprawozdania finansowego, gdy ono jest prawidłowe. Z kolei jeśli przyjmiemy hipotezę H_0, gdy jest fałszywa, to popełniamy błąd drugiego rodzaju, co oznacza, że akceptujemy sprawozdanie mimo że jest ono nieprawidłowe.

Postawione hipotezy są weryfikowane na podstawie tzw. statystyki testowej (sprawdzianu testu), który jest zwykle estymatorem $\tilde{e}_s$ parametru e_U. Ponadto zakładamy, że $\tilde{e}_s$ ma granicznie rozkład normalny z wartością oczekiwaną $E(\tilde{e}_s) = e_{U,i}$, $i = 1, 2$, przy czym $E(\tilde{e}_s) = e_0$, gdy hipoteza sprawdzana H_0 jest prawdziwa, natomiast $E(\tilde{e}_s) = e_1$, jeśli hipoteza H_1 jest prawdziwa. Zakładamy, że $V_s(\tilde{e}_s)$ jest zgodnym estymatorem wariancji $V_s(\tilde{e}_s)$ i $\tilde{e}_s \sim N(e_{U,i}, V(\tilde{e}_s))$, $i = 1, 2$, gdy $N \to \infty$, $n \to \infty$ i $N - n \to \infty$. Ponadto, te założenia pozwalają wnioskować na podstawie zestandaryzowanej wartości statystyki testowej, która ma postać:

$$Z_s = \frac{\tilde{e}_s - e_0}{\sqrt{V_s(\tilde{e}_s)}}. \tag{3.44}$$

W przypadku prób prostej, monetarnej i warstwowej (z proporcjonalną lokalizacją próby w warstwach) statystykę testową Z_s określają wzory (3.21) albo (3.33), albo (3.42).

Załóżmy, że liczebności N, n i $N-n$ są na tyle duże, iż rozkład prawdopodobieństwa statystyki Z_s jest dobrze aproksymowany rozkładem normalnym z jednostkową wariancją. Gdy hipoteza H_0 jest prawdziwa, to $Z_s \sim N(0, 1)$, a gdy hipoteza H_1 jest prawdziwa, to $Z_s \sim N(\delta, 1)$, przy czym $\delta = \frac{e_1 - e_0}{\sqrt{V(\tilde{e}_s)}}$. Wówczas postulaowaną wartość η ryzyka odrzucenia wiarygodnego (bilansującego się) sprawozdania określa wyrażenie:

$$\eta = P\left(\tilde{e}_s \leq -z_{\eta/2}\sqrt{V_s(\tilde{e}_s)} + e_0 \text{ albo } \tilde{e}_s \geq z_{\eta/2}\sqrt{V_s(\tilde{e}_s)} + e_0 | H_0\right), \tag{3.45}$$

co jest równoważne wyrażeniu:

$$\eta = P(|Z_s| \geq z_{\eta/2} | H_0) \tag{3.46}$$

lub

$$\eta = P(|Z| \geq z_{\eta/2}), \qquad Z \sim N(0, 1). \tag{3.47}$$

Korzystając ze wzoru (3.45), ryzyko κ przyjęcia niewiarygodnego (niebilansujcego się) sprawozdania określamy przez:

$$1-\kappa = P\left(\tilde{e}_s \le -z_{\eta/2}\sqrt{V_s(\tilde{e}_s)} + e_0 \ \text{ albo } \ \tilde{e}_s \ge z_{\eta/2}\sqrt{V_s(\tilde{e}_s)} + e_0 | H_1\right), \tag{3.48}$$

które to wyrażenie jest równoważne następującym:

$$1-\kappa = P\left(\frac{\tilde{e}_s - e_1}{\sqrt{V_s(\tilde{e}_s)}} \le -z_{\eta/2} + \frac{e_0 - e_1}{\sqrt{V_s(\tilde{e}_s)}} \right.$$
$$\left. \text{albo} \quad \frac{\tilde{e}_s - e_1}{\sqrt{V_s(\tilde{e}_s)}} \ge z_{\eta/2} + \frac{e_0 - e_1}{\sqrt{V_s(\tilde{e}_s)}} | H_1 \right),$$

$$1-\kappa = P\left(Z \le -z_{\eta/2} + \frac{e_0 - e_1}{\sqrt{V_s(\tilde{e}_s)}} \ \text{ albo } \ Z \ge z_{\eta/2} + \frac{e_0 - e_1}{\sqrt{V_s(\tilde{e}_s)}}\right),$$

gdzie $Z \sim N(0,1)$,

$$1-\kappa = \phi\left(-z_{\eta/2} + \frac{e_0 - e_1}{\sqrt{V_s(\tilde{e}_s)}}\right) + 1 - \phi\left(z_{\eta/2} + \frac{e_0 - e_1}{\sqrt{V_s(\tilde{e}_s)}}\right),$$

$$\kappa = \phi\left(z_{\eta/2} + \frac{e_0 - e_1}{\sqrt{V_s(\tilde{e}_s)}}\right) - \phi\left(-z_{\eta/2} + \frac{e_0 - e_1}{\sqrt{V_s(\tilde{e}_s)}}\right), \tag{3.49}$$

gdzie $\phi(z) = P(Z < z)$, $Z \sim N(0,1)$, czyli $\phi(z)$ jest dystrybuantą rozkładu normalnego standardowego. Przypomnijmy, że $\beta = 1 - \kappa$ jest mocą testu, czyli prawdopodobieństwem niepopełnienia błędu II rodzaju.

Na podstawie wzorów (3.57) i (3.49) można wykazać, że przy ustalonym ryzyku η wraz ze wzrostem liczebności próby ryzyko κ maleje do zera.

Proces podjęcia ostatecznej decyzji jest następujący. Jeśli wartość z_s sprawdzianu testu Z_s spełnia nierówność $|z_s| \ge z_{\eta/2}$, to hipotezę H_0 odrzucamy na korzyść H_1 z prawdopodobieństwem popełnienia błędnej decyzji (ryzykiem odrzucenia prawdłowego sprawozdania) równym η. W przeciwnym przypadku, gdy $|z_s| < z_{\eta/2}$, to hipotezę H_0 akceptujemy z prawdopodobieństwem popełnienia błędnej decyzji (ryzykiem przyjęcia nieprawidłowego sprawozdania) równym κ.

Drugi, równoważny powyższemu, sposób podejmowania decyzji wykorzystuje tzw. pojęcie p-wartości, którą w naszym przypadku określa wzór:

$$p = P(|Z| \ge |z_s|). \tag{3.50}$$

Wówczas, jeśli jest spełniona nierówność $p \le \eta$, to hipotezę H_0 odrzucamy na korzyść H_1 z ryzykiem odrzucenia prawidłowego sprawozdania równym

η. Gdy $p > \eta$, to hipotezę H_0 przyjmujemy z ryzykiem akceptacji nieprawidłowego sprawozdania równym κ.

W pewnych sytuacjach rozważa się następującą jednostronną hipotezę:

$$H_0 : \ e_U \leq e_0, \qquad H_1 : \ e_U \geq e_1 > e_0. \tag{3.51}$$

Dodajmy, że formalnie rzecz biorąc, można rozważać także hipotezy:

$$H_0' : \ e_U \geq e_0, \qquad H_1' : \ e_U \leq e_1 < e_0.$$

Jednak zauważmy, że jeśli zdefiniujemy $e_U = y_U - x_U$ zamiast $e_U = x_U - y_U$, to hipotezy H_0 i H_1 są równoważne odpowiednim hipotezom H_0' i H_1'. Zatem w praktyce nie ma potrzeby formułowania hipotez H_0' i H_1'.

Przykładowo niech $e_U = x_U - y_U$ jest różnicą między obserwowaną sumą x_U planowanych potrzeb finansowych danej gminy na dotacje na remont lub rozbudowę infrastruktury komunalnej a faktycznymi potrzebami wyrażającymi się kwotą y_U. Wartość $e_0 > 0$ można traktować jako tolerowane przeszacowanie potrzeb gminy. Z kolei e_1 jest już niedopuszczalnym przekroczeniem deklarowanych potrzeb, co m.in. może świadczyć o niskim poziomie procesu planowania wydatków gminy.

Zauważmy jeszcze, iż w szczególności gdy $e_0 = 0$, to:

$$H_0 : \ e_U \leq 0, \qquad H_1 : \ e_U \geq e_1 > 0. \tag{3.52}$$

Postulowaną wartość η ryzyka odrzucenia wiarygodnego sprawozdania (określającego sumę zapotrzebowań finansowych) określa wyrażenie:

$$\eta = P\left(\tilde{e}_s \geq z_\eta \sqrt{V_s(\tilde{e}_s)} + e_0 | H_0\right), \tag{3.53}$$

co jest równoważne wyrażeniu:

$$\eta = P(Z_s \geq z_\eta | H_0) \tag{3.54}$$

lub

$$\eta = P(Z \geq z_\eta), \qquad Z \sim N(0,1),$$

przy czym Z_s określają wzory (3.57) i (3.44). Korzystając ze wzoru (3.54), ryzyko przyjęcia niewiarygodnego sprawozdania określamy przez:

$$\kappa = P\left(\tilde{e}_s \leq z_\eta \sqrt{V_s(\tilde{e}_s)} + e_0 | H_1\right), \tag{3.55}$$

które to wyrażenie jest równoważne następującym

$$\kappa = P\left(\frac{\tilde{e}_s - e_1}{\sqrt{V_s(\tilde{e}_s)}} \leq z_\eta + \frac{e_0 - e_1}{\sqrt{V_s(\tilde{e}_s)}} | H_1\right),$$

$$\kappa = P\left(Z \leq z_\eta + \frac{e_0 - e_1}{\sqrt{V_s(\tilde{e}_s)}}\right). \tag{3.56}$$

Wprowadźmy upraszczające oznaczenie:

$$V_s(\tilde{e}_s) = \frac{V_s}{n}, \tag{3.57}$$

gdzie $v > 0$.

Wtedy, stąd i ze wzoru (3.56) wynika:

$$\kappa = \phi\left(z_\eta + \frac{e_0 - e_1}{\sqrt{V_s(\tilde{e}_s)}}\right) = \phi\left(z_\eta + \frac{e_0 - e_1}{\sqrt{V_s}}\sqrt{n}\right) = \phi(z_\kappa). \tag{3.58}$$

Z tego wynika, że przy ustalonym ryzyku η i v_s wraz ze wzrostem liczebności próby ryzyko κ maleje do zera.

Proces decyzyjny jest następujący. Jeśli wartość z_s sprawdzianu testu Z_s spełnia nierówność $z_s \geq z_\eta$, to hipotezę H_0 odrzucamy na korzyść H_1 z prawdopodobieństwem popełnienia błędnej decyzji (ryzykiem odrzucenia prawidłowego sprawozdania) równym η. W przeciwnym przypadku, gdy $z_s < z_\eta$, to hipotezę H_0 akceptujemy z prawdopodobieństwem popełnienia błędnej decyzji (ryzykiem przyjęcia nieprawidłowego sprawozdania) równym κ.

Drugi, równoważny powyższemu, sposób podejmowania decyzji wykorzystuje pojęcie p-wartości, którą w aktualnym przypadku określa wzór:

$$p = P(Z \geq z_s), \tag{3.59}$$

gdzie z_s jest zaobserwowaną wartością statystyki testowej (sprawdzianu testu) Z_s. Wówczas, jeśli nierówność $p \leq \eta$ jest spełniona, to hipotezę H_0 odrzucamy na korzyść H_1 z ryzykiem odrzucenia prawidłowego sprawozdania równym η. Gdy $p > \eta$, to hipotezę H_0 przyjmujemy z ryzykiem akceptacji niewiarygodnego sprawozdania równym κ.

Przykład 3.2

Przeprowadzono badanie zasadności planowanych kosztów wykonania projektów naukowo-badawczych zgłaszanych do finansowania przez Narodowy Fundusz Badań Naukowych (NFBN). Spośród $N = 3000$ wniosków zgłoszonych w 2005 roku wybrano za pomocą schematu odrzucającego losowania próbę monetarną o rozmiarze $n = 300$. Losowano z prawdopodobieństwami proporcjonalnymi do proponowanych we wniosku kosztów badań. Średni planowany koszt projektu wynosi $\bar{x}_U = 44000$ zł, a więc wartość globalna wszystkich wniosków wynosi $x_U = N\bar{x}_U = 132000000$.

Zachodzi podejrzenie, że suma (wartość globalna) planowanych kosztów realizacji projektowanych zadań jest znacznie zawyżona. Dysponent funduszu NFBN przeznaczonego na badania jest w stanie tolerować zawyżenie

wartości globalnej kosztów na poziomie $e_0 = 500000$ zł. Wynika to stąd, że przewiduje się zachować rezerwę na poziomie właśnie e_0. Z kolei nie dopuszcza się przekroczenia globalnych kosztów o wartość $e_1 = 2000000$. Tłumaczy to tym, że NFBN jest w stanie brakującą kwotę $e_1 - e_0$ przewyższającą rezerwę e_0 pokryć, zaciągając kredyt. Fachowa kontrola poprawności planowania kosztów jest droga i długotrwała. Zdecydowano się więc na przeprowadzenie audytu statystycznego na podstawie próby losowanej za pomocą monetarnego schematu losowania próby, bowiem jest rzeczą uzasadnioną, iż rozmiar zawyżonego kosztu badania jest proporcjonalny do jego deklarowanego poziomu. Zatem mamy do czynienia z weryfikacją hipotez określonych wzorem (3.51). Ustalono, że ryzyko odrzucenia prawidłowo zaplanowanej wartości globalnej kosztów badań wynosi $\eta = 0.05$.

W wyniku przeprowadzonego audytu zasadności planowanych kosztów zaobserwowanych w próbie wyliczono wartość określonego wzorem (3.24) estymatora Horvitza-Thompsona kosztów realizacji wniosków skorygowanych w wyniku audytu, który wynosi $\bar{y}_{HTs} = 41000$. Zatem wartość określonej wzorem (3.25) oceny błędu audytowego wynosi: $e_{HTs} = N(\bar{x}_U - \bar{y}_{HTs}) = 132000000 - 123000000 = 9000000$. Wyliczona na podstwie wyrażenia (3.30) wartość oceny błędu średniego szacunku (czyli pierwiastak z wariancji) estymatora e_{HTs} wynosi $\sqrt{V_s(e_{HTs})} = 6120000$. To już prowadzi do wyznaczenia zdefiniowanej wzorem (3.33) wartości sprawdzianu testu, która wynosi:

$$u_s = \frac{9000000 - 500000}{6120000} = 1.3889.$$

Przyjmując, że rozmiar próby jest na tyle duży, iż rozkład prawdopodobieństwa sprawdzianu testu można dostatecznie dokładnie aproksymować rozkładem normalnym, wyliczamy na podstawie wzoru (3.59) tzw. $p-$wartość testu, która wynosi $p = P(Z \geq 1.3889) = 1 - \phi(1.3889) = 0.0824 > 0.05 = \eta$. Zatem nie odrzucamy hipotezy $H_0 : \ e_U = e_0 = 500000$. Po to, by tę hipotezę przyjąć, trzeba ocenić ryzyko przyjęcia zawyżonych kosztów realizacji badań, czyli prawdopodobieństwo κ, co czynimy na podstawie wzoru (3.58). Po stosownych obliczeniach mamy:

$$\kappa = \phi\left(z_\eta + \frac{500000 - 2000000}{6120000}\right) = 0.9192,$$

przy czym $z_\eta = 1.6449$, ponieważ $\eta = 0.05 = P(Z \geq 1.6449)$. Stąd wynika, że wyliczony poziom $\kappa = 0.9192$ ryzyka akceptacji zawyżonej wartości globalnej planowanych kosztów jest nie do przyjęcia.

Szczególna postać estymatora $\tilde{e}_s$ i jego wariancji zależy m.in. od schematu losowania próby. Dlatego w przypadku próby prostej losowanej bezzwrotnie wyrażenia (3.20) i (3.58) prowadzą do następujących wzorów:

$$z_\kappa = z_\eta + \frac{e_0 - e_1}{V_s(\tilde{e}_s)},$$

$$n \geq n_{\eta,\kappa} = \left(\frac{1}{N} + \frac{(e_0 - e_1)^2}{(z_\kappa - z_\eta)^2 N^2 v_{*s}(y)} \right)^{-1} \approx \frac{N^2 v_{*s}(y)(z_\kappa - z_\eta)^2}{(e_0 - e_1)^2} = n_a, \tag{3.60}$$

przy czym $\phi(z_\kappa) = \kappa$ natomiast $\phi(z_\eta) = 1 - \eta$. Nieznaną wariancję z populacji $v_*(y)$ zastępujemy jej oszacowaniem pochodzącym z uprzednich badań audytowych kontrolowanej populacji dokumentów lub estymujemy ją na podstawie próby wstępnej.

Przykład 3.3

Audytor kontroluje sumę wartości zakupów sprzętu biurowego uzyskanego w wyniku przeprowadzenia procedur przetargowych. Ma uzasadnione podstawy do wysunięcia przypuszczenia, iż te procedury przetargowe nie były przeprowadzone rzetelnie w tym sensie, że można było dokonać tych zakupów za mniejsze środki. Zatem wysuwa hipotezę, że suma poniesionych wydatków na zakupy jest większa od sumy, która mogła być wydana oszczędniej. Poniesione wydatki wynoszą 1 040 000 i są sumą wartości 3000 faktur. Audytor jest skłonny tolerować nieuzasadnione wydatki do sumy $e_0 = 50000$, co oznacza różnicę między wydatkami poniesionymi a wydatkami oszczędnymi, wynikającymi z uczciwych przetargów. Z kolei niedopuszczalny poziom przekroczenia wydatków ustala się na poziomie $e_1 = 90000$. Na podstawie danych pochodzących z kontroli wcześniej prowadzonych przetargów oszacowano wariancję nieuzasadnionych poziomów przekroczenia racjonalnych wydatków, która wynosi $v_*(e) = 130^2$.

Z formalnego punktu widzenia mamy do czynienia z weryfikacją hipotezy określonej wyrażeniem (3.51). Zatem hipoteza sprawdzana głosi, iż nieuzasadnione wydatki na zakupy sprzętu nie przekraczają dopuszczalnego poziomu, czyli H_0 $e_U \leq e_0 = 50000$. Z kolei hipoteza alternatywna stanowi, iż nieuzasadnione wydatki przekraczają niedopuszczalny poziom, a zatem H_1 $e_U \geq e_1 = 90000$. Audytor ma zamiar zweryfikować próbę dokumentów (faktur), by podjąć decyzję o przyjęciu jednej z dwóch właśnie wyspecyfikowanych hipotez. Ryzyko uznania uczciwie przeprowadzonego przetargu jako nieprawidłowego (poziom istotności testu) ustala na poziomie $\eta = 0.05$. Z kolei ryzyko przyjęcia nieuczciwie przeprowadzonego przetargu ustala na poziomie $\kappa = 0.1$. Do weryfikacji postawionych hipotez użyje testu określonego wzorami (3.18)-(3.21), wyznaczanego na podstawie danych o popełnionych błędach obserwowanych w próbie prostej losowanej z całej populacji dokumentów $N = 3000$. Po to, by spełnić postawione postulaty, wyznaczamy próbę o rozmiarze określonym wzorem (3.60). Po stosownych rachunkach otrzymujemy: $z_\eta = 1.645$, $z_\kappa = -1.282$ i $n_{\eta,\kappa} = 171$.

Przypadek próby warstwowej z proporcjonalnym wariantem ustalania liczebności podprób losowanych bezzwrotnie:

$$n \geq n_{\eta,\kappa} = \left(\frac{1}{N} + \frac{(e_0 - e_1)^2}{(z_\kappa - z_\eta)^2 N^2 v_{*prop}(y)} \right)^{-1} \approx$$
$$\approx \left(\frac{z_\kappa - z_\eta}{e_0 - e_1} \right)^2 N^2 v_{*prop}(y) = n_{ap}, \quad (3.61)$$

przy czym

$$v_{*prop}(y) = \sum_{h=1}^{H} w_h v_{*s_h}(y), \qquad n_{\eta,\kappa} < n_{ap}.$$

Stąd wynika, że wyliczanie niezbędnej liczebności próby na podstawie wzoru przybliżonego prowadzi do nieznacznego zawyżenia jej rozniaru.

Porównując przybliżone wzory (3.60) i (3.61) wyznaczające niezbędną liczebność próby dla odpowiednio losowania bezzwrotnego prostego oraz losowania warstwowego proporcjonalnego, dochodzimy do wniosku, że $n_a \geq n_{ap}$, ponieważ $v_{*prop}(y) \leq v_*(y)$, por. np. Bracha (1996). Stąd wynika, że przy tych samych ryzykach η i κ potrzebny rozmiar próby warstwowej jest nie większy od niezbędnej liczebności próby prostej losowanej z całej populacji.

Przypuśćmy, że mamy ustaloną liczebność próby na poziomie n_d. Wtedy na podstawie wzoru przybliżonego $n_{\eta,\kappa} \approx n_{ap} = n_d$, przy założonym żądanym poziomie ryzyka κ wyznaczamy poziom ryzyka η w następujący sposób. Ze wzoru (3.61) mamy:

$$z_\eta = z_\kappa + \frac{|e_1 - e_0|\sqrt{n_d}}{N\sqrt{v_{*prop}(y)}},$$

$$\eta = P(Z > z_\eta) = P\left(Z > z_\kappa + \frac{|e_1 - e_0|\sqrt{n_d}}{N\sqrt{v_{*prop}(y)}} \right),$$

$$\eta = \eta(n, \kappa) = 1 - \phi\left(\phi^{-1}(\kappa) + \frac{|e_1 - e_0|\sqrt{n_d}}{N\sqrt{v_{*prop}(y)}} \right), \quad (3.62)$$

gdzie $\phi^{-1}(.)$ jest funkcją odwrotną do dystrybuanty $\phi(.)$.

Przykład 3.4

Zwykle sprawozdania z podatków osobistych są obciążone błędem niedoszacowania należności na rzecz danego urzędu skarbowego. Przypuśćmy, że w pewnym mieście jest zarejestrowanych $N = 120000$ podatników. Tę populację osób podzielono na cztery niepuste i rozłączne podpopulacje, zwane warstwami, które oznaczamy przez U_h, $h = 1, 2, 3, 4$, przy czym $\sum_{h=1}^{4} U_h = U$. Romiar warstwy U_h oznaczamy przez N_h, przy czym już teraz zakładamy, iż $N_h > 1$ $h = 1, \ldots, 4$. Warstwa U_1 składa się z kobiet w wieku do 40 lat i liczy $N_1 = 20000$ osób. Warstwę składającą

się z $N_2 = 30000$ kobiet w wieku powyżej 40 lat oznaczymy przez U_2. W końcu warstwy mężczyzn w wieku nieprzekraczającym 40 lat i powyżej tej granicy oznaczamy przez odpowiednio U_3 i U_4. Liczą one odpowiednio $N_3 = 50000$ i $N_4 = 20000$. Frakcje elementów przynależnych do kolejnych warstw są więc następujące: $w_1 = w_4 = 1/6$, $w_2 = 0.25$ i $w_3 = 5/12$. Na podstawie wcześniejszych badań audytowych tak powarstwowanej populacji oszacowano, że odchylenia standardowe skorygowanych wartości podatków w kolejnych warstwach wynoszą odpowiednio: 500, 800, 910 i 685. Suma (wartość globalna) zadeklarowanych przez podatników kwot wynosi $x_U = 240000000$. Wartość globalna należnych urzędowi skarbowemu podatków jest oznaczana przez y_U.

Zadanie polega na ocenie, czy podatnicy nie zaniżają wpłacanych należności na rzecz urzędu skarbowego. Aby przekonać się o tym, przeprowadzono badanie audytowe zadeklarowanych przez podatników należności obserwowanych. Chodzi o to, by urząd skarbowy nie stracił w sumie zbyt wielkich należnych mu kwot pieniędzy. Zatem należy ocenić różnicę między sumą (wartością globalną) należnych urzędowi podatków a wartością globalną zadeklarowanych podatków, co oznaczymy przez $e_U = y_U - x_U$. Urząd skarbowy jest w stanie tolerować niedobór sumy wpłacanych podatków na poziomie $e_0 = 250000$, natomiast absolutnie nie dopuszcza, by niedobór podatków przewyższał poziom $e_1 = 2000000$.

Nasze zadanie polega na takim wyznaczeniu liczebności próby, aby określoną wyrażeniem (3.51) hipotezę weryfikować z ryzykiem $\kappa = 0.05$ akceptacji niedopuszczalnego niedoboru globalnych wpływów z podatków oraz ryzykiem $\eta = 0.1$ nie uznania tego niedoboru podatkowego jako tolerowalnego. W związku z tym na podstawie wzoru (3.61) wyliczamy, że $\phi(-1.6449) = \kappa = 0.05$ i $\phi(1.2816) = 1 - \eta = 0.9$, co prowadzi już do następującego wyniku:

$$n \geq \left(\frac{1}{120000} + \frac{(250000 - 2000000)^2}{(-1.6449 - 1.2816)^2 \cdot 120000^2 \cdot 729.5661^2} \right)^{-1} = 18187.$$

Próba składa się z podpróbek losowanych z warstw proporcjonalnie do ich rozmiarów. Przypomnijmy, że frakcje elementów przynależnych do kolejnych warstw wynoszą $w_1 = w_4 = 1/6$, $w_2 = 1/4$ i $w_3 = 5/12$. Zatem, z poszczególnych warstw zgodnie ze wzorem $n_h = nw_h$, $h = 1, \ldots, 4$, koniecznie pamiętając, że wyniki należy zaokrąglać z nadmiarem do najbliższej liczby naturalnej, czyli $n_1 = n_4 = 18187/6 \approx 3032$, $n_2 = 18187/4 \approx 4547$, $n_3 = 5 \cdot 18187/12 \approx 7578$. W związku z tym trzeba wylosować próbę warstwową o rozmiarze co najmniej równym 18189 elementów po to, aby ryzyka $\eta = 0.1$ i $\kappa = 0.05$ nie zostały przekroczone.

Urząd skarbowy nie stać jednak na sprawdzenie aż 18187 zeznań podatkowych, a tylko $n = 5000$. Wówczas, korzystając ze wzoru (3.62), wyliczamy, że godząc się na zwiększenie ryzyka akceptacji niedopuszczalnego

niedoboru wpływów z podatku do poziomu $\kappa = 0.1$, ryzyko odrzucenia dopuszczalnego niedoboru wpływów z podatku wzrasta do poziomu $\eta = 0.448$.

3.3.2. Ustalone ryzyko odrzucenia wiarygodnego sprawozdania

Często w praktyce jest trudno określić poziom wartości parametru e_1. Wówczas określone wyrażeniem (3.43) hipotezy redukują się do postaci:

$$H_0 : \ |e_U| \leq e_0 \qquad H_1 : \ |e_U| > e_0. \tag{3.63}$$

Wtedy proces podejmowania decyzji jest taki sam, jak to wyżej opisano, lecz nie będzie możliwe wyliczenie wartości ryzyka κ. Zatem, jeśli zachodzi nierówność $p \leq \eta$, to hipotezę H_0 odrzucamy na korzyść H_1 z ryzykiem odrzucenia prawidłowego sprawozdania równym η, przy czym p-wartość jest wyznaczana ze wzoru (3.50). Z kolei gdy $p > \eta$, to nie ma podstaw do odrzucenia hipotezy H_0. Jednak w tym przypadku nie mamy również podstaw do przyjęcia hipotezy H_0, bo nie dało się oszacować ryzyka κ. Równoważne postępowanie polega na odrzuceniu hipotezy H_0, gdy $|z_s| \geq z_{\eta/2}$, gdzie z_s określają wzory (3.44) i (3.57), natomiast $z_{\eta/2}$ wyznacza wzór (3.47).

W szczególności gdy nie można ustalić wartości e_1, to hipotezy, dane wyrażeniem (3.51), upraszczają się do postaci:

$$H_0 : \ e_U \leq e_0, \qquad H_1 : \ e_U > e_0. \tag{3.64}$$

Wtedy nie jest możliwe wyliczenie ryzyka κ i decyzję podejmujemy dalej na podstawie wartości statystyki Z_s zdefiniowanej wzorami (3.57) i (3.44), oraz wyrażeń (3.54) i (3.59) określających odpowiednio wartość krytyczną z_η i p-wartość. Zatem jeśli $z_s \geq z_\eta$, co jest równoważne $p \leq \eta$, to hipotezę H_0 odrzucamy na korzyść H_1 z ryzykiem odrzucenia prawidłowego sprawozdania równym η. W przeciwnym wypadku nie ma podstaw do odrzucenia hipotezy H_0.

3.3.3. Ustalone ryzyko przyjęcia niewiarygodnego sprawozdania

W badaniach audytowych bardziej zwraca się uwagę na to, aby nie poddawać się zbyt dużemu ryzyku przyjęcia niewiarygodnego sprawozdania finansowego. Dlatego przyjmuje się, że to ryzyko, oznaczane przez κ, jest jednocześnie poziomem istotności testu, czyli prawdopodobieństwem popełnienia błędu I rodzaju polegającego na odrzuceniu prawdziwej hipotezy H_0 głoszącej, że odchylenie e_U osiągnie niedopuszczalny poziom e_1 na korzyść alternatywnej H_1 głoszącej, że parametr e_U jest mniejszy od poziomu dopuszczalnego e_0.

Najpierw rozważmy przypadek, że hipoteza sprawdzana H_0 głosi, że odchylenie e_U przekroczy niedopuszczalny poziom e_1 na korzyść alternatywnej

H_1 głoszącej, że parametr e_U nie przekracza poziomu e_1. Sformułowane hipotezy symbolicznie zapisujemy następująco:

$$H_0 : \ e_U \geq e_1, \qquad H_1 : \ e_U < e_1. \tag{3.65}$$

W rozważanym przypadku postulowaną wartość κ ryzyka akceptacji niewiarygodnego sprawozdania określa wyrażenie:

$$\kappa = P\left(\tilde{e}_s \leq z_\kappa \sqrt{V(\tilde{e}_s)} + e_1 | H_0\right), \tag{3.66}$$

co jest równoważne wyrażeniu:

$$\kappa = P(Z_s \leq z_\kappa | H_0) \tag{3.67}$$

lub

$$\kappa = P(Z \leq z_\kappa) = \phi(z_\kappa), \qquad Z \sim N(0,1). \tag{3.68}$$

Proces decyzyjny jest następujący. Jeśli jest spełniona nierówność $z_s \leq z_\kappa$, to hipotezę H_0 odrzucamy na korzyść H_1 z prawdopodobieństwem popełnienia błędnej decyzji (ryzykiem przyjęcia nieprawidłowego sprawozdania) równym κ. W przeciwnym przypadku, gdy $z_s > z_\kappa$, nie ma podstaw do odrzucenia hipotezy H_0.

Drugi, równoważny powyższemu, sposób podejmowania decyzji wykorzystuje wyliczoną p-wartość, którą w aktualnym przypadku określa wzór:

$$p = P(Z \leq z_s), \tag{3.69}$$

gdzie z_s jest zaobserwowaną wartością statystyki testowej (sprawdzianu testu) Z_s. Wówczas, jeśli jest spełniona nierówność $p \leq \kappa$, to hipotezę H_0 odrzucamy na korzyść H_1 z ryzykiem przyjęcia nieprawidłowego sprawozdania równym κ. Gdy $p > \kappa$, to nie ma podstaw do odrzucenia hipotezy H_0.

Przykład 3.5

Niniejsze rozważania stanowią analizę zagadnienia szczególnego w stosunku do problemu analizowanego w przykładzie 3.4 problemu weryfikacji istotności przekroczenia nietolerowalnego poziomu globalnego niedoboru wpływów z podatków osobistych. Do tego celu użyto poddanych kontroli audytowej sprawozdań podatkowych osób dobranych do próby warstwowej z proporcjonalnym wariantem ustalania liczebności próbek losowanych z poszczególnych warstw. Rozmiar populacji podatników wynosi 120 000.

Zakładamy, że urząd skarbowy dopuszcza ryzyko akceptacji sumy zaniżonych wpływów z podatków na poziomie prawdopodobieństwa $\kappa = 0.05$. To już prowadzi do następującego skonkretyzowania postaci hipotez określonych wzorem (3.65):

$$H_0 : \ e_U \geq e_1 = 2000000, \qquad H_1 : \ e_U \leq e_1. \tag{3.70}$$

Zapisana hipoteza sprawdzana głosi, iż wartość niedoboru wpływów podatkowych jest nie mniejsza od niedopuszczalnego (nietolerowalnego) poziomu $e_1 = 2000000$, a hipoteza alternatywna, że ten niedobór jest mniejszy (dopuszczalny). Przypuśćmy, że w wyniku dalej opisanej procedury testowej odrzucimy hipotezę H_0 na korzyść H_1, co będzie oznaczało, iż uznajemy, że niedobór wpływów z podatków nie przekroczył niedopuszczalnego poziomu e_1, przy czym ta decyzja jest obarczona ryzykiem κ.

Skontrolowano zeznania podatkowe w próbie składającej się z $n = 1200$ podatników wylosowanych bezzwrotnie z poszczególnych warstw w ilościach proporcjonalnych do ich rozmiarów. W związku z tym z kolejnych warstw wylosowano bezzwrotnie próby proste o liczebnościach odpowiednio $n_1 = 200$, $n_2 = 300$, $n_3 = 500$ i $n_4 = 200$. Ponadto, zgodnie ze wzorem (3.35) wyliczono oceny wartości średnich skorygowanych podatków w kolejnych warstwach, co dało wyniki: $\bar{y}_{s_1} = 1800$, $\bar{y}_{s_2} = 1200$, $\bar{y}_{s_3} = 2500$,$\bar{y}_{s_4} = 2200$. Te wyniki i wzory (3.34) i (3.35) prowadzą do ocen wartości średniej i globalnej należnych wpływów z podatków wynoszących odpowiednio $\bar{y}_{w,s} = 2008.33$ i $y_{w,s} = 241000000$. Suma zadeklarowanych wpływów z podatków wynosi 240 000 000. To oraz wzór (3.36) dają już ocenę niedoboru globalnych wpływów z podatków wynoszącego $e_{w,s} = 241000000 - 240000000 = 1000000$. Zgodnie ze wzorem (3.39) wyliczono nieobciążone oceny wariancji, które w kolejnych warstwach wynoszą $v_{*s_1}(y) = 15610$, $v_{*s_2}(y) = 48990$, $v_{*s_3}(y) = 24220$, $v_{*s_4}(y) = 31110$. Te dane oraz wzór (3.41) dają ocenę błędu średniego szacunku estymatora $e_{w,s}$, która jest równa $\sqrt{V_s(e_{w,s})} = 598375.55$. W końcu wartość określonego wzorem (3.42) sprawdzianu testu wynosi: $z_{w,s} = -1.6712$. Z konstrukcji postawionych hipotez wynika, że hipotezę H_0 odrzucamy, gdy wartość sprawdzianu jest istotnie mniejsza od zera przy ustalonym poziomie istotności równym poziomowi ryzyka κ oznaczającego akceptację w naszym przypadku stwierdzonej wartości globalnej niedoboru wpływów z podatku jako tolerowalnej. Ze wzorów (3.68) i (3.69) mamy: $p = 0.0473 = \phi(z_{w,s}) = \phi(-1.6712)$ oraz $z_\kappa = -1.6449$, ponieważ $\kappa = 0.05 = \phi(-1.6449)$. Stąd wynika że: $p = 0.0473 < 0.05 = \kappa$, co jest równoważne faktowi zachodzenia nierówności $z_{w,s} = -1.6712 < z_\kappa = -1.6449$. W związku z tym hipotezę H_0 odrzucamy na korzyść H_1 z prawdopodobieństwem popełnienia błędnej decyzji równym $\kappa = 0.05$, co oznacza, iż pojawiający się niedobór podatkowy uznajemy za nieistotny (tolerowalny) z ryzykiem podjęcia błędnej decyzji równym $\kappa = 0.05$.

Teraz rozważmy przypadek, gdy audytor ustala wartość krytyczną e_κ rozmiaru poziomu globalnego błędów, którego wynikiem jest możliwość podjęcia decyzji o przyjęciu sprawozdania księgowego, bądź jego odrzuceniu. Z formalnego punktu widzenia dalej mamy do czynienia z weryfikacją hipotez określonych wyrażeniem (3.65). Wówczas z wyrażenia (3.66) wynika,

że

$$e_\kappa = e_1 + z_\kappa \sqrt{V(\tilde{e}_s)}. \tag{3.71}$$

Przyjmując, że $V(\tilde{e}_s) = v/n$ mamy:

$$e_\kappa = \frac{e_1 + z_\kappa \sqrt{v}}{\sqrt{n}}.$$

Stąd już mamy:

$$n \geq \left[\frac{(e_1 + z_\kappa \sqrt{v})^2}{e_\kappa^2} \right] + 1 = n_o. \tag{3.72}$$

Zatem po to, by podjąć decyzję o ewentualnym przyjęciu sprawozdania księgowego z ryzykiem na poziomie κ i przy wartości krytycznej globalnego błędu księgowego ustalonej na poziomie e_κ, należy wylosować próbę o rozmiarze równym co najmniej n_o dokumentów. Wtedy, jeśli $\tilde{e}_s \leq e_\kappa$, to odrzucamy określoną wzorem (3.65) hipotezę sprawdzaną H_0 z prawdopodobieństwem popełnienia błędnej decyzji równym κ. W kontekście rozważanego badania audytowego odrzucenie hipotezy H_0 jest równoważne przyjęciu sprawozdania księgowego z ryzykiem nieprawidłowej decyzji na poziomie κ. W przeciwnym przypadku, gdy $\tilde{e}_s > e_\kappa$, nie ma podstaw do odrzucenia hipotezy H_0. W tej sytuacji z punktu widzenia naszych rozważań nie ma podstaw do odrzucenia stwierdzenia, że sprawozdanie księgowe jest nie do przyjęcia, czyli że sprawozdanie nie jest prawidłowe.

Przykład 3.6

Audytor ma sprawdzić, czy wartość globalna błędu księgowego przekracza poziom niedopuszczalny równy 10 000. Zbiór (populacja) dokumentów wynosi 42 000. Audytor zakłada, że ryzyko przyjęcia nieprawidłowego sprawozdania wynosi 0.2. Ponadto audytor ustala, że sprawozdanie przyjmie, gdy zaobserwowana w próbie ocena globalnego błędu księgowego nie przekroczy poziomu 100. W celu oszacowania wartości globalnej błędu księgowego ma zamiar wylosować bezzwrotnie próbę prostą. Zadanie polega na wyznaczeniu rozmiaru tej próby. Na podstawie wyników z wcześniej prowadzonych badań audytowych podobnych populacji dokumentów oszacowano, iż odchylenie standardowe błędów księgowych wynosi 9 zł.

Korzystając ze wzorów (3.71) i (3.19), mamy:

$$e_\kappa = e_1 + z_\kappa \sqrt{\frac{N(N-n)}{n} v_{*U}(e)}.$$

Stąd już mamy:

$$n \geq \left[N \left(1 + \frac{(e_\kappa - e_1)^2}{z_\kappa^2 N v_{*U}(e)} \right)^{-1} \right] + 1 = n_o. \tag{3.73}$$

Przyjmując teraz, że $N = 42000$, $e_1 = 10000$, $v_{*U}(e) = 81$, $\kappa = 0.02$ i $e_\kappa = 100$, wyliczamy, że $n_o = 1008$. Zatem krótko mówiąc, po to, aby podjąć decyzję o ewentualnym przyjęciu sprawozdania z ryzykiem wynoszącym 0.2, należy wylosować bezzwrotnie 1008 dokumentów z populacji w celu ich sprawdzenia.

BIBLIOGRAFIA

Ardilly P., Tille Y. (2006): *Sampling Methods. Excercises and Solutions.* Springer, New York.

Berger Y.G. (1998): *Rate of Convergence to Normal Distribution for the Horvitz-Thompson Estimator.* ,,Journal of Statistical Planning and Inference", Vol. 67, s. 209-226.

Bracha Cz. (1996): *Teoretyczne podstawy metody reprezentacyjnej.* Wydawnictwo Naukowe PWN, Warszawa.

Bracha Cz. (1998): *Metoda reprezentacyjna w badaniu opinii publicznej i marketingu.* Efekt, Warszawa.

Brewer K.R.W., Hanif M. (1983): *Sampling with Unequal Probabilities.* Springer Verlag, New York-Heidelberg-Berlin 1983.

Fuller W.A. (2009). *Sampling Statistics.* Wiley and Sons, Hoboken, New Jersey.

Gerstenkorn T., Śródka T. (1974). *Kombinatoryka i rachunek prawdopodobieństwa.* PWN, Warszawa.

Guy D.M., Carmichael D. R., Whittington R. (2002). *Audit Sampling. An Introduction.* John Wiley & Sons, New Jersey.

Guilford I.P. (1960). *Podstawowe metody statystyczne w psychologii i pedagogice.* PWN, Warszawa.

Hájek J. (1964): *Asymptotic Theory of Rejective Sampling with Varying Probabilities from a Finite Population.* ,,Annals of Mathematical Statistics", No. 35, s. 1491-1523.

Hájek J. (1981): *Sampling from a Finite Population.* Marcel Dekker, New York – Bassel.

Hartley H.O., Rao J.N.K. (1962): *Sampling with Unequal Probabilities and without Replacement.* ,,Annals of Mathematical Statistics", Vol. 33, s. 350-374.

Hołda A., Pociecha J. (2004): *Rewizja finansowa.* Wydawnictwo Akademii Ekonomicznej, Kraków.

Hołda A., Pociecha J. (2009): *Probabilistyczne metody badania sprawozdań finansowych.* Wydawnictwo Uniwersytetu Ekonomicznego, Kraków.

Horvitz D.G., Thompson D.J. (1952): *A Generalization of Sampling without Replacement from Finite Universe.* ,,Journal of the American Statistical Association", Vol. 47, s. 663-685.

Jakubowski J., Sztencel R. (2004): *Wstęp do teorii prawdopodobieństwa.* SCRIPT, Warszawa.

Karliński W. (2005): *Dobór próby w audycie.* Instytut Rachunkowości i Podatków, Warszawa.

Kiger J.E., Scheiner J.H. (1994). *Auditing.* Houghton Mifflin Company, Boston-Toronto-Geneva.

Klima D. (2005): *Statystyka dla audytorów w przykładach. Biblioteka Audytora 2.* InfoAudit, Warszawa.

Krzyśko M. (2000): *Wykłady z teorii prawdopodobieństwa.* WNT, Warszawa.

Muraszew M. (2011): *Algorytm ustalania arbitralnej aproksymacji całkowitoliczbowej wielowymiarowego wektora danych rzeczywistych w kontekście badań reprezentacyjnych.* ,,Wiadomości Statystyczne", nr 2, 1-8.

Owczarek K. (1996): *Badanie próbki jako narzędzie wspomagające rewizję. Próbkowanie według zmiennych - badanie według hipotezy biegłego.* ,,Rachunkowość", nr 9, s. 453-460.

Panel on Nonstandard Mixtures of Distributions (PNMDZ) (1989). *Statistical models and analysis in auditing.* ,,Statistical Science", No.4, 2-33.

Pawłowski Z. (1972): *Wstęp do statystycznej metody reprezentacyjnej.* PWN, Warszawa.

Pfaff J. (2008): *Wpływ rewizji finansowej na wiarygodność sprawozdania finansowego.* Wydawnictwo Akademii Ekonomicznej, Katowice.

Rao J.N.K., Hartley H.O., Cochran W.G. (1962): *On a Simple Procedure of Unequal Probability Sampling without Replacement.* ,,Journal of the Royal Statistical Association", Vol. 24, No. 2, s. 482-491.

Rao, J.N.K. (1965): *On Two Simple Schemes of Unequal Probability Sampling Without Replacement.* ,,Journal of the Indian Statistical Association" 3, 173-180.

Ryan T.P. (2013): *Sample Size Determination and Power.* John Wiley and Sons, Hoboken, New Jersey.

Sampford M.R. (1967): *On Sampling Without Replacement with Unequal Probabilities of Selection.* ,,Biometrika" 54, s. 499-513.

Steczkowski J. (1995): *Metoda reprezentacyjna w badaniu zjawisk ekonomiczno-spoecznych.* Wydawnictwo Naukowe PWN, Warszawa.

Tillé Y. (2006): *Sampling algorithms.* Springer.

Wywiał J. L. (1992): *Statystyczna metoda reprezentacyjna w badaniach ekonomicznych (Optymalizacja badań próbkowych).* Wydawnictwo Akademii Ekonomicznej, Katowice.

Wywiał J. L. (2010). *Wprowadzenie do metody reprezentacyjnej.* Wydawnictwo Uniwersytetu Ekonomicznego, Katowice.

Wywiał J. L. (2012): *On Limit Distribution of Horvitz-Thompson Statistic under the Rejective Sampling.* ,,Studia Ekonomiczne" Uniwersytetu Ekonomicznego w Katowicach, nr 120, s. 84-96.

PYTANIA I ZADANIA

1. Podać definicję oraz interpretację błędu średniego szacunku.
2. Określić względny średni błąd szacunku.
3. Podać określenie obciążenia estymatora.
4. Przedstawić dekompozycję błędu średniokwadratowego estymacji.
5. Podać określenia planu oraz schematu losowania próby.
6. Określić prawdopodobieństwa inkluzji rzędów pierwszego i drugiego.
7. Opisać plan i schemat bezzwrotnego losowania próby prostej.
8. Opisać plan i schemat losowania próby warstwowej.
9. Na czym polega proporcjonalny wariant losowania próby warstwowej?
10. Opisać plan losowania systematycznego próby.
11. Wyjaśnić termin losowania monetarnego (dolarowego, walutowego) próby.
12. Opisać plan losowania Stampforda.
13. Opisać plan losowania próby odrzucającego powtórzenia.
14. Podać własności średniej z próby prostej losowanej bezzwrotnie.
15. Podać własności średniej z próby warstwowej.
16. Opisać własności estymatora Horvitza-Thompsona.
17. Określić poziom ufności estymatora przedziałowego.
18. Określić ryzyko badania audytowego oraz jego czynniki.
19. Na czym polegają błędy pierwszego i drugiego rodzaju podczas weryfikacji hipotez statystycznych?
20. Zdefiniować poziom istotności testu oraz jego moc.
21. Na czym polega ryzyko nieprawidłowej akceptacji?
22. Na czym polega ryzyko nieprawidłowego odrzucenia?
23. Przedstawić związki między ryzykiem nieprawidłowej akceptacji albo nieprawidłowego odrzucenia, a poziomem istotności testu lub prawdopodobieństwem popełnienia błędu drugiego rodzaju.
24. Rozkład błędów w zapisach dokumentów jest normalny z nieznaną średnią i odchyleniem standardowym wynoszącym 100. Dopuszczalny błąd średni wynosi 1000 zł, a dyskwalifikujący 3000 zł. Ryzyko przyjęcia nieprawidłowego sprawozdania ustalono na poziomie 0.2, a ryzyko odrzucenia poprawnego sprawozdania na poziomie 0.1. W celu podjęcia decyzji

o przyjęciu lub odrzuceniu sprawozdania będzie losowana bezzwrotnie próba prosta z populacji liczącej 40 000 dokumentów. Korzystając z arkusza kalkulacyjnego Excel, wyznaczyć rozmiar tej próby potrzebny do podjęcia decyzji.

25. Populacja liczy 30 000 faktur, których wartości będą sprawdzane. Z wcześniejszych badań wynika, że odchylenie standardowe błędów wynosi 190 zł. Wyznaczyć niezbędną liczebność próby prostej losowanej bezzwrotnie zakładając, że błąd średni szacunku wartości globalnej faktur po dokonaniu korekt nie przekracza poziomu 200 000 zł.
26. Podjęto rozważany w podrozdziale 2.2.1 problem wydania decyzji o skuteczności systemu kontroli wewnętrznej przyjmując, że proporcja dopuszczalna nieprawidłowości nie przekracza poziomu 0.08, a proporcja dyskwalifikująca jest nie mniejsza od 0.12. Ryzyko dezaprobaty skutecznego systemu kontroli ustala się na poziomie 0.08, natomiast ryzyko aprobaty nieskutecznego systemu kontroli na poziomie 0.05. Potrzebna próba prosta poddawanych kontroli pozycji księgowych będzie losowana bezzwrotnie z populacji o rozmiarze 420 000 dokumentów. Korzystając z przybliżenia normalnego odpowiednich statystyk, wyznaczyć niezbędną liczebność próby oraz wartość krytyczną stosownego testu statystyczego pozwalającego podjąć właściwą decyzję o skuteczności systemu kontroli.
27. Audytor sprawdza jakość 400 dokumentów, które już przeszły stosowne procedury kontroli wewnętrznej. Ta kontrola zostanie uznana jako skuteczna jeśli w badanej zbiorowości dokumentów nie będzie więcej niż 4 dokumenty błędne, a liczbę 6 dokumentów już uważa się za niedopuszczalną. Audytor zakłada ryzyko dezaprobaty skutecznego systemu kontroli ustala się na poziomie 0.15, natomiast ryzyko aprobaty nieskutecznego systemu kontroli na poziomie 0.05. Korzystając z programu zamieszczonego w przykładzie 2.2, możemy wyznaczyć niezbędny rozmiar próby prostej losowanej bezzwrotnie z populacji dokumentów oraz wartość krytyczną testu pozwalającą na podjęcie stosownej decyzji tak by spełnione były określone postulaty co do ryzyka wydania nieprawidłowego werdyktu o jakości kontroli wewnętrznej.
28. Rozważmy ponownie określony w uprzednim zadaniu problem wyznaczania niezbędnej liczebności próby oraz wartości krytycznej testu lecz dla populacji liczącej 10 000 dokumentów. Teraz dopuszczalna proporcja nieprawidłowych dokumentów wynosi 0.06, natomiast dyskwalifikujący poziom tej proporcji jest równy 0.12. Ryzyko dezaprobaty skutecznego systemu kontroli ustala się na poziomie 0.1, natomiast ryzyko aprobaty nieskutecznego systemu kontroli na poziomie 0.04. Korzystając z programów zamieszczonych w przykładach 2.2 i 2.3, wyznaczymy niezbędny rozmiar próby prostej losowanej bezzwrotnie z populacji dokumentów oraz wartość krytyczną testu pozwalającą na podjęcie stosownej decyzji,

tak by spełnione były określone postulaty co do ryzyk wydania nieprawidłowego osądu jakości kontroli wewnętrznej.

29. Ponownie bieżemy pod uwagę określony w uprzednim zadaniu problem wyznaczania niezbędnej liczebności próby oraz wartości krytycznej testu lecz dla populacji liczącej 190 000 dokumentów. Dopuszczalna proporcja nieprawidłowych dokumentów wynosi 0.01, natomiast dyskwalifikujący poziom tej proporcji jest równy 0.03. Ryzyko dezaprobaty skutecznego systemu kontroli ustala się na poziomie 0.09, natomiast ryzyko aprobaty nieskutecznego systemu kontroli na poziomie 0.03. Program zamieszczony w przykładzie 2.4 wyznaczy niezbędny rozmiar próby prostej losowanej bezzwrotnie z populacji dokumentów oraz wartość krytyczną testu pozwalającą na podjęcie właściwej decyzji tak by były zachowane określone postulaty co do ryzyk wydania nieprawidłowego sądu o jakości kontroli wewnętrznej.
30. Audytor sprawdza jakość 1400 dokumentów na podstawie już wylosowanej bezzwrotnie próby prostej o rozmiarze 60. Dopuszczalna proporcja nieprawidłowych dokumentów wynosi 0.04. Ryzyko aprobaty nieskutecznego systemu kontroli ustala się na poziomie 0.05. Korzystając z procedur używanych w przykładach 2.5-2.7 należy wyznaczyć wartość krytyczną testu pozwalającą na podjęcie stosownej decyzji o jakości systemu kontroli tak by spełniony był postulat co do ryzyka wydania nieprawidłowego werdyktu.
31. Audytor sprawdza jakość 1000 dokumentów. W tym celu ma zamiar wylosować bezzwrotnie próbę prostą. Ponadto, audytor ustala, że jeśli zaobserwuje w tej próbie nie więcej niż 9 dokumntów z nieprawidłowościami, to uzna system kontroli wewnętrznej za skuteczny z ryzykiem wynoszącym 0.08. Dopuszczalna proporcja nieprawidłowych dokumentów wynosi 0.1. Korzystając z procedur używanych w przykładzie 2.9 należy wyznaczyć niezbędną liczebność próby pozwalającą na podjęcie stosownej decyzji o jakości systemu kontroli tak by spełnione były określone wcześniej postulaty.
32. Korzystając z programu dydaktycznego opisanego w przykładzie 1.19 przeprowadzić weryfikację hipotez o wartości średniej zmiennej losowej o rokładzie normalnym. Rozważyć różne warianty wartości parametrów zmiennej losowej, rozmiaru próby prostej losowej oraz wartości poziomu istotności testu.

ISBN 978-83-7875-160-1