J. ODERFELD (Warszawa)

STATYSTYCZNE BADANIE NA CECHY SKORELOWANE

§ 1. W praktyce statystycznej kontroli jakości dość często się zdarza, że badanie przedmiotów na ważną cechę Y jest takie kosztowne lub trudne, że z partii można zbadać tylko niewiele sztuk; natomiast badanie na mniej ważną cechę X może być tanie i łatwe. Oto kilka przykładów: X — twardość, a Y — wytrzymałość stali, X — twardość, a Y — ścieralność gumy, X — opór elektryczny, a Y — zawartość krzemu w stali. W tych wszystkich przypadkach badanie na cechę Y jest nie tylko kosztowne, ale niszczące.

Jeśli cechy X i Y są mocno skorelowane, nasuwa się myśl, żeby orzeczenia o partii dotyczące cechy Y oprzeć na obu rodzajach badania i przez to część trudnych badań zastąpić łatwiejszymi.

Taka myśl, w odniesieniu do estymacji, nie jest nowa. Np. w [1] można znaleźć przykłady i teorię losowania dwufazowego. Najpierw bada się dużą losową próbkę o licznosci N tylko na cechę X , a następnie małą losową próbkę o licznosci n na obie cechy i opiera się wnioskowanie o rozkładzie brzegowym cechy Y na obu zespołach informacji.

Dla przejrzystości w niniejszej pracy wprowadzono wzór (6) przy stosownych założeniach, nie korzystając z wzorów ogólnych podanych w [1]. Nowe wydaje się zastosowanie naszkicowanej zasady do postępowania odbiorczego. Paragraf 3 jest oparty na innej zasadzie niż wyłożona w [1].

Przypuśćmy, że z doświadczenia wiadomo, że dwuwymiarowa zmienna losowa X, Y ma stałe i znane odchylenia średnie σ_x i σ_y oraz współczynnik korelacji ρ i że średnie μ_x i μ_y są stałe i nieznanne.

Możliwe jest oszacowanie σ_x , σ_y i ρ z aktualnej próbki, prostsze jednak i lepsze jest oparcie się na uprzednich obszernych pomiarach, jeśli potwierdzają one dużą stateczność tych trzech liczb, charakteryzujących produkt wytwarzany w ustalonym procesie technologicznym. Sądzę, że uwagi te stosują się na przykład do twardości i wytrzymałości stali węglowej, nie dysponuję jednak materiałem liczbowym na poparcie tego przypuszczenia.

Pokażemy najpierw, jak można utworzyć przedział ufności dla μ_y , traktując partię produktu jako próbkę losową populacji o parametrach $\mu_x, \mu_y, \sigma_x, \sigma_y, \rho$. W tym celu sprawdzamy całą partię liczącą N (duża liczba) elementów na cechę X i otrzymujemy ciąg wartości $x_1, x_2, \dots, x_N$, po czym znajdujemy średnią

$$(1) \quad \bar{x}_N = \frac{1}{N} \sum_1^N x_i,$$

która ze względu na dużą licznosc próbki jest niemal dokładnym oszacowaniem μ_x .

Następnie pobieramy losowo próbkę o licznosci n (mała liczba) i znajdujemy średnie

$$(2) \quad \bar{x}_n = \frac{1}{n} \sum_1^n x_i, \quad \bar{y}_n = \frac{1}{n} \sum_1^n y_i.$$

Z punktu $(\bar{x}_n, \bar{y}_n)$ prowadzimy linię prostą równoległą do linii regresji drugiego rodzaju $\bar{y}$ na $\bar{x}$ i znajdujemy rzędną $\bar{y}_N$ jej przecięcia z linią o równaniu $x = \bar{x}_N$. Łatwo możemy sprawdzić, że

$$(3) \quad \bar{y}_N = \bar{y}_n - \rho \frac{\sigma_y}{\sigma_x} (\bar{x}_n - \bar{x}_N).$$

Kładąc z omówionym przybliżeniem $\bar{x}_N = \mu_x$ mamy

$$(4) \quad \bar{y}_N = \bar{y}_n - \rho \frac{\sigma_y}{\sigma_x} (\bar{x}_n - \mu_x).$$

Okażemy, że $\bar{y}_N$ określone przez (3) jest nieobciążonym estymatorem μ_y . Istotnie, $\bar{y}_N$ jest realizacją zmiennej losowej $\bar{Y}_N$ o kształcie

$$(5) \quad \bar{Y}_n = \frac{1}{n} \sum_1^n Y_i - \rho \frac{\sigma_y}{\sigma_x} \left(\frac{1}{n} \sum_1^n X_i - \frac{1}{N} \sum_1^N X_i \right).$$

Stąd

$$E(\bar{Y}_n) = \mu_y - \rho \frac{\sigma_y}{\sigma_x} (\mu_x - \mu_x) = \mu_y.$$

Teraz wyznaczmy wariancję zmiennej losowej $\bar{Y}_n$. Skorzystamy przy tym z przybliżenia (4), aby uniknąć komplikacji, do których prowadzi (3).

$$D^2(\bar{Y}_n) = \frac{1}{n^2} D^2 \left(\sum_1^n Y_i - \rho \frac{\sigma_y}{\sigma_x} \sum_1^n X_i \right) = \frac{1}{n^2} \left(n\sigma_y^2 + n\rho^2 \frac{\sigma_y^2}{\sigma_x^2} \sigma_x^2 - 2\rho n \rho \frac{\sigma_y}{\sigma_x} \sigma_x \sigma_y \right),$$

czyli

$$(6) \quad D^2(\bar{Y}_n) = \frac{1}{n} \sigma_y^2 (1 - \rho^2) \stackrel{\text{def}}{=} s^2.$$

A więc $\bar{Y}_n$ ma rozkład brzegowy ze średnią μ_y (dokładnie) i z odchyleniem średnim s (w przybliżeniu).

Teraz widać wyraźnie, dlaczego proponujemy szacować μ_y za pomocą estymatora $\bar{y}_N$, a nie za pomocą estymatora $\bar{y}_n$, który wprowadzie jest tak samo nieobciążony, ale ma wariancję σ_y^2/n , czyli większą niż (6). Jeśli ρ jest bliskie 1 lub -1 , redukcja wariancji jest znaczna. Poszukiwany przedział ufności dla μ_y jest

$$(7) \quad (\bar{y}_N - \alpha_1 s, \bar{y}_N + \alpha_2 s),$$

gdzie α_1, α_2 są współczynnikami zależnymi od rodzaju rozkładu i poziomu ufności.

§ 2. Wykorzystanie tych uwag do ułożenia przepisu odbiorczego jest proste. Wszystkie sztuki partii o liczności N sprawdzamy na cechę łatwą X , a wszystkie sztuki w próbie o liczności n sprawdzamy ponadto na cechę trudną Y . Przyjmujemy wartość krytyczną y_k . Stosując wzory (1), (2), (3) znajdujemy kolejno $\bar{x}_N, \bar{x}_n, \bar{y}_n, \bar{y}_N$.

Obszar krytyczny tworzymy zależnie od tego, czy jest pożądane, żeby Y przyjmowało wartości możliwie duże (przypadek a), czy też możliwie małe (przypadek b).

W przypadku a partię uznajemy za dobrą lub niedobłą zależnie od tego, czy $\bar{y}_N \geq y_k$, czy też $\bar{y}_N < y_k$.

W przypadku b partię uznajemy za dobrą lub niedobłą zależnie od tego, czy $\bar{y}_N \leq y_k$, czy też $\bar{y}_N > y_k$.

Wyznamy teraz moc testu przyjmując, że zmienna losowa X, Y ma dwuwymiarowy rozkład normalny.

Zacniemy od przypadku a. Oznaczmy przez y_0 najmniejszą wartość cechy Y , przy której można jeszcze uznać sztukę za dobrą, przez w — wadliwość partii ze względu na cechę Y , to znaczy frakcję sztuk, których cecha jest mniejsza od y_0 , przez Q — prawdopodobieństwo uznania partii za niedobłą. Ponadto wprowadźmy funkcję F^{-1} , odwrotną do dystrybuanty F rozkładu normalnego $N(0, 1)$, określoną następująco:

$$(8) \quad F^{-1}(F(A)) = A.$$

Przy tych oznaczeniach moc testu jest określona przez układ równań

$$(9) \quad \frac{\mu_y - y_0}{\sigma_y} = F^{-1}(w), \quad \frac{\mu_y - y_k}{s} = F^{-1}(Q).$$

Po wyrugowaniu z (9) parametru μ_y i uwzględnieniu (6) znajdujemy moc testu w postaci ostatecznej:

$$(10) \quad y_k - y_0 = \sigma_y \left[F^{-1}(w) - F^{-1}(Q) \frac{\sqrt{1 - \varrho^2}}{\sqrt{n}} \right].$$

Ponieważ w (10) są dwa parametry obieralne, y_k i n , można nałożyć dwa warunki, na przykład takie:

$$(11) \quad \text{jeśli } w = w_1, \text{ to } Q = Q_1, \quad \text{jeśli } w = w_2, \text{ to } Q = Q_2.$$

Z (10) i (11) otrzymujemy

$$(12) \quad n = (1 - \varrho^2) \left[\frac{F^{-1}(Q_1) - F^{-1}(Q_2)}{F^{-1}(w_1) - F^{-1}(w_2)} \right]^2,$$

$$(13) \quad y_k = y_0 + \sigma_y \frac{F^{-1}(w_2)F^{-1}(Q_1) - F^{-1}(w_1)F^{-1}(Q_2)}{F^{-1}(Q_1) - F^{-1}(Q_2)}.$$

W przypadku b y_0 oznacza największą wartość cechy Y , przy której można jeszcze uznać sztukę za dobrą, a wadliwość w — frakcję sztuk, których cecha jest większa od y_0 . Niechaj, jak poprzednio, Q będzie prawdopodobieństwem uznania partii za niedobłą. Niechaj wreszcie obowiązują nadal warunki (11). Przy tych umowach, wobec symetrii rozkładu normalnego, wzory (10), (12) i (13) pozostają w mocy.

Łatwo można ułożyć tablice lub nomogramy uwalniające praktyka od rachunków związanych ze stosowaniem wzorów (12) i (13), nie będziemy się jednak tym zajmowali.

Porównując ten test z klasycznym testem „jednostronnym”, który byłby oparty tylko na badaniu Y i na znajomości σ_y , widzimy, że w naszym teście, badając n sztuk na cechę Y , uzyskuje się tę samą moc co w teście klasycznym przy badaniu na Y próbki o liczności $n' = n/(1 - \varrho^2)$. Który z testów jest ekonomiczniejszy, zależy od kosztów badania.

Przyjmijmy na przykład, że koszt naszego testu jest $C + k_1N + k_2n$, a koszt testu klasycznego o tej samej mocy jest $C + k_2n'$, czyli $C + k_2n/(1 - \varrho^2)$, przy czym C , k_1 i k_2 są stałe. Nasz test opłaca się, gdy

$$(14) \quad \frac{k_2}{k_1} > \frac{N}{n} \cdot \frac{1 - \varrho^2}{\varrho^2}.$$

Dla ilustracji podamy przykład odnoszący się do przypadku b.

Niech przedmiotem kontroli będą opony samochodowe badane na cechę główną, ścieralność Y , i cechę zastępczą, twardość X . Przypuśćmy, że największa dopuszczalna ścieralność y_0 dobrej opony jest równa 200. Żądamy, żeby dla $w_1 = 10\%$ było $Q_1 = 95\%$, a dla $w_2 =$

$= 2,5\%$ było $Q_2 = 5\%$. Przypuśćmy, że z doświadczenia wiadomo, że zmienna losowa X , Y jest normalna, współczynnik ρ korelacji między X a Y wynosi $-0,89$, a odchylenia średnie wynoszą odpowiednio $\sigma_x = 8$, $\sigma_y = 30$.

Z tablic rozkładu normalnego odczytujemy

$$\begin{aligned} F^{-1}(w_1) &= -1,284, & F^{-1}(Q_1) &= 1,645, \\ F^{-1}(w_2) &= -1,960, & F^{-1}(Q_2) &= -1,645. \end{aligned}$$

Podstawiając te wartości do (12) i (13) znajdujemy

$$n = 5, \quad y_k = 151,3.$$

Według (10) moc testu ma postać

$$\frac{151,3 - 200}{30} = F^{-1}(w) - F^{-1}(Q) \sqrt{\frac{0,21}{5}},$$

czyli

$$F^{-1}(Q) = 4,88[F^{-1}(w) + 1,622],$$

co prowadzi do następującej tabliczki:

100 w	0	1	2	4	6	8	10
100 Q	0	0,03	1,7	27	63	86	95

Nawiasem mówiąc, nasz test jest pod względem mocy mniej więcej równoważny z pojedynczym testem alternatywnym o liczności próbki 100, w którym partię uznaje się za niedobłą, gdy w próbie pokaże się 6 lub więcej sztuk niedobrych.

Z (3) znajdujemy $\bar{y}_N$:

$$(15) \quad \bar{y}_N = \bar{y}_5 + 0,89 \frac{30}{8} (\bar{x}_5 - \bar{x}_N) = \bar{y}_5 + 3,34 (\bar{x}_5 - \bar{x}_N).$$

Przypuśćmy teraz, że stosując ten test w praktyce do paruset opon samochodowych znaleziono $\bar{x}_N = 71$, $\bar{x}_5 = 67,5$, $\bar{y}_5 = 160$. Z (15) znajdujemy

$$\bar{y}_N = 160 + 3,34(67,5 - 71) = 148.$$

Ponieważ $y_k = 151,3$, więc $\bar{y}_N < y_k$ i partię należy uznać za dobrą.

§ 3. Opiszemy teraz inny test oparty na własnościach statystyk pozycyjnych. Dla skrócenia wywodów będziemy uważali za dobrą sztukę taką, której cecha Y jest niemniejsza od pewnej wartości y_0 ; niech

ponadto cecha Y , trudna do badania, będzie skorelowana dodatnio z cechą X łatwą do badania. Inne przypadki będą wymagały tylko nieistotnych zmian.

Z partii przedmiotów pobieramy próbkę o dużej liczności N i sprawdzamy ją na cechę X łatwą do badania. Następnie porządkujemy próbkę niemalejąco i wybieramy tę sztukę, u której zaobserwowana wartość x jest m -tą od lewej strony, to znaczy „ m -tą najgorszą” sztukę ze względu na X . Na cechę Y badamy tylko tę właśnie sztukę. Niech zaobserwowana wartość Y będzie y . Jeśli $y \geq y_k$, gdzie y_k jest wartością krytyczną, całą partię uznajemy za dobrą; w przeciwnym przypadku za niedobłą.

Przy tym postępowaniu korzystamy więc z informacji o X , ale tylko do tendencyjnego wyboru jedynej sztuki sprawdzanej na Y . Jest jasne, że rezygnujemy przy tym z pewnej części informacji zawartej w liczbowych wartościach $x_1, x_2, \dots, x_N$ i że nie można wiele informacji wydobyc z jednego pomiaru cechy Y . Mimo to wydaje się, że test może być użyteczny dzięki swej prostocie, ekonomiczności i — jak zobaczymy — nienajgorszym własnościom statystycznym. Warto zwrócić uwagę, że czasem uporządkowanie n sztuk ze względu na cechę X może być szczególnie łatwe. Tak będzie wówczas, gdy uda się organoleptycznie wyodrębnić tę część sztuk, wśród których niewątpliwie znajdują się sztuki „najgorsze” ze względu na X . Przykład: niech cechą X będzie luz w łożyskach, cechą Y — trwałość, wymagająca paruset godzin próbowania. Otóż nieraz bez dokładnego mierzenia można z paruset łożysk wybrać kilkanaście szczególnie luźnych, sprawdzić ich luzy i próbie na trwałość poddać tylko „trzecie najgorsze” łożysko.

Rozpatrzmy teraz zmienną losową normalną dwuwymiarową $N(0, 0, 1, 1, \rho)$. Rozkład względny jest normalny $N(\rho x, \sqrt{1 - \rho^2})$; oznaczmy jego gęstość przez $f(y|x)$; rozkłady brzegowe są normalne $N(0, 1)$. Asymptotyczny rozkład m -tej od lewej strony statystyki uporządkowanej w próbce o liczności N ma (por. np. [2]) gęstość $\varphi_m(x_m)$ określoną następująco:

$$(16) \quad F(u_m) = \frac{m}{N}, \quad a_m = \frac{N}{m} f(u_m), \quad z_m = a_m(x_m - u_m),$$

$$\varphi_m(x_m) = \frac{m^m}{\Gamma(m)} a_m \exp(mz_m - me^{z_m});$$

F i f oznaczają odpowiednio dystrybuantę i gęstość rozkładu $N(0, 1)$. Tablicę $\varphi_m(x_m)$ dla $m = 1$ opublikowano w [2]; przez jej łatwą transformację można otrzymać gęstość dla innych wartości m .

Asymptotyczny rozkład brzegowy cechy Y_m elementu m -tego od lewej strony ze względu na cechę X ma gęstość $\psi_m(x_y)$ i dystrybuantę $\Psi_m(x_y)$; są one określone przez wzory

$$(17) \quad \psi_m(y) = \int_{-\infty}^{\infty} f(y|x_m) \varphi_m(x_m) dx_m, \quad \Psi_m(y) = \int_{-\infty}^y \psi_m(t) dt.$$

Badanie ogólne rozkładu określonego przez (17) wydaje się trudne i dla tego poprzestaliśmy na numerycznym zbadaniu przypadku szczególnego, gdy $N = 50$, $\rho = 0,9$. Wyniki⁽¹⁾ są zawarte w tablicy 1 i pokazane na rysunku 1.

TABLICA 1

Dystrybuanta zmiennej losowej
określonej przez (17)
dla $N = 50$, $\rho = 0,9$, $m = 1, 2, 3$.

y	Dystrybuanta		
	$100 \Psi_1(y)$	$100 \Psi_2(y)$	$100 \Psi_3(y)$
-5,0	0,0	0,0	0,0
-4,5	0,1	0,0	0,0
-4,0	0,6	0,0	0,0
-3,5	2,4	0,2	0,0
-3,0	8,3	1,5	0,4
-2,5	24,0	8,4	3,4
-2,0	51,2	29,2	17,3
-1,5	79,6	62,0	47,8
-1,0	95,4	88,4	80,0
-0,5	99,6	98,5	96,3
0	100,0	100,0	99,8
0,5	100,0	100,0	100,0
$E(Y_m)$	-2,05	-1,69	-1,48
$D^2(Y_m)$	0,423	0,307	0,267

Rys. 1

Zgodnie z przewidywaniem wariancja rozkładu określonego przez (17) jest mniejsza od wariancji w rozkładzie brzegowym wynoszącej 1. Ponadto wariancja maleje ze wzrostem m .

Teraz możemy wyznaczyć moc testu opisanego na początku tego paragrafu. Postulujemy, że ρ i σ_y są stałe i znane, μ_y jest stałe i nieznane, y_0 wynika z wymagań technicznych, y_k jest parametrem mocy. Zakładamy, że $N = 50$, $\rho = 0,9$.

⁽¹⁾ Obliczeniami kierował E. Fidelis, rachunki wykonała K. Olsztyńska z działu SKJP Inst. Mat. PAN.

Oznaczmy przez F^{-1} funkcję odwrotną do dystrybuanty F rozkładu $N(0, 1)$, przez Ψ_m^{-1} funkcję odwrotną do dystrybuanty określonej przez tablicę 1, przez w wadliwość partii, to znaczy frakcję sztuk, w których $Y < y_0$, oraz przez Q prawdopodobieństwo uznania partii za nie-dobrą. Przy tych oznaczeniach

$$(18) \quad \frac{y_0 - \mu_y}{\sigma_y} = F^{-1}(w),$$

$$(19) \quad \frac{y_k - \mu_y}{\sigma_y} = \Psi_m^{-1}(Q).$$

Wprowadźmy jeszcze oznaczenie

$$(20) \quad k = \frac{y_0 - y_k}{\sigma_y}.$$

Wtedy, z (18), (19) i (20),

$$(21) \quad F^{-1}(w) - \Psi_m^{-1}(Q) = k.$$

Jest to moc testu. Ponieważ ma tylko jeden parametr k , więc można na nią nałożyć jeden warunek, na przykład żeby dla $Q = Q_2$ było $w = w_2$.

Użyteczne będzie porównać moc naszego testu z mocą testu klasycznego, w którym nie byłoby wstępnego badania na cechę X , a orzeczenie o partii opierałoby się na zmierzeniu cechy Y u jednej, losowo wykazanej sztuki; gdyby wynik tego pomiaru był y' , partię uznanoby za dobrą pod warunkiem, że $y' \geq y_k$. Moc testu klasycznego miałaby postać

$$(22) \quad F^{-1}(w) - F^{-1}(Q) = k.$$

Na rysunku 2 pokazano krzywe operacyjno-charakterystyczne testu klasycznego i testów zaproponowanych dla $m = 1, 2$ i 3 . Wszystkie krzywe odpowiadają tym samym warunkom: $w_2 = 2\%$, $Q_2 = 5\%$. Jak widzimy, nowy test jest selektywniejszy od klasycznego; jego selektywność rośnie z m .

§ 4. Test opisany w paragrafie poprzednim jest tylko przykładem wielu wariantów, których w tej pracy nie badamy szczegółowo. Poprzestaniemy na wymienieniu paru możliwości.

1° Test z § 3 dla różnych wartości ρ i m .

2° Test oparty na średniej — powiedzmy — z pierwszej, drugiej i trzeciej skrajnej wartości cechy trudnej.

Rys. 2

3° Test oparty na informacjach w rodzaju 1° lub 2° uzupełnionych danymi o średniej $\bar{x}_N$ z dużej liczby pomiarów cechy łatwej.

4° Postępowanie mieszane, to znaczy sortująco-kwalifikujące. Można by na przykład najpierw zbadać całą partię na cechę łatwą X i usunąć z niej te przedmioty, w których zaobserwowano wartości iksów pozwalające przypuszczać, że wartości igreków są szczególnie niezadowalające: następnie testowanoby tylko partię uszlachetnioną. Ten projekt podała E. Pleszczyńska.

Prace cytowane

[1] W. G. Cochran, *Sampling techniques*, New York 1953, str. 278.

[2] E. J. Gumbel, *Statistical theory of extreme values*, National Bureau of Standards, series 33 (1954).

INSTYTUT MATEMATYCZNY POLSKIEJ AKADEMII NAUK

Praca wpłynęła 27. 12. 1957

Я. ОДЕРФЕЛЬД (Варшава)

СТАТИСТИЧЕСКОЕ ИССЛЕДОВАНИЕ КОРРЕЛИРОВАННЫХ ХАРАКТЕРИСТИК

РЕЗЮМЕ

Пусть X, Y — случайная величина распределение которой имеет известное среднее отклонение, известный коэффициент корреляции и неизвестные средние μ_x и μ_y .

Пусть X описывает характеристику продукта легко доступную для исследования, но мало важную, а Y — трудно доступную для исследования характеристику, но решающую вопрос пригодности продукта. Характеристику X можно проверять, следовательно, при большой выборке и даже в целой партии получая почти точную оценку μ_x . Характеристику же Y можно проверять только при малой выборке.

Сопоставляя соответственно эти информации и предполагая, что распределение X, Y двумерно и нормально, в работе приведен критерий приема служащий для проверки статистических гипотез, касающихся браков в целой партии. Приведена функция мощности критерия и показано, что предложенный критерий является в некоторых условиях более экономный чем классический.

Затем построен другой критерий, в котором исследование на X служит только для упорядочения выборки по отношению к характеристике X , после чего исследуется на Y только один предмет имеющий m -тое крайнее значение икса в выборке ($m = 1, 2, 3$).

И наконец, приведен набросок нескольких возможностей обобщения второго критерия и соединения его с первым.

J. ODERFELD (Warszawa)

STATISTICAL TESTING FOR CORRELATED CHARACTERISTICS

SUMMARY

Let the random variable X, Y have a distribution with known mean deviations, a known correlation coefficient and unknown means μ_x and μ_y .

Let X denote a characteristic of the product that is easy to test but not very important, and Y a characteristic that is difficult to test but decisive as regards the usefulness of the product. We can thus test characteristic X in a large sample, or even in the whole lot, obtaining an almost exact estimate of μ_x . Characteristic Y , however, can be tested only in a small sample.

Combining these data in a suitable way and assuming that the distribution of X, Y is a two-dimensional normal one, the author describes an acceptance test serving to verify statistical hypotheses concerning the fraction defective of the lot. The functions of power are given, and it is shown that the test suggested is in certain conditions more economical than the classical one.

The author then constructs another test, in which testing for X serves only for ordering the sample with respect to characteristic X and only one item, having the m -th extreme value of X in the sample ($m = 1, 2, 3$) is tested for Y .