ANALIZA SYSTEMOWA I ZARZĄDZANIE

Książka jubileuszowa
z okazji
50-lecia pracy naukowej

ROMANA KULIKOWSKIEGO

ISBN 83-85847-34-0

Druk: "ARGRAF" Agencja Poligraficzno-Wydawnicza, Warszawa
Skład: Barbara Kotuszewska

DECYZJE STATYSTYCZNE W BADANIACH SYSTEMOWYCH

Olgierd Hryniewicz

Instytut Badań Systemowych PAN

1 Badania systemowe jako narzędzie wspomagania decyzji

Początków badań systemowych należy szukać we wczesnych latach pięćdziesiątych w Stanach Zjednoczonych, gdy powołany został zespół doradczy mający zaproponować najlepsze rozwiązanie dotyczące dyslokacji strategicznych sił lotniczych USA. Ponieważ praca nad tym zadaniem została oceniona bardzo pozytywnie, badania tego typu kontynuowano, przeważnie na zlecenie sił zbrojnych Stanów Zjednoczonych. Z badaniami systemowymi tego okresu wiąże się, przede wszystkim, działalność słynnego instytutu badawczego RAND Corporation, będącego zapleczem intelektualnym rządu USA. Z biegiem czasu badania systemowe, lub utożsamiana z tym pojęciem analiza systemowa, znalazły zastosowania cywilne. Znane są wyniki badań dotyczących systemu przeciwpowodziowej ochrony Holandii, budowy systemu szybkiej kolei w Japonii, problemu transgranicznych przepływów zanieczyszczeń atmosferycznych i wielu innych [2]. Znaczna część tych prac wiąże się z działalnością powstałego w połowie lat siedemdziesiątych Międzynarodowego Instytutu Stosowanej Analizy Systemowej (IIASA) w Laxenbergu pod Wiedniem.

W Polsce badania systemowe znalazły uznanie w latach siedemdziesiątych, gdy utworzono pod kierunkiem profesora Romana Kulikowskiego Instytut Badań Systemowych Polskiej Akademii Nauk. Ich dalszy rozwój związany był także z działalnością wielu naukowców i zespołów badawczych współpracujących z IIASA, a jednym z najważniejszych projektów był realizowany na przełomie lat siedemdziesiątych i osiemdziesiątych program rozwoju regionu Górnej Noteci. Informacje o realizowanych - również przy współudziale polskich naukowców - aktualnych projektach IIASA można znaleźć w pracy [1].

Badania systemowe jako niezależna dyscyplina naukowa nie zostały zdefiniowane w sposób jednoznaczny. Różni badacze pojmują je w różny sposób, dokonując własnych rozgraniczeń pomiędzy bardzo zbliżonymi obszarami badań naukowych jakimi są: analiza systemowa, teoria modelowania systemów oraz badania operacyjne. Wszystkie te dziedziny nauki sytuują się na styku matematyki stosowanej, informatyki, automatyki, nauk o zarządzaniu oraz nauk ekonomicznych. Wynika stąd istotna trudność w precyzyjnym określeniu czym naprawdę są badania systemowe. Nie ulega jednak wątpliwości, że podstawową dziedziną badań systemowych jest analiza systemowa, która wg Quade'a [2]

"... jest formalnym i jawnym badaniem wspomagającym działanie osób odpowiedzialnych za decyzje lub linię postępowania w określonej sytuacji charakteryzującej się niepewnością. Analiza systemowa ma na celu określenie pożądanego działania lub linii postępowania przez rozpoznanie i rozważenie dostępnych wariantów oraz porównanie ich przewidywanych następstw"

Analizując treść powyższej definicji zwróćmy uwagę na dwa jej istotne aspekty. Po pierwsze, problemy będące przedmiotem zainteresowania analizy systemowej odznaczają się związaną z nimi niepewnością. W większości przypadków jest to niepewność związana z losowością opisywanych zjawisk. Typowym przykładem mogą być analizy dotyczące złożonych problemów zabezpieczenia przed skutkami katastrof naturalnych, takich jak powodzie. Innym przykładem niepewności związanej ze zjawiskami losowymi jest analiza celowości wprowadzania energetyki nuklearnej. Należy w niej uwzględnić ryzyko związane z awarią reaktora nuklearnego, którego to zjawiska nie można, niestety, przewidzieć. Istnieje również inny element niepewności, charakterystyczny dla zadań analizy systemowej. Jest on związany z koniecznością dokonywania prognoz, a także z budowaniem modeli matematycznych złożonych zjawisk społecznych i gospodarczych na podstawie danych empirycznych zawierających składowe o charakterze losowym.

Uwzględnienie w analizie systemowej elementu niepewności wymaga zastosowania odpowiednich narzędzi, jakimi są rachunek prawdopodobieństwa i statystyka matematyczna. W przypadku gdy wykorzystujemy dane empiryczne, statystyka jest podstawowym narzędziem pracy specjalisty badań systemowych. Ze względu na różnorodność i złożoność analizowanych problemów w badaniach systemowych możemy wykorzystać różne metody statystyczne. Niektóre z nich, związane z budową modeli matematycznych opisujących analizowane problemy, opisane są w napisanych przez K.Mańczaka rozdziałach książki [2].

Drugi istotny aspekt analizy systemowej, który również dotyczy całej dziedziny badań systemowych, związany jest z podstawowym jej zadaniem - wspomaganiem decyzji. Ostatecznym celem przeprowadzania jakiejkolwiek analizy systemowej powinno być wskazanie właściwemu decydentowi odpowiednich w danej sytuacji decyzji. Propozycje te powinny być właściwie umotywowane, w tym także przez przedstawienie uzasadniających ich celowość analiz, w których wykorzystane są dane statystyczne. Odpowiednie narzędzia do tego celu dostarcza teoria decyzji statystycznych.

W niniejszym artykule przedstawiony zostanie jeden, stosunkowo wąski, fragment teorii decyzji statystycznych, który powinien znaleźć zastosowanie w badaniach systemowych. Dotyczy on decyzji o charakterze alternatywnym (TAK lub NIE), które należy podjąć na podstawie uzyskanych danych statystycznych mających taki sam charakter. Typowym przykładem opisanej tu metodologii jest analiza ankiet oceniających celowość proponowanych rozwiązań. Ponieważ wiele decyzji, które są podejmowane z wykorzystaniem wyników badań systemowych, ma swoje konsekwencje dla jakiejś grupy ludzi (np. mieszkańców pewnego regionu) konieczne jest wykonanie odpowiednich badań. Właściwe zastosowanie metod statystycznych jest tu rzeczą niezwykle istotną, gdyż od poprawności analizy wyników tych badań może zależeć powodzenie całego przedsięwzięcia. W rozdziale drugim niniejszego artykułu zostaną przedstawione podstawowe metody pozwalające na właściwą interpretację wyników takich badań.

Z innym przypadkiem danych statystycznych o charakterze alternatywnym mamy do czynienia w przypadku zjawisk zachodzących stosunkowo rzadko. Wielkością wykorzystywaną w analizie systemowej jest tu prawdopodobieństwo zajścia takiego, lub podobnego, zdarzenia. Ze względu na ubogość jednorodnego materiału empirycznego nie jesteśmy zazwyczaj w stanie skorzystać z klasycznych metod statystyki matematycznej. W związku z tym, przy analizie zjawisk tego typu należy korzystać z różnych źródeł informacji, w tym również z opinii ekspertów. Podstawowe problemy związane z podejmowaniem decyzji w takich przypadkach przedstawione zostały w rozdziale trzecim niniejszego artykułu.

2 Decyzje statystyczne przy danych wg klasyfikacji alternatywnej

Najprostszym przypadkiem danych statystycznych wykorzystywanych w analizach prowadzonych w ramach badań systemowych są dane określane wg klasyfikacji alternatywnej. Ogólnie rzecz biorąc, są to dane dotyczące sytuacji, w której interesuje nas posiadanie lub nieposiadanie przez dany obiekt pewnego atrybutu. Typowym przykładem takich danych są wyniki ankiet, w których badane osoby mają do wyboru dwie możliwe odpowiedzi: TAK lub NIE. Innym przykładem może być analiza możliwości wykonania określonych zadań przez jednostki administracji terytorialnej, gdzie odpowiedź na postawione pytanie może być formułowana w kategoriach TAK lub NIE. Podobne przykłady można mnożyć, gdyż są one charakterystyczne dla sytuacji, gdy interesuje nas liczba (lub frakcja) jednostek wyróżniających się od pozostałych jednostek danej zbiorowości (populacji) jakimś atrybutem.

Oznaczmy przez N liczność interesującej nas zbiorowości (populacji), wśród której D jednostek odznacza się występowaniem pewnego atrybutu. Załóżmy, że przedmiotem analizy systemowej jest zbadanie celowości podjęcia pewnej decyzji, która to celowość uzależniona jest od liczby D opisującej daną populację. Dla ustalenia uwagi przyjmijmy, że podjęcie interesującej nas decyzji można uznać za wskazane, gdy liczba wyróżniających się danym atrybutem jednostek nie przekracza pewnej wielkości granicznej d_0. Interesuje nas, oczywiście, przypadek, gdy liczba D nie jest nam znana. Najprostszym, ale na pewno nie najtańszym, sposobem rozwiązania problemu byłoby przebadanie całej populacji w celu określenia liczby jej jednostek posiadających interesujący nas atrybut. Takie rozwiązanie jest możliwe praktycznie tylko wtedy, gdy posiadamy pełną wiedzę o danej zbiorowości jednostek, a ponadto mamy wystarczająco dużo czasu i środków na przeprowadzenie wymaganej analizy. Jednakże w przypadku gdy analiza stanu poszczególnych jednostek rozpatrywanej zbiorowości wymaga dużych nakładów pracy, przebadanie wszystkich N jednostek tej zbiorowości w stosunkowo krótkim czasie może być nawet niewykonalne. Do odpowiedzi na nurtujące nas pytanie dotyczące celowości podjęcia rozpatrywanej decyzji musimy więc wykorzystać metody statystyczne, pozwalające na wyciąganie uzasadnionych wniosków na podstawie wyników badań pewnej próby reprezentatywnej dla danej zbiorowości (zagadnienie reprezentatywności badanej próby przedstawione jest w sposób przystępny w książce [3]).

Powyższe rozważania teoretyczne zilustrujmy następującym przykładem. Interesuje nas wprowadzenie pewnego elementu reformy finansowania szkolnictwa średniego, realizowanego na poziomie powiatu. Przyjmijmy, że wprowadzenie tego elementu można uznać za celowe, gdy może być on wprowadzony w przynajmniej 90% wszystkich powiatów ziemskich. Dla uproszczenia przyjmijmy, że powstaje $N = 300$ powiatów ziemskich i interesuje nas liczba powiatów odznaczających się występowaniem atrybutu *"brak możliwości wprowadzenia rozpatrywanego elementu reformy systemu"*. Postawione wymaganie można sformułować formalnie w sposób następujący: liczba D powiatów, w których nie można wprowadzić danego elementu reformy nie powinna przekraczać wielkości $d_0 = 30$.

Załóżmy, że nie dysponujemy środkami pozwalającymi na przebadanie wszystkich 300 powiatów, a możliwe jest przebadanie próby o liczności n jednostek. Wnioski będziemy wyciągać na podstawie zaobserwowanej w tej próbie liczby jednostek k, dla których stwierdzono występowanie interesującego nas atrybutu. Jest rzeczą oczywistą, że wynik takiego badania będzie wynikiem losowym. Jeżeli liczność badanej próby spełnia warunek $n \leq D$, to w skrajnych przypadkach możemy zaobserwować $k = 0$ albo $k = n$ jednostek odznaczających się występowaniem danego atrybutu.

Oznaczmy przez K losową liczbę jednostek z danym atrybutem, które obserwujemy wśród n badanych jednostek. Rozkład prawdopodobieństwa zmiennej losowej K jest w opisanym przez nas przypadku rozkładem hipergeome-

k	$D = 20$	$D = 30$	$D = 40$	$D = 50$
0	0.1129	0.0356	0.0107	0.0031
1	0.2700	0.1331	0.0558	0.0209
2	0.2951	0.2313	0.1361	0.0670
3	0.1960	0.2488	0.2071	0.1345
4	0.0885	0.1858	0.2211	0.1904
5	0.0289	0.1025	0.1761	0.2025

Tabela 1: Prawdopodobieństwa wyników badania

trycznym o funkcji prawdopodobieństwa

$$P(K = k|N, D, n) = \frac{\binom{D}{k}\binom{N-D}{n-k}}{\binom{N}{n}}, \quad k = \max(0, n+D-N), ..., \min(D, n) \tag{1}$$

Losowość wyniku badania próby zilustrujmy na przykładzie rozpatrywanego przez nas problemu wprowadzania reformy finansowania szkolnictwa średniego. Niech $N = 300$ oraz $n = 30$. Oznacza to, że liczba powiatów, w których badamy czy możliwe jest wprowadzenie danego elementu reformy stanowi 10% całej interesującej nas zbiorowości. W Tablicy 1 podane zostały prawdopodobieństwa zaobserwowania w tej próbie k jednostek odznaczających się interesującym nas atrybutem dla czterech przypadków: $D = 20$, $D = 30$, $D = 40$ oraz $D = 50$. Pierwszy przypadek oznacza, że liczba powiatów, w których nie można wprowadzić danego elementu reformy jest znacznie niższa od przyjętej wartości krytycznej. W drugim przypadku liczba ta jest równa wartości krytycznej, a w pozostałych wyraźnie większa.

Jak łatwo zauważyć, w przypadku gdy rzeczywista liczba posiadających interesujący nas atrybut jednostek jest znacznie większa od przyjętej wartości krytycznej d_0, istnieje spore prawdopodobieństwo zaobserwowania w badanej próbie niewielkiej liczby takich jednostek. Z drugiej strony, jeżeli liczba takich jednostek jest wyraźnie mniejsza od liczby krytycznej (np. gdy $D = 20$), prawdopodobieństwo zaobserwowania w próbie o liczności $n = 30$ jednostek więcej niż trzech takich przypadków jest większe od 10%.

Zauważmy teraz, że postawione wymaganie d_0 jest równoważne wymaganiu na frakcję $w_0 = d_0/N$ jednostek posiadających interesujący nas atrybut. Z kolei, obserwując w próbie k takich jednostek, obserwujemy frakcję $w = k/n$. Z teorii statystyki wynika, że zaobserwowana w próbie wartość w jest realizacją nieobciążonego estymatora $\hat{w}$ nieznanej frakcji $W = D/N$. Tak więc, gdy chcemy na podstawie wyników badania próby o liczności n jednostek ocenić nieznaną liczbę D, to powinniśmy skorzystać z jej estymatora $\hat{D}$ określonego zależnością

$$\hat{D} = \hat{w}N = (K/n)N \tag{2}$$

Niestety, podejmowanie decyzji poprzez porównanie zaobserwowanej realizacji estymatora $\hat{D}$ z wymaganiem d_0, lub - równoważnie - porównywanie zaobserwowanej w próbie frakcji w z wymaganiem w_0, nie gwarantuje akceptowalnego poziomu prawdopodobieństwa popełnienia błędnej decyzji. Dane przedstawione w Tablicy 1 świadczą o tym w sposób jednoznaczny. Konieczne jest więc takie zaprojektowanie statystycznej reguły decyzyjnej, by prawdopodobieństwa podjęcia błednej decyzji było wystarczająco małe.

Niestety, nie istnieje prosta metoda pozwalająca na analityczne rozwiązanie postawionego powyżej zadania w każdym przypadku. Możemy jednak do tego celu wykorzystać procedurę zaprojektowaną dla celów statystycznej kontroli odbiorczej i podanej w Polskiej Normie PN-ISO 2859-2 [6], będącej tłumaczeniem odpowiedniej normy międzynarodowej. W normie tej podane są plany badania przeznaczone do odbiorczej kontroli jakości, które można wykorzystać w rozpatrywanym przez nas przypadku. Aby zastosować tę metodę w praktyce zauważmy, że

- nie istnieje statystyczna procedura decyzyjna pozwalająca w sposób bezbłędny ocenić czy spełnione jest wymaganie dotyczące interesującej nas frakcji W;
- prawdopodobieństwa podjęcia błędnych decyzji (tzn. przyjęcia hipotezy $W \leq w_0$, gdy słuszna jest hipoteza przeciwna lub przyjęcia hipotezy $W > w_0$ w przypadku gdy rzeczywista frakcja jednostek posiadających interesujący nas atrybut jest mniejsza od wartości krytycznej) można zmniejszyć wyłącznie przez zwiększenie liczności próby;
- jeżeli dopuścimy pewne małe prawdopodobieństwo błędnego odrzucenia hipotezy $W \leq w_0$, to musimy zgodzić się na przyjęcie tej hipotezy (z założonym małym prawdopodobieństwem) w przypadku granicznym, gdy $W = \acute{w} > w_0$.

Żeby skorzystać z normy PN-ISO 2859-2 powinniśmy wybrać z ciągu podanych w niej wartości parametru LQ (*Limiting Quality*) taką, by $LQ > w_0$. Jeżeli kilka wartości LQ spełnia powyższy warunek, to przyjęcie najmniejszej z nich wiąże się z najmniejszym prawdopodobieństwem błędnego odrzucenia hipotezy $W \leq w_0$. Z kolei, utożsamiamy liczność badanej populacji z wymienioną w normie licznością kontrolowanej partii, a następnie w podanej w normie tablicy znajdujemy wymaganą liczność próby n oraz liczbę krytyczną

Liczność partii		Jakość graniczna (LQ) w %				
		...	8,0	12,5	20	....
... ...	n Ac	... ...	... ...	... ...	... ..	... ...
151 do 280	n Ac	... ...	28 0	20 0	20 1	... ...
281 do 500	n Ac	... ...	32 0	32 1	20 1	... ...
501 do 1200	n Ac	... ...	50 1	32 1	32 3	... ...
... ...	n Ac	... ...	... ...	... ...	... ...	... ...

Tabela 2: Plany badania normy PN-ISO 2859-2

Ac. Jeżeli zaobserwowana w próbie liczba jednostek posiadających interesujący nas atrybut k jest nie większa od wartości krytycznej Ac, to istnieje duże prawdopodobieństwo, że rzeczywista wartość frakcji W nie przekracza wartości krytycznej w_0. Fragment podanej w normie PN-ISO 2859-2 tablicy z zalecanymi planami badań podano w Tablicy 2.

W rozpatrywanym przez nas przykładzie liczność populacji wynosi 300 jednostek, a graniczna frakcja jednostek posiadających dany atrybut wynosi $w_0 = 10\%$. W Tablicy 2 należy więc wybrać plan badania odpowiadający kontroli odbiorczej partii o liczności od 281 do 500 jednostek i parametrze LQ większym od 10%. Powyższy warunek spełnia kilka planów zamieszczonych w odpowiedniej tablicy normy PN-ISO 2859-2, z których rozpatrzmy dwa, podane w Tablicy 2. Są to plany: ($n = 32$, $Ac = 1$) - odpowiadający parametrowi $LQ = 12,5\%$, oraz plan ($n = 20$, $Ac = 1$) - odpowiadający parametrowi $LQ = 20\%$. Przyjmując pierwszy z tych planów ograniczamy prawdopodobieństwo błędnego przyjęcia hipotezy $W \leq w_0$ gdy rzeczywista wartość liczby D przekracza postawione przez nas wymaganie (w naszym przypadku $d_0 = 30$). Niestety, w takim przypadku zwiększamy również prawdopodobieństwo błędnego odrzucenia tej hipotezy, gdy w rzeczywistości postawiony przez nas warunek będzie spełniony. Drugi z wymienionych powyżej planów badania zapewnia małe prawdopodobieństwo błędnego odrzucenia hipotezy $W \leq w_0$, gdy w rzeczywistości postawiony przez nas warunek będzie spełniony. Z drugiej jednak strony, hipotezę tę możemy przyjąć z dużym prawdopodobieństwem nawet wtedy, gdy liczba jednostek posiadających interesujący nas atrybut jest znacznie większa od przyjętej przez nas wartości granicznej. W Tablicy 3 podane zostały przykładowe wyniki obliczeń prawdopodobieństwa

Plan	D						
badania	10	20	25	30	40	50	60
$n = 32,\ Ac = 1$	0.711	0.346	0.225	0.141	0.052	0.017	0.005
$n = 20,\ Ac = 1$	0.862	0.608	0.488	0.383	0.223	0.123	0.063

Tabela 3: Prawdopodobieństwa przyjęcia hipotezy dla wybranych planów badania

przyjęcia hipotezy $W \leq w_0$ dla obu interesujących nas planów badania i różnych wartości liczby D.

Wybór właściwego planu badania zależy od tego, jakiego typu błędu nie chcemy popełniać. Jeżeli postawione wymaganie jest rzeczywiście bardzo istotne wybierzemy plan badania odpowiadający parametrowi $LQ = 12,5\%$, godząc się z dużym prawdopodobieństwem podjęcia błędnej decyzji, gdy w rzeczywistości warunek ten jest spełniony. Jeżeli jednak postawiony warunek nie musi być spełniony w sposób rygorystyczny, lepszym rozwiązaniem wydaje się przyjęcie drugiego z rozpatrywanych planów badania. W przypadku, gdy wyznaczone prawdopodobieństwa podjęcia błędnych decyzji są zbyt duże, należy poszukać innych rozwiązań, zwiększając w znaczny sposób liczność badanej próby n.

W przypadku gdy konieczne staje się indywidualne zaprojektowanie planu badania możliwe jest skorzystanie ze związku pomiędzy przedziałami ufności dla nieznanych parametrów rozkładów prawdopodobieństwa a statystycznymi procedurami decyzyjnymi. Jeżeli, na przykład, dla nieznanego parametru D w rozkładzie hipergeometrycznym wg (1) określimy jednostronny przedział ufności $[k, \overline{D}]$ na poziomie ufności β, to porównując wartość $\overline{D}$ z wymaganiem d_0 weryfikujemy statystycznie hipotezę $D \leq d_0$ na poziomie istotności $\alpha = 1-\beta$. Jeżeli $\overline{D} \leq d_0$, to przyjmujemy hipotezę $D \leq d_0$. W przeciwnym przypadku, tzn. gdy $\overline{D} > d_0$, nie mamy podstaw do przyjęcia tej hipotezy.

Wyznaczenie przedziału ufności dla D nie jest proste i wymaga wykonania wielu żmudnych obliczeń z wykorzystaniem wzoru (1). W praktyce jest to możliwe wyłącznie przy wykorzystaniu komputera. Możemy jednak skorzystać z aproksymacji rozkładu hipergeometrycznego innymi rozkładami. Jedną z tych aproksymacji zaproponował Wise [7], który do obliczania wartości dystrybuanty w rozkłądzie hipergeometrycznym wykorzystał następującą aproksymację rozkładem dwumianowym

$$P(K \leq k|N, D, n) \approx \sum_{i=0}^{k} \binom{n}{i} \acute{p}^i (1-\acute{p})^{n-i} \tag{3}$$

gdzie

$$\acute{p} = \frac{2D - k}{2N - n + 1} \tag{4}$$

Korzystając z powyższej aproksymacji możemy wyznaczyć przedział ufności dla parametru $\acute{p}$ korzystając ze wzorów i tablic podanych np. w książce Firkowicza [5] lub tablicach statystycznych Zielińskiego [8], a następnie, wykorzystując (4) wyznaczyć odpowiedni przedział ufności dla D. Jeżeli jednak nieznana wartość D oraz liczność próbki są duże, tak by oczekiwana liczba jednostek w próbie, które posiadają interesujący nas atrybut była wystarczająco duża (zwykle większa od pięciu), to możemy skorzystać z aproksymacji rozkładu dwumianowego rozkładem normalnym. Korzystając z metody opisanej m.in. w książce Barnetta [3] oraz wzoru (4) otrzymujemy następujący przybliżony wzór na górną granicę jednostronnego przedziału ufności dla D na poziomie ufności β.

$$\overline{D} \approx \frac{\overline{p}(2N - n + 1) + k}{2} \tag{5}$$

gdzie

$$\overline{p} = \frac{n(2k + y_\beta^2) + y_\beta\sqrt{n^2 y_\beta^2 + 4kn(n-k)}}{2n(y_\beta^2 + n)} \tag{6}$$

przy czym y_β jest kwantylem rzędu β w standaryzowanym rozkładzie normalnym, stabelowanym w praktycznie wszystkich książkach poświęconych rachunkowi prawdopodobieństwa i statystyce matematycznej.

Na przykład, niech liczność próby wynosi $n = 80$ jednostek, a w próbie zaobserwowano $k = 5$ jednostek posiadających interesujący nas atrybut. Korzystając ze wzorów (5) i (6) możemy wyznaczyć górną granicę jednostronnego przedziału ufności na poziomie $\beta = 0.90$, która w tym przypadku wynosi $\overline{D} = 31$. Oznacza to, że na tym poziomie ufności nie możemy dać gwarancji, że wymaganie $D \leq 30$ jest spełnione, chociaż wszystko raczej na to wskazuje.

Powyższe rozważania oraz analiza podanych przykładów sugerują, że w przypadku podejmowania decyzji na podstawie rezultatów badań statystycznych należy bardzo wnikliwie przeanalizować możliwość podjęcia błędnych decyzji. Ogólnie rzecz ujmując, im mniej liczna jest badana próba losowa, tym wynik badania musi wyraźniej wskazywać na słuszność podejmowanej decyzji.

3 Bayesowskie decyzje statystyczne przy danych wg klasyfikacji alternatywnej

Wiele problemów znajdujących się w obszarze zainteresowań badań systemowych dotyczy zjawisk losowych, które zachodzą stosunkowo rzadko lub też dotyczą stosunkowo małej liczby obiektów. Typowym przykładem mogą tu być katastrofy naturalne, takie jak powodzie, trzęsienia ziemi lub huragany. Jest rzeczą oczywistą, że podejmowane są kroki mające na celu ograniczenie skutków takich katastrof, przy czym podejmowane działania muszą mieć zazwyczaj bardzo szeroki zasięg. Wspomaganie osób, które muszą podjąć odpowiednie decyzje w tym zakresie jest typowym zadaniem badań systemowych. Istotnym parametrem wykorzystywanym do modelowania problemu decyzyjnego jest prawdopodobieństwo wystąpienia katastrofy w ustalonym przedziale czasu. Na przykład, może to być prawdopodobieństwo wystąpienia w danym miejscu dużej (tzn. o zadanym przyroście poziomu wody w rzece) wiosennej powodzi. Na pierwszy rzut oka prawdopodobieństwo to można ocenić na podstawie analizy wieloletnich zapisów historycznych. Należy jednak brać pod uwagę zmiany środowiska geograficznego, takie jak np. budowę zbiornika retencyjnego. W przypadku wybudowania takiego zbiornika powyżej miejsca, dla którego ocenia się prawdopodobieństwo wystąpienia powodzi, bezpośrednie wykorzystanie danych historycznych z okresu poprzedzającego budowę takiego zbiornika może prowadzić do błędnej oceny prawdopodobieństwa wystąpienia powodzi. Z kolei, okres po wybudowaniu zbiornika może być zbyt krótki, by wiarogodnie ocenić prawdopodobieństwo wystąpienia powodzi. Załóżmy na przykład, że przez pięć kolejnych lat po wybudowaniu danego zbiornika nie zaobserwowano w danym miejscu powodzi o katastrofalnych rozmiarach. Punktowa ocena prawdopodobieństwa wystąpienia takiej powodzi wynosi, oczywiście, $p = 0$; jednakże jednostronny 95% przedział ufności dla tej oceny wynosi $(0, 0.522)$, co można łatwo odczytać z tablicy zamieszczonej np. w książce Firkowicza [5]. Niestety, jest to największa możliwa dokładność oceny tego prawdopodobieństwa, jeżeli do tego celu wykorzystamy wyłącznie ostatnie dane statystyczne.

Sytuacja badacza przeprowadzającego analizę systemową danego problemu może jednak ulec radykalnej poprawie, gdy w swoich analizach wykorzysta inne źródła informacji, a w tym subiektywne oceny ekspertów. Aby tego dokonać konieczne jest zastosowanie innego podejścia do analizy danych statystycznych, zwanego podejściem bayesowskim. W podejściu bayesowskim zakładamy, że interesujące nas prawdopodobieństwo nie jest stałym, choć nieznanym, parametrem, ale zmienną losową o zadanym rozkładzie prawdopodobieństwa, zwanym *rozkładem a priori* $F(w)$ nieznanego prawdopodobieństwa W. Przy pomocy rozkładu a priori $F(w)$ opisujemy naszą, zwykle subiektywną, wiedzę o wartościach jakie może przyjąć interesujące nas prawdopodobieństwo. Na przykład, jeżeli jesteśmy w znacznym stopniu przekonani, że oceniane prawdopodobieństwo może przyjąć wartości bliskie jakiejś liczby, to za rozkład a priori możemy przyjąć rozkład o gęstości prawdopodobieństwa skupionej wokół spodziewanej przez nas wartości. Z kolei, gdy nic nie możemy powiedzieć o spodziewanej wartości ocenianego prawdopodobieństwa, to jako rozkład $F(w)$

możemy przyjąć rozkład równomierny na przedziale $[0,1]$.

Rozpatrzmy najprostszy przypadek bayesowskiej analizy decyzyjnej i załóżmy, że analizie podlega k możliwych decyzji, przy czym każda z nich opisana jest przez pewną funkcję strat (lub zysków) $C_i(W)$, $i = 1,\ldots,k$. Znając wartość prawdopodobieństwa W moglibyśmy podjąć decyzję wiążącą się z minimalnymi stratami lub maksymalnym zyskiem. Ponieważ wartości W zazwyczaj nie znamy, możemy jedynie obliczyć oczekiwaną wartość funcji strat dla danego rozkładu rozkładu a priori $F(w)$. Wielkość tę wyznaczamy ze wzoru

$$R_i = \int_{-\infty}^{\infty} C_i(w)dF(w), \quad i = 1,\ldots,k \tag{7}$$

i nazywamy *ryzykiem a priori* związanym z podjęciem danej decyzji. Tak więc, możemy wybrać właściwą decyzję bazując wyłącznie na naszej dotychczasowej wiedzy, opisanej rozkładem a priori $F(w)$. W pewnych przypadkach może to wystarczać, ale zazwyczaj konieczne jest przeanalizowanie dodatkowych danych statystycznych, które mogą potwierdzić lub też zakwestionować prawidłowość przyjętej decyzji. Bayesowskie podejście do analizy danych statystycznych umożliwia połączenie dotychczasowej informacji a priori z informacją zawartą w nowych danych statystycznych.

Załóżmy, że możemy wykorzystać wyniki n obserwacji (w szczególnym przypadku może być $n = 1$), z których x odznacza się występowaniem interesującego nas atrybutu. Niech zmienna losowa X oznacza liczbę takich obserwacji w badanej próbie losowej. Jeżeli prawdopodobieństwo obserwacji odznaczającej się występowaniem danego atrybutu wynosi w, to rozkład zmiennej losowej X, określony funkcją prawdopodobieństwa $g(x|w)$ jest rozkładem dwumianowym. Korzystając ze znanego w teorii prawdopodobieństwa twierdzenia Bayesa możemy określić warunkowy rozkład ocenianego prawdopodobieństwa o gęstości $f(w|x)$ dla warunku, że w próbie losowej o liczności n zaobserwowano x jednostek odznaczających się występowaniem interesującego nas atrybutu.

$$f(w|x) = \frac{g(x|w)f(w)}{\int_0^1 g(x|w)dF(w)} \tag{8}$$

Rozkład opisany funkcją prawdopodobieństwa (8) nazywamy *rozkładem a posteriori* ocenianego prawdopodobieństwa W. Znając rozkład a posteriori możemy wyznaczyć *ryzyka a posteriori* zastępując we wzorze (7) rozkład a priori o gęstości $f(w)$ rozkładem a posteriori o gęstości $f(w|x)$. Decyzja, dla której ryzyko a posteriori jest minimalne nazywana jest *decyzją bayesowską*, a odpowiadające jej ryzyko nazywane jest *ryzykiem bayesowskim.* Ogólne podejście do zagadnienia określania decyzji bayesowskich przedstawione jest m.in. w książce [4].

Wykorzystywany w powyższej analizie rozkład a priori $F(w)$ może być dowolnym rozkładem prawdopodobieństwa określonym na przedziale $[0,1]$. W praktyce korzystamy zazwyczaj z tzw. *sprzężonych rozkładów a priori*. Jeżeli wyniki obserwacji w próbie opisane są rozkładem prawdopodobieństwa $g(x|w)$, to rozkład a priori o gęstości $f(w)$ należy do rodziny sprzężonych rozkładów a priori, gdy wyznaczony wg (8) rozkład a posteriori również należy do tej samej rodziny. Dla obserwacji opisanych rozkładem dwumianowym rodziną sprzężonych rozkładów a priori są *rozkłady beta* o funkcji gęstości

$$f(w) = \begin{cases} \frac{\Gamma(u+v)}{\Gamma(u)\Gamma(v)} w^{u-1}(1-w)^{v-1}, & 0 \leq w \leq 1,\ u, v > 0 \\ 0, & \text{w pozostałym przypadku} \end{cases} \tag{9}$$

Rodzina rozkładów beta jest bogata i obejmuje rozkłady prawdopodobieństwa o funkcjach gęstości różnej postaci. W zależności od wartości parametrów (u, v) mogą to być gęstości: malejące, rosnące, jednomodalne z maksimum oraz jednomodalne z minimum. Szczególnym przypadkiem rozkładu beta ($u = 1, v = 1$) jest rozkład równomierny na przedziale $[0, 1]$. Tak więc rozkład beta może być łatwo wykorzystany do przedstawienia wiedzy a priori o możliwych wartościach nieznanego prawdopodobieństwa W.

Jeżeli rozkład a priori o gęstości $f(w)$ jest rozkładem beta o parametrach (u, v) i gęstości wg (9), to rozkład a posteriori $f(w|x)$ jest także rozkładem beta o parametrach $(u+x, v+n-x)$. Powyższą własność możemy wykorzystać przy sekwencyjnym napływie danych empirycznych. Rozkład a posteriori uzyskany w wyniku uwzględnienia wyniku obserwacji pierwszej próby (w szczególnym przypadku będącej pojedynczą obserwacją) możemy wykorzystać jako nowy rozkład a priori wykorzystywany na następnych etapach analizy bayesowskiej.

Rozkład a posteriori o gęstości $f(w|x)$ możemy również wykorzystać do wyznaczenia *bayesowskiego estymatora* nieznanego prawdopodobieństwa W. Można wykazać, patrz [4], że w pewnych przypadkach wartość oczekiwana w rozkładzie a posteriori o gęstości $f(w|x)$ jest optymalnym estymatorem bayesowskim prawdopodobieństwa W. Jeżeli rozkład a posteriori jest rozkładem beta o parametrach $(u + x, v + n - x)$, to bayesowski estymator ocenianego prawdopodobieństwa dany jest zależnością

$$\hat{w}_B = \frac{u + x}{u + v + n} \tag{10}$$

Powyższy estymator możemy wykorzystać wtedy, gdy posiadamy informację a priori o możliwych wartościach ocenianego prawdopodobieństwa oraz pewne informacje empiryczne. Stosując powyższe podejście należy jednak pamiętać, że uzyskana ocena może w znacznej mierze zależeć od przyjętej przez nas informacji a priori. Jeżeli gęstość rozkładu a priori jest skoncentrowana wokół pewnej wartości, to ocena różniąca się od tej wartości może być uzyskana po uwzględnieniu wyniku badań próby o wystarczająco dużej liczności.

4 Podsumowanie

Przedstawione w niniejszej pracy proste metody statystyczne mogą znaleźć szerokie zastosowanie w typowych analizach badań systemowych. Dotyczą one najprostszego przypadku danych o postaci alternatywnej. W przypadku, gdy dane empiryczne są innego typu, a także gdy decyzje dotyczą wielkości opisywanych na skali ciągłej (np. wysokość nakładów) konieczne jest wykorzystanie innych modeli matematycznych. Statystyka matematyczna może dostarczyć właściwych narzędzi w większości spotykanych w praktyce przypadków. Użycie niewłaściwych metod w analizie problemów odznaczających się występowaniem niepewności o charakterze losowym może prowadzić do podejmowania błędnych decyzji. Zastosowanie opisanych w tej pracy metod statystycznych pozwala ograniczyć prawdopodobieństwo zaistnienia takiej sytuacji.

Referencje

[1] *Analiza systemowa i jej zastosowania.* Materiały z konferencji "Dni Międzynarodowego Instytutu Stosowanej Analizy Systemowej", Warszawa, 20-21 kwietnia, 1993. Red. Jan W. Owsiński, IBS PAN, Warszawa 1993.

[2] *Analiza systemowa - podstawy i metodologia.* Red. W. Findeisen. PWN, Warszawa 1985.

[3] Barnett, V.: *Elementy teorii pobierania prób.* PWN, Warszawa 1982.

[4] DeGroot, M.H.: *Optymalne decyzje statystyczne.* PWN, Warszawa 1981.

[5] Firkowicz, S.: *Statystyczne badanie wyrobów.* WNT, Warszawa 1970.

[6] PN-ISO 2859-2 (1996). *Procedury kontroli wyrywkowej metodą alternatywną. Plany badania na podstawie jakości granicznej (LQ) stosowane podczas kontroli partii izolowanych.*

[7] Wise, M.E.: *A quickly convergent expansion for cumulative hypergeometric probabilities. Direct and inverse.* Biometrika, vol. 41, 1954, 317-329.

[8] Zieliński, R.: *Tablice statystyczne.* PWN, Warszawa 1972.

ISBN 83-85847-34-0