E. FIDELIS i J. ODERFELD (Warszawa)

DWUSTOPNIOWA KONTROLA Z UWZGLĘDNIENIEM BŁĘDÓW POMIARU

Jeżeli ostateczny odbiór produktu opiera się na postępowaniu statystycznym uzgodnionym pomiędzy dostawcą a odbiorcą, to w wielu przypadkach odbiorca dla ostrożności poprzedza odbiór właściwy wewnętrzną kontrolą statystyczną.

Różne warianty i aspekty tego zagadnienia były niejednokrotnie referowane w piśmiennictwie statystycznym. Np. w [1] autorzy zakładają, że wyniki kontroli wewnętrznej i ostatecznej są niezależne. Wspominamy o tym tylko dla porządku, gdyż zajmujemy się innym zagadnieniem.

Przyjmujemy mianowicie, że produkt sztukowy jest scharakteryzowany przez cechę X o rozkładzie normalnym. Cechę tę sprawdza się metodą mierniczą, która jest obciążona błędem przypadkowym Y o rozkładzie normalnym. X i Y są z założenia niezależne.

Na każdym egzemplarzu mierzy się jeden lub dwa razy cechę X . Niech wynik pierwszego pomiaru będzie m_1 , a drugiego m_2 . Oznaczmy przez A_1 i A_2 odpowiednio dwa domknięte przedziały ustalone z góry. Każdy egzemplarz zalicza się do klasy I lub II według następującego schematu:

Wynik pierwszego pomiaru	Wynik drugiego pomiaru	Decyzja
$m_1 \notin A_1$	nie mierzy się	zaliczyć do II klasy
$m_1 \in A_1$	$m_2 \notin A_2$	zaliczyć do II klasy
	$m_2 \in A_2$	zaliczyć do I klasy

Aby zilustrować celowość takiego postępowania, wyobraźmy sobie, że klasa I oznacza produkt eksportowy, a klasa II produkt przeznaczony na rynek wewnętrzny. Przypuśćmy ponadto, że pierwszy pomiar przeprowadza kontroler fabryczny, a pomiar drugi kontroler obcy.

Przyjmijmy wreszcie, że zaliczenie egzemplarza do II klasy pociąga za sobą dla producenta małe straty, gdy nastąpi przy kontroli fabrycznej, a duże straty — gdy nastąpi przy kontroli zewnętrznej. W tej sytuacji producentowi może opłacać się kontrola fabryczna — do której zresztą nie jest zobowiązany — przed oficjalną kontrolą zewnętrzną.

Opisany schemat postępowania przedstawił (w nieco innej formie) J. Obalski na zebraniu seminarium Działu Zastosowań Przemysłowych odbytym w kwietniu 1960 r., w Instytucie Matematycznym PAN. Ponadto rozpatrzył on niektóre jego konsekwencje i ogłosił swoje rozważania w pracy [2]. Jest tam mowa o wodomierzach, które oczywiście mogą posłużyć za ilustrację i do niniejszej pracy.

Obalski buduje swój model matematyczny następująco: pierwszy etap kontroli opiera się na losowaniu z populacji o znanym rozkładzie, po tym etapie następuje ucięcie populacji i dalej losowanie odbywa się z populacji uciętej. Jest to model intuicyjnie najprostsz, jednakże jego opracowanie wymaga żmudnych obliczeń, szacowania zbieżności szeregow itd.

Autorzy niniejszej noty wybrali odmienną drogę analizy, mianowicie opartą na opracowaniu ogłoszonym przez J. Oderfelda i S. Zubrzyckiego w roku 1954 w pracy [3]. Ponieważ od tej pory powstały liczne tablice i nomogramy rozkładu dwuwymiarowych zmiennych losowych, można było zrezygnować ze żmudnych rachunków i uzyskać wyniki dość ogólne i nadające się do natychmiastowego wykorzystania w praktyce.

Sformalizujemy reguły postępowania używając zwykłych oznaczeń:

$N(\dots)$ — rozkład normalny jednowymiarowej zmiennej losowej o parametrach podanych w nawiasie,

$p(u)$ — gęstość rozkładu zmiennej losowej U ,

$p(u|v)$ — gęstość rozkładu warunkowego zmiennej losowej U , gdy zmienna V przyjęła wartość v ,

$p(u, v)$ — gęstość rozkładu łącznego zmiennych losowych U i V .

Niech X będzie kontrolowaną cechą egzemplarza produktu. Załóżmy, że w nieskończonej populacji wszystkich egzemplarzy, X ma rozkład normalny $N(0, \sigma_1)$, gdzie σ_1 jest stałe i znane. Gdyby wartość średnia była różna od zera, to w dalszych rozważaniach wystarczyłoby wprowadzić trywialne zmiany.

Niech zmienna losowa Y będzie błędem przypadkowym metody sprawdzania. Zmienna ta jest określona w nieskończonej populacji wszystkich wyników pomiaru wykonanych na jednym egzemplarzu produktu. Jest ona z założenia jednakowa dla wszystkich egzemplarzy produktu i normalna $N(0, \sigma_2)$.

Zmienne losowe X i Y są z założenia niezależne.

Niech M_1 i M_2 będą odpowiednio zmiennymi losowymi, które są wynikami pierwszego (fabrycznego) i drugiego (odbiorczego) pomiaru. Odpowiednio m_1 i m_2 niech będą realizacjami tych zmiennych losowych.

Nie zmniejszając ogólności rozważań można zawsze dokonać takiego przekształcenia liniowego, żeby można było egzemplarz produktu uważać za dobry przy pierwszym pomiarze, gdy $|m_1| \leq a_1$, a przy drugim, gdy $|m_2| \leq a_2$. Wielkości a_1 i a_2 oznaczają odpowiednio połowy długości prze-

działów A_1 i A_2 , poprzednio zdefiniowanych. Do drugiego pomiaru dopuszcza się tylko te egzemplarze, które uznano za dobre przy pierwszym pomiarze.

Przed wszystkim zajmiemy się prawdopodobieństwem P_{12} zdarzenia polegającego na tym, że egzemplarz produktu przejdzie pomyślnie przez obie kontrole. Zachodzi tożsamość

$$(1) \quad p(m_1)p(x|m_1) = p(x)p(m_1|x).$$

Występujące w niej gęstości mają następującą postać i znaczenie:

$$(2) \quad \begin{cases} p(m_1) & \text{— gęstość bezwarunkowego rozkładu } N(0, \sqrt{\sigma_1^2 + \sigma_2^2}); \text{ dotyczy zmiennej losowej } M_1 = X + Y \text{ (wynik pierwszego pomiaru);} \\ p(x) & \text{— gęstość bezwarunkowego rozkładu } N(0, \sigma_1); \text{ dotyczy zmiennej losowej } X \text{ (cecha produktu);} \\ p(m_1|x) & \text{— gęstość warunkowego rozkładu } N(x, \sigma_2); \text{ dotyczy zmiennej losowej będącej wynikiem pierwszego pomiaru tych egzemplarzy, u których cecha } X = x; \\ p(x|m_1) & \text{— gęstość warunkowego rozkładu (o postaci na razie nieznanej) zmiennej losowej będącej cechą tych egzemplarzy, które dały } m_1 \text{ jako wynik pierwszego pomiaru.} \end{cases}$$

Z (1) i (2) wynika, że

$$(3) \quad p(x|m_1) \text{ jest gęstością rozkładu } N\left(m_1 \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}, \frac{\sigma_1 \sigma_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right).$$

Rachunek potrzebny do tego wyniku sprowadza się do dodania i uporządkowania wykładników w potęgach liczby e występującej w gęstościach rozkładów normalnych wymienionych w (2).

Gęstość $p(m_2|m_1)$ rozkładu warunkowego drugiego pomiaru tych egzemplarzy, które dały m_1 jako wynik pierwszego pomiaru, znajdujemy uwzględniając, że oba pomiary robi się na tym samym egzemplarzu, dla którego cecha $X = x$. Wskutek tego

$$(4) \quad p(m_2|m_1) \text{ jest gęstością rozkładu } N\left(m_1 \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}, \sqrt{\frac{\sigma_1^2 \sigma_2^2}{\sigma_2^2 + \sigma_2^2} + \sigma_2^2}\right).$$

Teraz można już obliczyć gęstość $p(m_1, m_2)$ łącznego rozkładu dwuwymiarowej zmiennej losowej (M_1, M_2) będącej wynikiem pierwszego i drugiego pomiaru. Opieramy się tu na tożsamości

$$(5) \quad p(m_1, m_2) = p(m_1)p(m_2|m_1).$$

Rachunek polega znowu na dodaniu i uporządkowaniu wykładników w potęgach liczby e występującej w wyrażeniach po prawej stronie tożsamości (5). Okazuje się ostatecznie, że zmienna losowa (M_1, M_2) ma

rozkład normalny o obu średnich równych zeru, obu odchyleniach średnich równych $\sqrt{\sigma_1^2 + \sigma_2^2}$ i współczynniku korelacji

$$(6) \quad \varrho = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}.$$

Wprowadźmy unormowane długości przedziałów:

$$(7) \quad a_1 = \frac{a_1}{\sqrt{\sigma_1^2 + \sigma_2^2}}, \quad a_2 = \frac{a_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}.$$

Wtedy prawdopodobieństwo P_{12} ostatecznego uznania egzemplarza wyboru za dobry oblicza się jako

$$(8) \quad P_{12} = \frac{1}{2\pi\sqrt{1-\varrho^2}} \int_{-a_1}^{a_1} \int_{-a_2}^{a_2} \exp\left\{-\frac{1}{2(1-\varrho^2)}(u^2 - 2\varrho uv + v^2)\right\} du dv.$$

Tak więc zależności (6), (7) i (8) stanowią rozwiązanie postawionego zagadnienia: mając σ_1 , σ_2 , a_1 , a_2 oblicza się kolejno ϱ , a_1 , a_2 i wreszcie P_{12} .

Dla wygody obliczyliśmy⁽¹⁾ w tabeli 1 niektóre wartości $P_{12}(a_1, k, \varrho)$, gdzie $k = a_2/a_1$. Korzystaliśmy przy tym z [7]. Przy grubszej interpolacji można stosować wykresy na rysunku 1, który jest odpowiednikiem tabeli 1.

TABELICA 1

$$P(a_1, k, \varrho) = \frac{1}{2\pi\sqrt{1-\varrho^2}} \int_{-a_1}^{a_1} \int_{-a_2}^{a_2} \exp\left\{-\frac{1}{2(1-\varrho^2)}(u^2 - 2\varrho uv + v^2)\right\} du dv$$

$\varrho \backslash k$ a_1	0,7			0,8			0,9			0,95		
	1,1	1,2	1,3	1,1	1,2	1,3	1,1	1,2	1,3	1,1	1,2	1,3
1,0	0,561	0,587	0,606	0,587	0,614	0,627	0,620	0,637	0,648	0,648	0,659	0,668
1,5	0,805	0,827	0,838	0,823	0,837	0,842	0,833	0,846	0,850	0,847	0,853	0,856
2,0	0,934	0,940	0,947	0,938	0,944	0,948	0,943	0,947	0,950	0,949	0,950	0,951
2,5	0,982	0,984	0,986	0,983	0,985	0,986	0,984	0,985	0,986	0,985	0,986	0,986

Z zależności (6), (7) i (8) można również z łatwością odpowiedzieć na wiele innych pytań praktycznych, na przykład:

1. Znane są wielkości σ_1 i σ_2 charakteryzujące rozkład cechy X i błąd metody sprawdzania, ponadto znana jest wielkość a_2 określająca wymagania odbiorcy. Jak dobrać wielkość a_1 , aby mieć zagwarantowane prawdopodobieństwo P_{12} pomyślnego przejścia przez próbę odbiorczą? Na to pytanie odpowiadamy obliczając kolejno ϱ , a_2 , a_1 , a_1 .

⁽¹⁾ Znaczną część obliczeń wykonała p. K. Olsztyńska.

Rys. 1. Wykresy $P_{12}(\alpha_1; k, \rho)$

Dalsza seria trzech naturalnych pytań producenta wymaga dodatkowych rozważań, zresztą zupełnie prostych. Przyjmujemy tu, że $\sigma_1, \sigma_2, a_1, a_2$ są znane, czyli że producent zna swoje możliwości produkcyjne, dokładność pomiaru i przepisy kontroli zarówno odbiorczej jak fabrycznej. Pytania i odpowiedzi są następujące:

2. Jaka będzie po pierwszym sprawdzeniu przeciętna frakcja P_1 egzemplarzy uznanych za dobre?

Odpowiedź jest

$$P_1 = 2\Theta(\alpha_1) = \frac{2}{\sqrt{2\pi}} \int_0^{\alpha_1} e^{-t^2/2} dt,$$

gdzie α_1 jest określone przez (7).

3. Jaka będzie po pierwszym sprawdzeniu przeciętna frakcja Q'_1 sztuk, które zdyskwalifikowano, mimo że wartość $x \in \langle -a_1, a_1 \rangle$.

Najpierw znajdziemy gęstość $p(x, m_1)$ łącznego rozkładu dwuwymiarowej zmiennej losowej (X, M_1) :

$$p(x, m_1) = p(x)p(m_1|x).$$

Gęstości występujące po prawej stronie są określone przez (2). Elementarny rachunek wykazuje, że zmienna losowa (X, M_1) ma rozkład normalny. Po unormowaniu tej zmiennej znajduje się ostatecznie, że

$$(10) \quad Q'_1 = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{|u| < a_2/\sigma_1} \int_{|v| > a_1} \exp\left\{-\frac{1}{2(1-\rho^2)}(u^2 - 2\rho_1 uv + v^2)\right\} du dv,$$

gdzie a_1 jest określone przez (7), a $q_1 = \sigma_1 / \sqrt{\sigma_1^2 + \sigma_2^2}$. Wartości Q'_1 można znaleźć korzystając z tablicy 1 i rysunku 1.

4. Jaka będzie po drugim sprawdzeniu przeciętna frakcja P_2 egzemplarzy zakwalifikowanych jako dobre, w odniesieniu do liczby egzemplarzy przedstawionych do drugiego sprawdzenia?

Jeśli przed pierwszym sprawdzeniem było N egzemplarzy, to do drugiego sprawdzenia zakwalifikuje się przeciętnie $P_1 N$ sztuk. Po obu sprawdzeniach zakwalifikowanych będzie przeciętnie $P_{12} N$ sztuk. Wskutek tego

$$(11) \quad P_2 = \frac{P_{12}}{P_1},$$

gdzie P_{12} i P_1 są określone odpowiednio przez (8) i (9).

Znajomość P_{12} , P_1 , Q'_1 , P_2 , ewentualnie również jednostkowych kosztów związanych ze zdarzeniami scharakteryzowanymi przez te wielkości, może depomóc producentowi w wyborze właściwej taktyki; może mu na przykład ułatwić ocenę, czy warto stosować dodatkową kontrolę fabryczną, a jeśli tak, to jak ustalić wielkość a_1 .

Pokażemy to na przykładzie.

Niech $\sigma_1 = 1$, $\sigma_2 = 0,23$, $a_2 = 1,33$.

Obliczamy kolejno:

$$Z (6) \quad q = 0,95,$$

$$z (7) \quad a_2 = 1,3.$$

Przyjmijmy $a_1 = 1$, czyli z (7) $a_1 = 1,023$. Wtedy $k = 1,3$ i z tablicy 1, dla $q = 0,95$, $k = 1,3$, $a_1 = 1$ odczytujemy $P_{12} = 0,668$; z (9) znajdujemy $P_1 = 2\Theta(1) = 0,683$ i z (11) obliczamy $P_2 = 0,988$.

Pozostawiając niezmiennione σ_1 , σ_2 , a_2 , a więc również q oraz a_2 , przyjmijmy $k = 1,1$. Mamy teraz $a_1 = a_2/k = 1,182$ oraz $a_1 = a_2/1,1 = 1,209$.

Z wykresu na rysunku 1 odczytujemy (dla $k = 1,1$, $q = 0,95$, $a_1 = 1,182$): $P_{12} = 0,735$.

Z (9) znajdujemy $P_1 = 2\Theta(1,182) = 0,763$ i z (11) obliczamy $P_2 = 0,965$.

Wreszcie przyjmijmy $a_1 = \infty$, czyli zrezygnujmy zupełnie z kontroli fabrycznej zachowując nadal niezmiennione σ_1 , σ_2 , a_2 . Mamy nadal $a_2 = 1,3$ i oczywiście $P_{12} = P_2 = 2\Theta(1,3) = 0,838$.

Wygodnie jest zebrać wyniki w następującej tabliczce:

a_1	P_{12}	P_2
∞	0,838	0,838
1,209	0,735	0,965
1,023	0,668	0,988

Jak widzimy, zmniejszenie α_1 , zmniejsza P_{12} a zwiększa P_2 . Znaczy to, że zaostrzanie kontroli fabrycznej zmniejsza przeciętną frakcję egzemplarzy przyjętych przez kontrolę zewnętrzną i zwiększa przeciętną frakcję tych egzemplarzy odniesioną do liczby egzemplarzy przedstawionych do kontroli zewnętrznej.

Znając koszty związane z niezakwalifikowaniem się egzemplarza w próbie fabrycznej i w próbie zewnętrznej, producent może z łatwością ustalić optymalną wartość α_1 .

Warto zwrócić uwagę, że w naszym przykładzie odchylenie średnie σ_2 błędu było małe wobec σ_1 , a wskutek tego współczynnik ρ korelacji był duży. Gdybyśmy zbadali przypadek małego ρ , okazałoby się że producent niewiele może zyskać: ze zmniejszaniem się α_1 , wielkość P_2 wzrasta tylko nieznacznie, a wielkość P_{12} maleje bardzo silnie. Jest to zrozumiałe, bo gdy σ_2 jest duże, kontrola fabryczna niezbyt skutecznie podnosi jakość produktu.

Wprost z tablicy 1 można odczytać, że dwustopniowa kontrola jest mało celowa również wtedy, gdy α_1 jest duże. Istotnie, w dolnym wierszu tej tablicy, dla $\alpha_1 = 2,5$, wartości P_{12} nie zależą praktycznie ani od ρ , ani od k ; zresztą są zawsze wysokie, a oczywiście zawsze zachodzi relacja $P_2 > P_{12}$.

Tak więc opisana metoda najbardziej nadaje się w tych przypadkach, gdy jakość produkcji jest niewysoka w porównaniu z wymaganiami (małe α_1) i gdy błędy pomiarowe są niewielkie (duże ρ).

Prace cytowane

- [1] Ö. Eltetö and K. Sarkadi, *On preliminary sampling with special reference to the microbiological control of tomato concentrates*, Biometrics Vol. 16, No. 3, str. 339-347.
- [2] J. Obalski, *O wyznaczaniu granic dopuszczalnych błędów przy wstępnym sprawdzaniu narzędzi mierniczych*, Pomiar-Automatyka-Kontrola, Nr. 8, sierpień 1960 r., str. 307-310.
- [3] J. Oderfeld i S. Zubrzycki, *O sprawdzaniu wodomierzy*, Zastosowania Matematyki 1(1954), str. 125-137.
- [4] D. B. Owen, *Tables for computing bivariate normal probabilities*, Ann. Math. Stat. 27 (1956), str. 1075.
- [5] D. B. Owen and J. M. Wiesen, *A method of computing bivariate normal probabilities*, The Bell System Technical Journal, March 1959, str. 553.
- [6] Е. С. Вентцель, *Теория вероятности*, Москва 1958.
- [7] E. Fidelis i R. Zieliński, *O pewnej metodzie obliczania prawdopodobieństw dwuwymiarowej zmiennej normalnej za pomocą siatki prawdopodobieństw* (praca w przygotowaniu do druku).

Praca wpłynęła 30. 12. 1960

Е. ФИДЕЛИС и ЯН ОДЕРФЕЛЬД (Варшава)

ДВУСТЕПЕННЫЙ КОНТРОЛЬ С УЧЁТОМ ОШИБОК ИЗМЕРЕНИЙ

РЕЗЮМЕ

Производитель лично производит статистическую проверку производимого продукта и представляет к окончательному статистическому отбору только те экземпляры, которые успешно прошли внутренний контроль. При каждой из проверок используются измерительные приборы, обладающие случайными ошибками.

В работе исследованы вероятностные последствия этой схемы и представлены результаты, пригодные к непосредственному использованию.

E. FIDELIS and J. ODERFELD (Warszawa)

TWO-STEP CONTROL TAKING INTO ACCOUNT THE MEASUREMENT ERRORS

SUMMARY

The producer himself controls his product statistically and presents for final statistical acceptance only those pieces which have passed the internal control successfully. In both controls a measuring device biased with a random error is employed.

The authors examine the probabilistic consequences of this procedure and present results in a form which can be directly applied.