METODY REPREZENTACYJNE WSPOMAGANE KOMPUTEROWO W BADANICH AUDYTORSKICH

ANDRZEJ OBECNY

WSTĘP

Celem niniejszego artykułu[1] jest wprowadzenie Czytelnika w zagadnienia dotyczące wykorzystania metod reprezentacyjnych w audycie z zastosowaniem techniki komputerowej. Dla osiągnięcia zamierzonego celu posłużymy się przykładami, w których Czytelnik powinien dostrzec zalety użycia narzędzia w postaci odpowiednio zaprojektowanego programu komputerowego. Liczymy również, że uda się przybliżyć problematykę metod reprezentacyjnych tym osobom, które o teorii tej słyszały dotąd niewiele.

Zamiarem autora było, by treść i forma artykułu była od początku do końca przystępna i jasna oraz – w efekcie końcowym - zachęciła do głębszego zainteresowania się poruszonym w artykule tematem. Świadomie zrezygnowano z posługiwania się symboliką matematyczną stosowaną w tej teorii. Nie zaprezentowano żadnego wzoru, czy ścisłej definicji. W to miejsce pojawił się język opisowy. Pojęcia, o których będziemy mówić są jedynie małym wycinkiem rozległej i niełatwej teorii. Osoby przeprowadzające badanie finansowe teorii tej znać nie muszą, dobrze byłoby jednak, by znały chociaż jej logikę i by potrafiły zastosować ją w praktyce. Spróbujemy omówić tylko te elementy metod reprezentacyjnych, które są niezbędne dla właściwego zrozumienia dwóch konkretnych zadań, jakie będziemy chcieli rozwiązać posługując się programem komputerowym. Postaramy się pokazać jak w rozwiązaniu tych zadań przydatne może być tego typu narzędzie. Nim do tego przejdziemy wyjaśnienia wymaga kilka niezbędnych pojęć.

BADANIA PEŁNE I NIEPEŁNE

Jest kilka powodów, dla których metoda reprezentacyjna jest warta zastosowania. Wymieńmy dwa – z punktu widzenia firmy audytorskie – najważniejsze.

Po pierwsze badanie reprezentacyjne pozwala prędzej uzyskać obraz interesującej nas zbiorowości, niżby to stać się mogło w przypadku badania pełnego. Czas badania skrócić się może nawet dziesięciokrotnie w stosunku do czasu, jaki wymagałoby badanie stuprocentowe. Skoro można zaoszczędzić na czasie, tym samym oszczędzamy na kosztach badania. Najczęściej koszt badania jest rosnącą funkcją ilości badanych elementów.

Drugim argumentem jest obiektywizm oceny poprzez zastosowanie w badaniu rewizyjnym naukowej metody. Fakt, którego nie można nie brać pod uwagę oceniając markę firmy audytorskiej.

[1] Recenzent: prof. dr hab. Zygmunt Przybycin, AE Katowice.

METODY REPREZENTACYJNE

Metody reprezentacyjne są działem statystyki matematycznej, w której obszarem badań są sposoby wyboru losowych prób ze skończonej zbiorowości statystycznej oraz metody wnioskowania o własnościach całej zbiorowości na podstawie poczynionych w wyniku losowania obserwacji.

Przenosząc zaraz ową definicje na grunt badań audytorskich widzimy pole do zastosowań tejże dyscypliny naukowej. Mamy bowiem - badając wiarygodność oraz zgodność sprawozdania finansowego - do czynienia z określonym zbiorem dokumentów stanowiącym właśnie zbiorowość statystyczną. W praktyce liczebność tego zbioru uniemożliwia przeprowadzenie pełnego badania, zatem odpowiednie procedury rewizyjne przeprowadza się do wybranych elementów badanego zbioru. Na podstawie oceny tychże elementów wnioskuje się ostatecznie na temat całej zbiorowości.

Skąd jednak pewność, że owe uogólnienie oddaje prawdziwy obraz całej zbiorowości? Otóż korzystając właśnie z metod reprezentacyjnych może taką pewność uzyskać, jednak tylko w określonym prawdopodobieństwem zakresie. Wyjaśnijmy na początek termin **próba losowa.**

PRÓBA LOSOWA

Otóż próba losowa**,** to zbiór otrzymany w wyniku stosowania pewnego mechanizmu losowego. W technice komputerowej tym mechanizmem jest obecnie tzw. generator liczb losowych. Dawniej posługiwano się w tym celu specjalnym układem liczb zwanym tablicą liczb losowych. Próba losowa, to taka część zbiorowości generalnej, w której pojawienie się w niej każdego z elementów zbiorowości jest określone przez dodatnie prawdopodobieństwo. Tak więc każdy z elementów zbiorowości może się do próby dostać, choć nie zawsze musi to być szansa równa dla wszystkich elementów.

Od próby losowej wymagamy także, by była reprezentatywna. Wiemy bowiem, że każda wybrana próba bardziej lub mniej różni się jednak od badanej zbiorowości. Zasadnicze w tym momencie pytanie brzmi: na ile ta różnica jest znacząca? Jeśli przyjąć, że różnica ta jest dla badanej zbiorowości nieistotna, to mówimy wówczas, że próbka jest **reprezentatywna**. Metody reprezentacyjne gwarantują nam, że z dowolnie małym zadanym przez nas ryzykiem popełnienia błędu osiągniemy ową nieistotność.

Użyliśmy wyżej mało precyzyjnego pojęcia nieistotność. Czy da się ją wyrazić jakąś konkretną obiektywną miarą? W tym miejscu zaczynamy dotykać już samej istoty metod reprezentacyjnych. Można mianowicie postawić pytanie, czy jest możliwy taki dobór próbki, aby owa nieistotność - którą nazywać będziemy od tej chwili **dopuszczalnym błędem** – była zawsze zagwarantowana? Innymi słowy, czy można uzyskać stuprocentową pewność, że każda próbka, którą otrzymamy da nam szacunek badanej wartości z co najwyżej zadanym przez nas odchyleniem (błędem)? Otóż odpowiedź na to pytanie brzmi: nie! Nie możemy. Nie jest możliwe uzyskanie takiej pewności w oparciu o jedną konkretną próbę. (Ani zresztą, o każdą inną próbkę z osobna.) Możemy jednak rozmiar możliwego błędu określić, w ten oto sposób, iż jest z góry zadane prawdopodobieństwo wylosowania próby dającej większy błąd. W tym miejscu Czytelnik musi przyjąć za fakt, iż istnieje matematyczna zależność między dokładnością szacunku, jaką możemy sobie zadać - czyli dopuszczalnym błędem - a prawdopodobieństwem wylosowania próby dającej większy błąd. Tym samym potrafimy

określić ryzyko wydania błędnej oceny, co do szacowanego parametru w porównaniu z jego faktyczną wartością.

Otóż prawdopodobieństwo to określa się terminem **ryzyko błędu** i podawane jest w postaci procentowej. Z reguły waha się ono między 0,5%, a 10%. Używa się także zamiennie terminu, który jest przeciwieństwem ryzyka, mianowicie **współczynnik ufności.** Wyraża się on liczbą z zakresu od zera do jeden. W praktyce spotyka się wartości między 0,95 a 99,9. Podobnie zamiennie używa się terminu **precyzja** zamiast dopuszczalnego błędu. Precyzję z reguły podaje się w postaci wartości bezwzględnej, rzadziej procentowej. Przy czym wartość procentowa – jeśli jest użyta - odnosi się do szacowanego parametru, którego prawdziwej wartości nie znamy, a nie do podanej przez firmę wartości księgowej.

PRZEDZIAŁ UFNOŚCI

Innym ważnym pojęciem jest **przedział ufności**. Jest to losowy przedział liczbowy, który z zadanym prawdopodobieństwem pokrywa nieznaną wartość, którą szacujemy na podstawie próby. Dla jego wyznaczenia konieczna jest znajomość przyjętego ryzyka oraz precyzji. Konkretny przedział powstanie w oparciu o konkretną wylosowaną próbkę. Gdyby próbka była inna, przedział byłby najprawdopodobniej również inny. Lecz rzecz w tym, iż tworząc takich przedziałów – w serii testów - np. 100 (zawsze na próbkach o tej samej liczebności!), średnio tylko w tylu przypadkach, ile wynosi ryzyko błędu przedziały te nie pokryłyby szacowanego parametru. Można zatem powiedzieć, że wszystkie przedziały z prawdopodobieństwem równym współczynnikowi ufności pokrywają prawdziwą wartość szacowanej cechy.

Konkludując, stwierdzić możemy, że badania reprezentacyjne gwarantują odpowiednią dokładność wyników. Stosując właściwy sposób losowania próby oraz dobierając właściwą liczbę elementów do próby, mamy praktyczną pewność, iż błąd szacunku interesującego nas parametru nie przekroczy z góry zadanego dopuszczalnego poziomu. Oznacza to tyle, że dzięki losowemu charakterowi próby dokładność wnioskowania jest pod kontrolą - można ją z góry ustalać.

NIEZBĘDNA LICZEBNOŚĆ PRÓBY I BADANIE WSTĘPNE

O rozmiarze (liczebności) próby, którą należy pobrać z badanej zbiorowości decydują: dopuszczalny błąd szacunku, ryzyko popełnienia błędu w badaniu, a także – o czym zaraz powiemy – schemat losowania. Zbiorowość cechują pewne parametry takie jak suma, wartość średnia, odchylenie od średniej wyrażane np. poprzez odchylenie średnie oraz wiele innych. Mówią one o „kształcie” rozkładu badanej zmiennej. Chcąc wyliczyć wymaganą liczebność próbki trzeba znać pewne parametry zbiorowości generalnej. Na ogół ich nie znamy, pozostaje zatem je oszacować na podstawie jakiegoś **badania wstępnego**. Badanie to nie jest jeszcze badaniem właściwym (nie będziemy się przyglądać dokumentom w tym losowaniu wskazanym). Da nam ono jednak pojęcie o całej zbiorowości. Próbka wstępna, zwana też pilotową nie jest wielka (najczęściej liczy ona od 30 to 50 elementów), zdarzyć się więc może, że parametr na jej podstawie szacowany będzie daleki (oczywiście względnie) od prawdziwego. Nie przeszkadza to jednak w ustaleniu liczebności próbki, gdyż tę będziemy – w razie potrzeby – powiększać w miarę jak nasze szacunki będą ulegać korektom. (Dokładne wyjaśnienie tej zależności wykracza poza ramy tego artykułu).

POLE BADAŃ DLA METOD REPREZENTACYJNYCH

Jak powiedziano na wstępie, z odpowiednio dobranej próby losowej można wyciągnąć wnioski o całej zbiorowości. Pytanie, jakiego rodzaju informacje nas interesują. Co chcemy wiedzieć. W metodach reprezentacyjnych rozróżnia się dwa podstawowe rodzaje badań. W jednym przypadku może nas interesować aspekt jakościowy (test istotności) np. poprawnie pod względem formalnym zaksięgowana faktura. Innym razem interesować nas może aspekt ilościowy (test wiarygodności) np. kwota należności na fakturze. Z kolei w tym przypadku całą zbiorowość możemy opisać (badać) na wiele sposobów, w zależności od tego, co chcemy wiedzieć o tej zbiorowości. Może nas interesować suma wszystkich należności, na którą składają się poszczególne pozycje, lub znać ich średnią wartość. Mówimy wtedy o badaniu zbiorowości ze względu na badaną cechę. Jeśli interesuje nas liczba elementów z określoną cechą np. saldem mniejszym od zadanej wartości, to jest to test istotności. Możemy być zainteresowani poznaniem rozkładu poszczególnych wartości w zbiorze, tj. opisem jak często pojawiają się konkretne wartości w całym badanym zbiorze (rozkład zmiennej). To jest pierwsza grupa cech pod względem, których możemy badać całą zbiorowość. Jest ona podstawowym celem badania reprezentacyjnego. Drugą grupę stanowią takie parametry jak: odchylenie standardowe czy współczynnik zmienności określające stopień zmienności. Inną kategorią są miary asymetrii, a jeszcze inne badają stopień koncentracji w zbiorze. Możliwości i sposoby opisu badanej zbiorowości jest wiele.

PLANOWANIE ZADANIA REPREZENTACYJNEGO

Zanim przystąpimy do przeprowadzenia badania należy określić przede wszystkim cel badania. Następnie należy określić zbiorowość generalną, czyli zbiór danych, z których zostanie pobrana próba. Później należy przyjąć założenia, co do ryzyka badania oraz dopuszczalnego błędu, a także określić schemat losowania. Pozostaje w tej chwili omówić ten ostatni element.

SCHEMAT LOSOWANIA I JEGO RODAJE

Schemat losowania to procedura tworzenia próbki reprezentatywnej określona konkretnymi wymogami. Jest kilka podstawowych schematów. Spośród nich wymieńmy dwa, które znajdą zastosowanie w naszych przykładach. Są nimi:

- **Losowanie proste zależne**. – Jest to losowanie, w którym wybrany do próby element nie może już zostać wybrany ponownie oraz takie, że prawdopodobieństwo wylosowania każdego z elementów jest takie samo i jest ono dodatnie. Ponadto, jeżeli jednostka jest losowana z całej zbiorowości oraz wylosowanie jej nie ogranicza wylosowania innej jednostki to schemat taki nazywamy **nieograniczonym.**
- **Losowanie warstwowe**. – Jest to losowanie, w którym dzieli się całą zbiorowość na kilka rozłącznych zbiorów zwanych **warstwami**. Kryterium podziału na warstwy powinno zapewnić zróżnicowanie rozkładu między warstwami. Same zaś warstwy powinny być w swej strukturze jak najbardziej jednorodne. Losowanie próby dokonuje się z każdej warstwy niezależnie. Próbę losową stanowi suma elementów wylosowanych z każdej warstwy. Ten schemat losowania jest więc schematem **ograniczonym.**

Koncepcja warstwowania wzięła się stąd, że istnieją pewne sytuacje – tzn. pewne rozkłady, których charakterystyki zamierzamy szacować – gdzie zróżnicowanie ze względu na badaną cechę jest zbyt duże. To znaczy jest ono na tyle duże, że może zaistnieć sytuacja

(prawdopodobieństwo takiego zdarzenia wcale nie jest bliskie zeru), w której pewna część zbiorowości nie będzie reprezentowana dostatecznie licznie. Dla uniknięcia takiego niebezpieczeństwa ograniczono prawdopodobieństwo zaistnienia takiej próbki do zera przez dokonanie podziału na warstwy. W dwóch przykładach jakie zaprezentujemy, oba te schematy znajdą zastosowanie. Jest spora ilość schematów losowania, które w zależności od konkretnej sytuacji są najlepsze do przeprowadzenia badania. O wyborze tego czy innego schematu nie można z góry przesądzić, że jest ono najlepsze. Zależy to od wielu czynników, takich np., jak: jaką cechę zamierzamy badać, jak jednorodny jest charakter danych itp.

WYKORZYSTANIE PROGRAMU KOMPUTEROWEGO DO BADANIA SPRAWOZDANIA FINANSOWEGO

Zaprezentowaną metodę reprezentacyjną można stosować z powodzeniem w procedurach audytorskich. Potrzeba jednak już nie tylko ogólnej wiedzy na temat stosowania tychże metod, lecz konieczne są do tego celu konkretne wzory, wykonane obliczenia oraz algorytmy. W bogatej literaturze na ten temat można znaleźć wszystkie niezbędne do tego celu informacje. Jednak – z góry uprzedzając Czytelnika – nie jest to lektura łatwa. Trzeba też dodać, że od ogólnej wiedzy teoretycznej do faktycznej, praktycznej umiejętności jej zastosowania jest droga daleka. Tutaj może nam przyjść z pomocą gotowe – zaprojektowane na te potrzeby – oprogramowanie komputerowe. Nieistotne jest w tym miejscu - z punktu widzenia użytkownika - jakim programem się posłużymy do rozwiązania tego zadania. Istotne jest, aby narzędzie takie było bardzo dobrze przetestowane zanim trafi do rąk odbiorcy. Wcześniejsze wprowadzenie w terminologię stosowaną w metodach reprezentacyjnych jest dla umiejętności korzystania z takiego specjalistycznego oprogramowania niezbędne. Obok tej terminologii nie mniej ważne jest „wyczucie" samej istoty, która leży u podstaw tej dziedziny nauki, a którą jest pojęcie **prawdopodobieństwa**. Drugorzędną sprawą jest – i w artykule tym nie będzie o tym mowy – właściwa obsługa tego czy innego programu, przygotowanie danych wejściowych i inne sprawy techniczno-eksploatacyjne.

PRZYKŁAD PIERWSZY Z WYKORZYSTANIEM SCHEMATU LOSOWANIA PROSTEGO ZALEŻNEGO NIEOGRANICZONEGO

Rozważmy następujący przykład. Należy zweryfikować pewną pozycję bilansu (saldo), której składnikami jest 3000 operacji (dokumentów). W tym celu należy zbudować przedział ufności przyjmując dwa parametry:

1. ryzyko
2. precyzję szacunku

Przyjmijmy następujące założenia:

1. Wartość księgowa: 15 000 000 zł

Jest to wartość, którą chcemy zweryfikować (została podana przez badaną firmę). Prawdziwa wartość może być inna!

2. Współczynnik ryzyka: 5 %
3. Precyzję 750 000 zł

Inaczej mówiąc zakładamy 95 % pewności tego, iż przedział ufności, który powstanie pokryje szacowana wartość z tolerancją błędu ± 750 000 (5 % kwoty podanej przez badaną firmę). Kwestia, czy w przedziale, który powstanie, znajdzie się podana przez badaną firmę kwota księgowa 15 000 000 zł. jest w tej chwili otwarta. I poniekąd nie jest to „zmartwienie" rewidenta, a raczej audytowanej firmy. Jeśli bowiem okaże się, że kwota ta nie mieści się w przedziale ufności, to biegły ma podstawę do wydania opinii negatywnej z przeprowadzonego procesu weryfikacji.

OPIS DZIAŁANIA PROGRAMU

Nie wchodząc w szczegóły, opiszemy po krótce, jaki jest algorytm działania programu w przypadku, gdy dane charakteryzują się wymaganą jednorodnością (patrz Wykres 1). Powiedzieliśmy wcześniej, że chcąc wyznaczyć wymaganą liczebność próby należy znać wartość pewnego parametru badanego rozkładu, a mianowicie odchylenie standardowe. Na ogół parametru tego nie znamy, zatem pozostaje jego oszacowanie. W tym celu wybieramy losową próbę wstępną. Jest ona potrzebna tylko na „wewnętrzny użytek" programu (biegły rewident nie kontroluje jeszcze żadnych dokumentów). Następuje wyznaczenie wymaganej liczebności próbki i rzeczywiste jej losowanie. Do rąk biegłego trafia pierwsza partia wylosowanych elementów (np. w postaci wydruku). Przystępuje on do badania. W którego trakcie może się okazać, że jakąś pozycję należy skorygować, zmieniając jej wartość. Biegły po sprawdzeniu wszystkich wylosowanych elementów przechodzi do następnej fazy programu, w której następuje ustalenie (znowu wg teoretycznych wzorów dla tej metody stosowanych), czy próbka jest już reprezentatywna. Jeżeli tak, to zostaje zbudowany przedział ufności oraz wydana opinia końcowa o wyniku badania. W przypadku, gdy próbka nie jest jeszcze dostatecznie reprezentatywna (nie spełnia naszych wymogów, co do wartości dopuszczalnego błędu) następuje dalsze losowanie, ale już tylko takiej ilości elementów (do tych, które są wylosowane), aby spełnić zadość wyznaczonej nowej liczebności próby. Następnie rewident bada dolosowaną partię elementów i dokonuje – jeżeli jest taka potrzeba – stosownych zmian. Należy podkreślić, że w tym etapie działania programu, może on dokonywać zmian tylko w ostatnio dolosowanej partii elementów! Zmiany takie mogą być na tyle istotne, że zajdzie konieczność dolosowania – w kolejnym kroku – znacznej ilości nowych elementów. Procedurę tę biegły wykonuje tak długo, aż próbka, którą uzyska spełni założoną precyzję.
Tabela 1 pokazuje kolejne etapy działania programu. Przy czym, w pierwszym przypadku (lewa część Tabeli 1) biegły nie wykonał żadnych zmian w wylosowanych elementach, bądź zmiany te były drobne (np. rzędu kilku, kilkunastu procent wartości pierwotnej i to w niewielu elementach). Charakterystyczne jest wtedy to, iż w kolejnych krokach programu nowe wymagane liczebności próby rosną coraz wolniej. Jest tak, gdyż już pierwsze losowanie dało w miarę dokładny obraz całej zbiorowości, a kolejne dolosowania ten obraz już tylko „wyostrzają".

W przypadku drugim (prawa część Tabeli 1) widać charakterystyczny skok ilości elementów w próbce. Stało się tak, gdyż biegły dokonał poważnej zmiany w jednej lub kilku pozycjach, co (mówiąc obrazowo) zburzyło powstający obraz szacowanej zbiorowości i trzeba go niejako kreślić od nowa.

Tabela 1
Przebieg realizacji programu: dobieranie elementów do próbki.

Wariant bez 'poważniejszych' zmian	Liczebność próbki	Wariant z 'dużymi' zmianami	Liczebność próbki
próbka wstępna	50	próbka wstępna	50
pierwsze dolosowanie	177	pierwsze dolosowanie	153
drugie dolosowanie	193	drugie dolosowanie	203
trzecie dolosowanie	199	trzecie dolosowanie	209
próbka reprezentatywna	*199*	czwarte dolosowanie	**287**
----------	---	*próbka reprezentatywna*	*287*

Źródło: Opracowanie własne.

Na koniec następuje – co pokazuje Tabela 2 - wyznaczenie przedziału ufności wraz z końcową oceną przeprowadzonego badania.

Tabela 2
Opinia biegłego i ocena końcowa badania.

	wariant I	wariant II
prawdziwa wartość szacowanego salda	14000000	15100000
Wartość podana przez firmę	15000000	15000000
przedział ufności	13 462 541 – 14 208 746	14 568 482 - 15 321 559
ocena biegłego	negatywna	pozytywna

Źródło: Opracowanie własne.

W wariancie I prawdziwa wartość salda znacznie różni się od tej podanej przez firmę. Nasz szacunek prowadzi do wydania opinii negatywnej, gdyż podana kwota 15 000 000 nie znalazła się w powstałym przedziale ufności.
Podkreślmy z naciskiem, że może – przy innym składzie próbki – powstać przedział ufności, w którym kwota 15 000 000 będzie się mieściła! Rzecz w tym, iż jest to możliwe statystycznie w 5 na 100 przypadków. Wtedy wydając ocenę pozytywną rozminięto by się z rzeczywistością. Sytuację taką określa się jako „nieprawidłowa akceptacja".

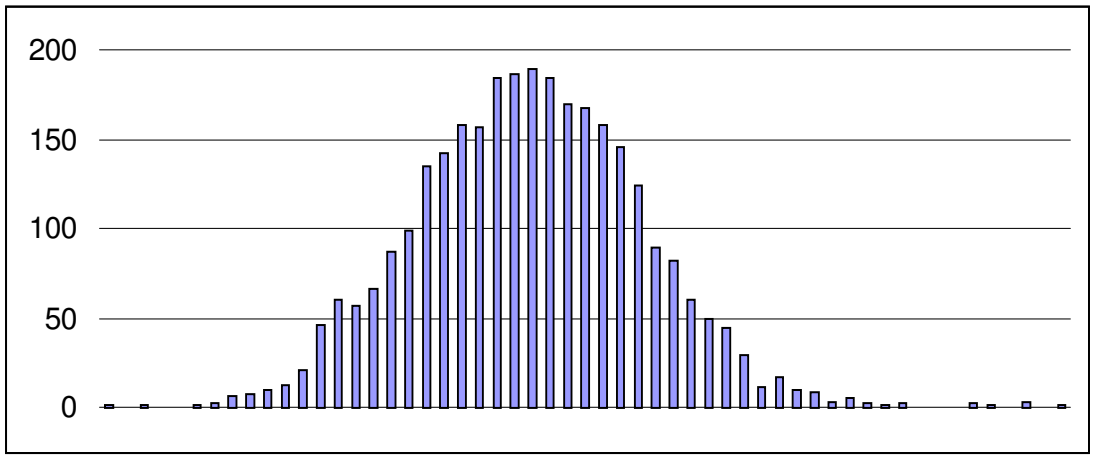
W wariancie II podana przez firmę kwota znalazła się wewnątrz przedziału ufności. Wydano zatem opinie pozytywną. Jednak i tutaj może się zdarzyć, że próbka, którą otrzymamy doprowadzi nas do fałszywego wniosku. Będzie tak, jeśli powstanie przedział ufności, w którym nie znajdzie się podana przez firmę kwota 15 000 000. Jest to możliwe, ale tylko z prawdopodobieństwem równym 0,05! Wtedy, wydając ocenę negatywną popełniono by również błąd. Sytuację taką określa się jako „nieprawidłowe odrzucenie".

Rozważmy teraz różne kombinacje wartości ryzyka i precyzji. Zobaczymy jak kształtują się liczebności próbki. Precyzję wyrazimy (dla czytelności tabeli) w procentach zamiast w wartościach bezwzględnych, pamiętając jednak, że w teorii precyzja procentowa odnosi się do szacowanej, nieznanej wielkości! Przeprowadzono w tym celu 500 niezależnych testów. Każda kombinacja ryzyka oraz precyzji została obliczona jako średnia arytmetyczna pięciu

testów. Wszystkie otrzymane wyniki odnoszą się do przyjętych wcześniej założeń, czyli do szacowania salda o wartości 15 000 000 zł. Zakładamy ponadto, że zbiór danych w trakcie weryfikacji nie uległ żadnym zmianom.

Dla uzupełnienia dodajmy, iż dane te mają rozkład normalny o średniej 5000 i odchyleniu standardowym 1000. Jest to o tyle istotne, że program komputerowy będzie analizował dane wejściowe i – w zależności od wyników tej analizy – proponował, czy wręcz narzucał inny schemat losowania. Na Wykresie 1 pokazany jest rozkład częstości występowania liczb w zbiorze wejściowym. Charakteryzuje go duża jednorodność, co jest istotne z punktu widzenia wyboru schematu losowania.

Wykres 1
Częstości występowania liczb w zbiorze danych wejściowych.



Źródło: Opracowanie własne.

W tabeli 3 przedstawione zostały uzyskane wyniki.

Tabela 3
Liczebność próbki w zależności od wartości parametrów ryzyka i precyzji

ryzyko \ precyzja	0,5	1	1,5	2	2,5	3	3,5	4	4,5	5	fałszywe wnioski
1	x	x	x	610	393	309	**219**	168	130	125	0
2	x	x	782	501	350	274	**173**	117	105	86	0
3	x	x	740	427	267	**207**	153	132	100	81	0
4	x	x	632	417	286	**189**	143	120	94	61	2
5	x	x	623	367	238	**197**	142	100	89	58	1
6	x	x	556	357	**211**	165	125	94	89	61	0
7	x	x	520	340	**223**	160	130	88	63	59	1
8	x	x	498	301	**199**	147	122	85	65	55	3
9	x	869	500	291	**181**	132	96	82	68	55	5
10	x	822	479	263	**170**	121	86	68	61	50	7

Źródło: Opracowanie własne.

Znak „x” oznacza, że próbka byłaby zbyt duża względem przyjętych założeń, tj. miałaby powyżej 30 % wszystkich elementów badanego zbioru. Z punktu widzenia praktycznego

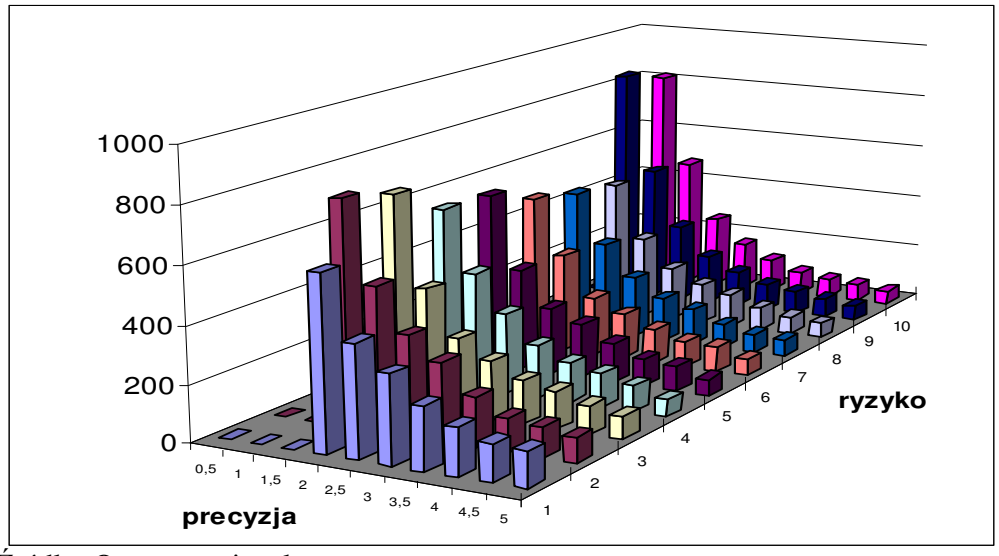
odrzuca się takie przypadki i zaleca poszerzenie przedziału ufności przez zwiększenie precyzji.

Z analizy uzyskanych danych bardzo wyraźnie rysują się – zgodnie zresztą z intuicją – dwie prawidłowości, a mianowicie:

- przy stałej wartości precyzji wzrost ryzyka powoduje obniżanie wartości próbki
- przy stałej wartości ryzyka wzrost precyzji powoduje obniżenie wartości próbki.

W przypadku stałej wartości ryzyka zależności te nie są charakteru liniowego, co pokazuje Wykres 2, a których można się doszukać przy stałej wartości precyzji.

Wykres 2
Wymagana liczebność próbki dla różnych wartości ryzyka i precyzji



Źródło: Opracowanie własne.

Ponadto da się zauważyć, iż ilość nietrafnych ocen wzrasta wraz ze wzrostem ryzyka. Przykładowo w wierszu 9 tabeli na 10*5 testów otrzymaliśmy pięć przypadków, gdzie w przedziale ufności nie mieściła się kwota 15 000 000 zł, podczas gdy wiadomo, iż rzeczywista wartość wynosi właśnie 15 000 000 zł! Jednak ilość nietrafnych ocen jest bliska teoretycznej wartości prawdopodobieństwa wydania błędnej oceny przy ryzyku równym 9 %.
Mamy bowiem:
(z testów)- 5 błędów na 50 prób, czyli 5/50=0,1
(z teorii) - ryzyko 9 %, co wartościowo oznacza 0,09.
Zatem zachodzi przybliżona równość (0,1 ~ 0,09).
Zarysowała się jeszcze jedna prawidłowość. Jeżeli poszukamy w każdym wierszu tabeli wartości najbliższej liczbie np. 200, to zaobserwujemy, że tworzą one wznoszącą się linię, patrząc od lewej strony wiersza 10 tabeli. Liczby te mieszczą się w zakresie 200 +- 15 %. Gdybyśmy zbudowali podobną tabelę o 100 wierszach i 100 kolumnach, krzywa byłaby jasno zarysowana. Co z tego faktu wynika? Otóż zdarzają się sytuacje w praktyce zawodowej biegłego rewidenta, że dobór wartości ryzyka i precyzji jest uzależniony od kosztów, jakie możemy ponieść lub czasu, który możemy przeznaczyć na badanie w konkretnej firmie. Innymi słowy, jeżeli znamy szacunkowo koszt badania „jednego dokumentu" lub „czas potrzebny na jego kontrolę" wiemy jak liczną próbkę możemy pobrać. To z kolei wskaże nam „krzywą" (podobną do tej z przykładu), której jakiś punkt możemy wybrać. W naszym

przykładzie wybór może paść na następujące wartości: ryzyko = 8 % i precyzja = 2,5 % (tj. 375 000).

W przykładzie tym próbowaliśmy wykazać korzyści, jakie biegły rewident mógłby odnieść gdyby, stosując metody reprezentacyjne w swej pracy, korzystał ze specjalistycznego oprogramowania komputerowego. Czytelnik sam stwierdzi, czy możliwość wykonania wstępnej symulacji komputerowej pomogłaby by mu w zaplanowaniu procedury badania finansowego. W przykładzie drugim zobaczymy jak od jakości programu komputerowego zależeć może bezpośrednio koszt badania audytorskiego.

PRZYKŁAD DRUGI Z WYKORZYSTANIEM SCHEMATU LOSOWANIA PROSTEGO ZALEŻNEGO WARSTWOWEGO

Zjawiska ekonomiczne charakteryzują się dużą różnorodnością. Nie jest zatem możliwe, by dla każdego napotkanego przez biegłego rewidenta przypadku stosować zawsze ten sam schemat losowana (tzn. ten sam algorytm programu). W poprzednim przykładzie założyliśmy przypadek prosty do analizy, lecz rzadko spotykany. Obecnie postawimy problem inaczej. Wykażemy – przez porównanie dwóch metod – jak ważny może być sposób warstwowania z punktu widzenia minimalizacji kosztów badania. Niech celem badania będzie – podobnie jak w przykładzie pierwszym - oszacowanie salda (wartości księgowej) zadanej zbiorowości, której składnikami jest 6000 operacji (dokumentów). Należy zbudować przedział ufności przyjmując następujące parametry:

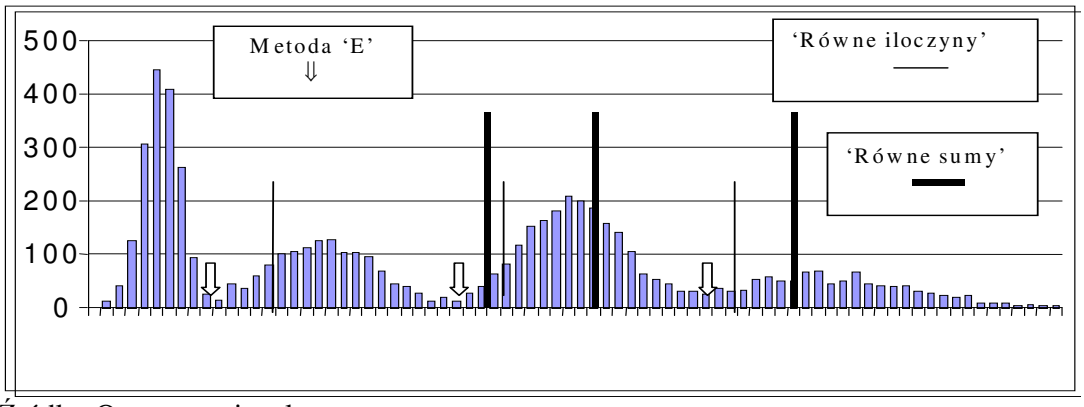
1. Wartość księgowa: 18 000 000 zł
2. Współczynnik ryzyka: 5 %
3. Precyzję 450 000 zł
4. Kwota brzegowa: 10 000 zł.

OPIS DZIAŁANIA PROGRAMU

W przypadku, gdy dane wejściowe charakteryzują się zbyt dużą niejednorodnością (patrz Wykres 3) algorytm programu będzie wyglądał następująco. Zgodnie ze schematem losowania warstwowego dokonamy podziału zbioru na warstwy w ten sposób, aby, sumy elementów w warstwach były w miarę równe (patrz Wykres 3). Jest to najprostszy sposób warstwowania, gdyż po posortowaniu elementów w porządku niemalejącym i po ustaleniu ilości warstw (o tym z kolei decyduje współczynnik zmienności zbioru danych), wyznacza się średnią sumę wartości warstwy. Następnie dokonuje się podziału elementów w ten sposób, że kolejne elementy od najmniejszego do największego „wchodzą” do warstwy tak długo, aż suma tych elementów osiągnie wyznaczoną średnią sumę warstwy. Tworzy się dodatkowo warstwę pełną (zwaną tez stuprocentową), w skład, której wejdą wszystkie elementy „nietypowe” tj. równe zero lub ujemne, a także wszystkie elementy większe od zadanej przez biegłego wartości. Może się jednak zdarzyć, że warstwa ta będzie pusta. W naszym przykładzie warstwa ta liczy 15 elementów. Są to wszystkie pozycje większe od kwoty 10 000 zł. Metoda ta spełnia podstawowy postulat warstwowania, tzn. tworzy warstwy zróżnicowane między sobą (średnie w warstwach mocno się różnią) oraz warstwy są

wewnętrznie mało zróżnicowanie (w każdym razie, mniej niż cała zbiorowość rozpatrywana razem). Z powyższego opisu wynika, że metoda ta jest łatwa do zastosowania w programie komputerowym. Osobnym zagadnieniem jest – po wyznaczeniu warstw – metoda pobierania próbek. Metod tych jest również kilka. W naszych przykładach zastosowano jedną z bardziej popularnych, mianowicie metodę proporcjonalnego do ilości elementów w każdej warstwie doboru elementów, korygowaną przez odchylenie średnie badanej cechy w warstwie. Nie jest to więc ściśle proporcjonalny podział elementów do próbki, ale zależy też od jej wewnętrznego zróżnicowania. Co do samej procedury uzyskiwania próbki reprezentatywnej, to algorytm jest już bardzo podobny do opisanego w poprzednim przykładzie. Korzystając rzecz jasna z innych wzorów, wykonując kolejne kroki, uzyskujemy efekt końcowy, czyli przedział ufności i opinię końcową o wyniku badania. Podział na warstwy przedstawia lewa część Tabeli 4. Rozwiązanie zaś ilustruje Tabela 5.

Wykres 3
Częstości występowania liczb w zbiorze wejściowym oraz podział na warstwy



Źródło: Opracowanie własne.

Tabela 4
Struktura warstw dla dwóch wariantów warstwowania

wyrównywanie sum			wyrównywanie iloczynów			
warstwa	liczebność warstwy	suma elementów warstwy	warstwa	liczebność warstwy	suma elementów warstwy	iloczyn warstwy
I	15	217381	I	15	217381	-
II	3201	4447542	II	2100	1797479	95
III	1148	4448939	III	1198	2994255	98
IV	945	4448586	IV	1821	7590290	101
V	691	4437552	V	866	5400595	95
razem	6000	18000000	razem	6000	18000000	-

Źródło: Opracowanie własne.

Tabel 5.

prawdziwa wartość szacowanego salda	18000000
wartość podana przez firmę	18000000
przedział ufności	17 701 976 - 18 598 030
ocena biegłego	pozytywna

Źródło: Opracowanie własne.

Rozważmy teraz wariant, w którym warstwowanie przeprowadzone zostanie inną metodą. Metoda ta polega na tym, by wyrównywać warstwy nie wg równych sum elementów warstw, lecz wg równych iloczynów następujących dwóch czynników:

- frakcji warstwy
- odchylenia średniego w warstwie.

Frakcja warstwy, to liczba wyrażająca stosunek liczby elementów w warstwie do liczby wszystkich elementów. Odchylenie średnie, to „rozrzut" poszczególnych wartości w warstwie względem średniej. W poprzedniej metodzie z samej jej definicji wynika, że ilość elementów pobieranych z warstwy pierwszej (w naszych przykładach jest to warstwa nr II) do próbki będzie największa. Musi tak być dlatego, że wewnętrzne zróżnicowanie w tej warstwie jest zawsze największe w porównaniu z pozostałymi warstwami oraz ilość elementów w tej warstwie jest zawsze największa. Mankament ten w nowym sposobie warstwowania spróbowano wyeliminować przez wyrównanie poziomu zróżnicowania między warstwami. Wynik prezentuje część prawa Tabeli 4. Aby zobrazować, jak sam wybór sposobu warstwowania może wpłynąć na wymaganą liczebność próbki przeprowadzona została następująca symulacja komputerowa:

Dla każdego z wariantów warstwowania przeprowadzono 30 testowych badań rewizyjnych szacujących podaną przez badaną firmę wartość księgową. Tabela 6 prezentuje uzyskane wyniki.

Tabela 6

Liczebność próbki w zależności od wartości ryzyka

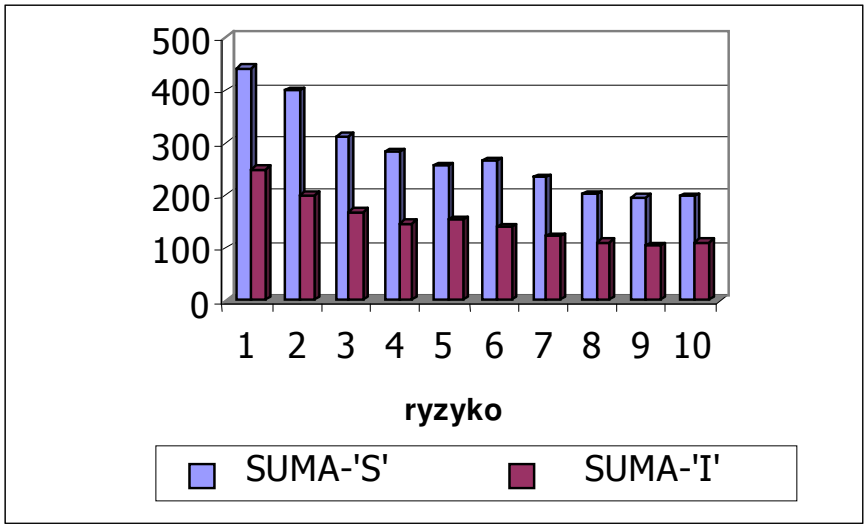
precyzja: 450 000 (2,5 %)										
metoda:	wyrównywanie sum'					wyrównywanie 'iloczynów'				
ryzyko	II	III	IV	V	SUMA-'S'	II	III	IV	V	SUMA-'I'
1	328	26	46	44	444	71	64	63	53	251
2	289	23	44	45	400	56	48	55	41	200
3	228	19	36	31	313	49	43	43	35	170
4	200	16	34	34	284	39	40	37	29	146
5	186	13	32	27	257	42	36	41	34	153
6	191	15	31	27	265	42	33	35	30	139
7	169	13	28	24	234	33	30	36	22	122
8	145	12	25	21	203	32	28	29	20	110
9	139	11	24	22	196	28	26	28	22	104
10	142	11	24	22	199	30	27	31	24	111

Źródło: Opracowanie własne.

Z analizy powyższej tabeli wynika jednoznacznie, że można bardzo poważnie ograniczyć ilość wymaganych (przez precyzję) elementów w próbce. W przykładzie tym obniżenie próbki sięga średnio 50 %.

Ilustruje to Wykres 4.

Wykres 4
Liczebność próbki w dwóch wariantach warstwowania.



Źródło: Opracowanie własne.

Suma' oznacza wariant warstwowania wg wyrównywania sum.
Umai' oznacza wariant warstwowania wg wyrównywania iloczynów.

Tabela 7 przedstawia podobne zestawienie, gdzie stałą jest tym razem wartość ryzyka, a zmienia się wartość precyzji.

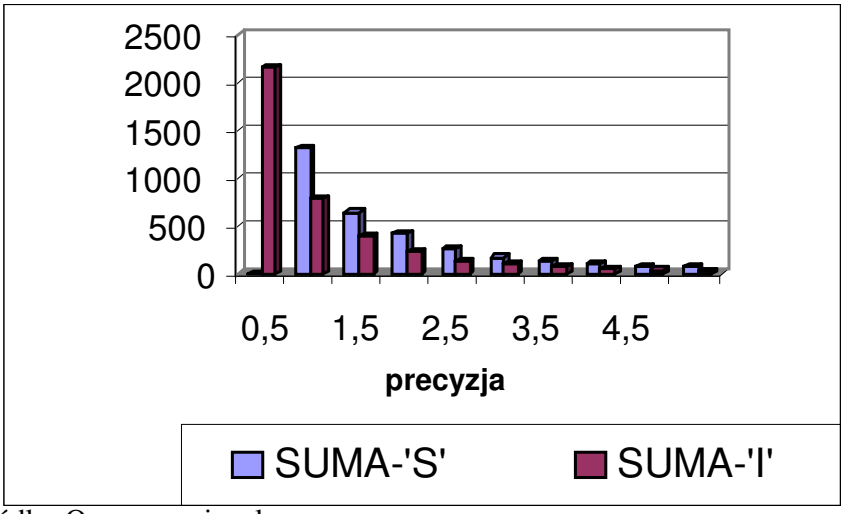
Tabela 7
Liczebność próbki w zależności od wartości precyzji

ryzyko: 5 %										
metoda:	wyrównywanie sum'					wyrównywanie iloczynów'				
precyzja	II	III	IV	V	SUMA-'S'	II	III	IV	V	SUMA-'I'
0,5	x	x	x	x	0	602	573	562	422	2159
1	934	77	164	131	1307	214	201	204	168	787
1,5	469	39	74	61	643	115	102	98	71	386
2	308	24	48	41	422	69	58	58	46	231
2,5	185	14	33	27	258	37	35	39	27	138
3	124	11	21	18	173	31	25	24	18	99
3,5	100	8	16	14	137	22	19	19	15	75
4	71	6	12	12	101	15	15	12	10	52
4,5	64	5	10	8	87	10	10	10	10	40
5	52	4	8	8	72	9	7	8	5	30

Źródło: Opracowanie własne.

W zestawieniu powyższym zauważamy, że wartość próbki zmienia się nie w sposób liniowy tak, jak to miało miejsce przy stałej precyzji

Wykres 5
Liczebność próbki w dwóch wariantach warstwowania.

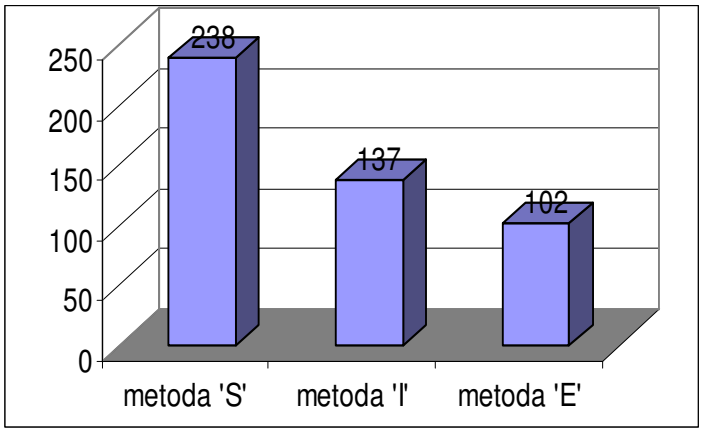


Źródło: Opracowanie własne.

METODY EKSPERCKIE

Metoda „wyrównywania iloczynów" jest bardzo trudna do zastosowania w programie komputerowym w tym sensie, aby program sam dokonał właściwego podziału na warstwy. Powinien on tak długo „przymierzać" się do warstw, aż owe iloczyny będą w miarę równe. Gdyby program mógł otrzymać jakieś „wskazówki" od użytkownika sprawa byłaby prostsza. Ponadto przyniosłyby jeszcze inną korzyść. Na wykresie 3 zaznaczono (strzałkami) punkty, które (kierując się doświadczeniem i intuicją) być może nadawałyby się jeszcze lepiej do dokonania w nich podziału na warstwy. Tak też jest w rzeczywistości, co potwierdziły stosowne obliczenia. Ich efekt zaprezentowany jest na Wykresie 6, gdzie widać, że udało się jeszcze poprawić efektywność warstwowania przez zastosowanie innego podziału. Przy tym podziale założenie o równaniu iloczynów musiało ustąpić metodzie optymalnej (dla tego konkretnego przykładu). Może tak być, gdyż w proponowanych schematach warstwowania ich autorzy dodają, że sprawdzają się one najlepiej w pewnych tylko kategoriach (klasach) zbiorowości. I tutaj dochodzimy do zagadnienia, które w informatyce określa się jako metody **eksperckie i heurystyczne**. Otóż tam, gdzie samo rachowanie nie wystarcza, gdyż nie znamy algorytmu, albo jego opracowanie jest zbyt kosztowne, przychodzi nam z pomocą heurystyka. W algorytmie heurystycznym nie ma analizy problemu z gotowym jego rozwiązaniem. Decydując się na użycie heurystyki, wyposażamy komputer w ogólne tylko zasady wynikające z konkretnej teorii (tutaj metody reprezentacyjne) oraz z naszego dotychczasowego doświadczenia zdobytego w tym zakresie. Z przykładu drugiego widać jasno jak pomocna dla końcowego efektu byłaby ingerencja doświadczonego użytkownika w przebieg działania programu. Musimy zdać sobie sprawę z tego, że w codziennej praktyce rewizyjnej spotkamy dziesiątki czy setki (rozkładów) zbiorów danych, które będą się mocno od siebie różniły. Konieczne jest tworzenie oprogramowana typu eksperckiego, gdzie stosując metody heurystyczne, aktywny udział użytkownika doprowadzi do wyboru najlepszej w danych warunkach i przyjętych założeniach metody badania.

Wykres 6
Liczebność próbki w trzech schematach warstwowania



Źródło: Opracowanie własne.

PODSUMOWANIE

Zaprezentowany materiał przekonał – jak sądzimy – Czytelnika, co do dwóch rzeczy. Po pierwsze przy wykorzystaniu metod reprezentacyjnych korzyści przynieść może wspomaganie komputerowe w postaci prostego nawet oprogramowania. Jeżeli chcemy zaoszczędzić na czasie podczas wykonywania żmudnych obliczeń, jeżeli nie chcemy ryzykować popełnienia błędu w obliczeniach, jeżeli nie zamierzamy korzystać z tablic liczb losowych, to z takiego programu z pewnością warto skorzystać.

Drugim argumentem jest to, co nazwaliśmy wyżej „efektywnością" próbki. Dla osiągnięcia tych celów z pomocą przyjść może oprogramowanie, o bardziej już skomplikowanej budowie. Oprogramowanie, którym będzie można przeprowadzić różnego rodzaju symulacje, szukając odpowiednich parametrów dla właściwego badania. Program, dzięki któremu da się wyszukać optymalny – w danym badaniu - schemat losowania i metodę pobierania elementów do próbki. Program taki może być ukierunkowany na wąską kategorię badań audytorskich, bądź też mieć charakter uniwersalny. Wiemy bowiem, że obszar zastosowań metod reprezentacyjnych wspomaganych komputerowo może być bardzo szeroki.

W tym miejscu postawimy nawet tezę bardziej odważną - zaprzeczając poniekąd samemu tytułowi artykułu. – Mianowicie efektywne stosowanie metod reprezentacyjnych z całym ich aparatem badawczym nie jest możliwe bez specjalistycznego oprogramowania typu eksperckiego. Nie chodzi tu już tylko o „wspomaganie" komputerowe, lecz o zupełnie nową jakość, która wynika z nowej metodologii badań. Chcąc zatem mówić o zastosowaniu metod reprezentacyjnych w audycie trzeba mówić także o systemach informatycznych na metodach tych opartych. Żywimy zatem nadzieję, że Czytelnik zechce zainteresować się bliżej poruszonym tutaj tematem. A stosując metody reprezentacyjne i właściwe do nich narzędzia programistyczne odniesie w swej pracy zawodowej wymierne korzyści i satysfakcję.