ZASTOSOWANIA MATEMATYKI
V (1960)

E. PLESZCZYŃSKA (Warszawa)

ODSIEWANIE W BADANIU STATYSTYCZNYM NA CECHY SKORELOWANE

1. W pracy opisano postępowanie, które jest uogólnieniem testu przedstawionego w [1], w § 2. Przyjęto te same założenia i utrzymano oznaczenia cytowanej pracy: dana jest zmienna losowa normalna (X, Y); ϱ, σ_x i σ_y są znane, μ_x i μ_y stałe, lecz nieznane. Partie pobrane z populacji (X, Y) poddajemy badaniu statystycznemu na cechę Y, która nas wyłącznie interesuje, a której wartości w przeciwieństwie do wartości cechy X są trudne do wyznaczenia. Cała partia zostaje sprawdzona na cechę łatwą X, co pozwala poznać z dużą dokładnością wartość przeciętną cechy X. Autor pracy [1] wykorzystuje tę informację do wyznaczenia estymatora wartości przeciętnej cechy Y. W niniejszej pracy znajomość rozkładu cechy X w partii zostaje wykorzystana także do odsiewania partii, tj. usunięcia z niej tych sztuk, w których zaobserwowana wartość x-ów każe się spodziewać niezadowalającej wartości y-ów.

Wszystkie rozważania stosują się do przypadku III spośród czterech dopuszczonych, ponieważ do niego odnosi się przykład rozpatrywany w [1]:

I. $\varrho > 0$, sztukę uważamy za dobrą, jeśli $Y < y_0$,

II. $\varrho > 0$, sztukę uważamy za dobrą, jeśli $Y > y_0$,

III. $\varrho < 0$, sztukę uważamy za dobrą, jeśli $Y < y_0$,

IV. $\varrho < 0$, sztukę uważamy za dobrą, jeśli $Y > y_0$;

w pozostałych przypadkach istnieje pełna analogia.

Dla rozpatrywanego przypadku odsiewanie polega więc na odrzuceniu sztuk, dla których $X < \xi$, gdzie ξ jest parametrem testu.

Przede wszystkim pokażę, że takie odsiewanie powoduje zmniejszenie wadliwości.

Jeśli w populacji generalnej o wadliwości $w = P(Y > y_0)$ odrzuca się elementy, dla których $X < \xi$, powstaje populacja ucięta o wadliwości $w' = P(Y > y_0, X > \xi)/P(X > \xi)$.

Obliczam różnicę

$$w-w' = P(Y > y_0) - \frac{P(Y > y_0, X > \xi)}{P(X > \xi)} =$$

$$= \frac{P(Y > y_0)P(X > \xi) - P(Y > y_0, X > \xi)}{P(X > \xi)} \left\{ \begin{array}{ll} < 0 & \text{dla } \varrho > 0, \\ = 0 & \text{dla } \varrho = 0, \\ > 0 & \text{dla } \varrho < 0. \end{array} \right.$$

W rozpatrywanym przypadku III $\varrho < 0$, więc $w > w'$.

W rzeczywistości ucięcie następuje w partiach o liczności N pobranych z populacji generalnej. Wadliwość takich odsianych partii jest więc ilorazem zmiennych losowych L/M, gdzie L jest liczbą sztuk, dla których $Y > y_0$, $X > \xi$, M zaś liczbą sztuk, dla których $X > \xi$. Zmienna losowa L ma rozkład dwumianowy o prawdopodobieństwie sukcesu

$$p_L = P(Y > y_0, X > \xi).$$

Zmienna losowa M ma rozkład dwumianowy o prawdopodobieństwie sukcesu

$$p_M = P(X > \xi).$$

Łączny rozkład zmiennych losowych L i M ma następującą postać:

$$(1) \qquad P(L = j, M = k) = P(L = j \mid M = k) P(M = k) =$$

$$= \binom{k}{j} \left(\frac{p_L}{p_M}\right)^j \left(1 - \frac{p_L}{p_M}\right)^{k-j} \binom{N}{k} p_M^k (1-p_M)^{N-k}.$$

Z (1) wynika, że $P(L = 0, M = 0) = (1-p_M)^N > 0$, trzeba więc zdefiniować wartość L/M dla $L = 0$ i $M = 0$. Najprościej wydaje się przyjąć $w' = p_L/p_M$. Wtedy wartość przeciętna i wariancja zmiennej losowej L/M wyrażą się wzorami:

$$(2) \qquad E\left(\frac{L}{M}\right) = \frac{p_L}{p_M}(1-p_M)^N + \sum_{k=1}^{N} \sum_{j=1}^{k} \frac{j}{k} P(L = j, M = k) =$$

$$= \frac{p_L}{p_M}(1-p_M)^N + \sum_{k=1}^{N} \frac{1}{k} \left[\sum_{j=1}^{k} j \binom{k}{j} \left(\frac{p_L}{p_M}\right)^j \left(1 - \frac{p_L}{p_M}\right)^{k-j} \right] \times$$

$$\times \binom{N}{k} p_M^k (1-p_M)^{N-k} =$$

$$= \frac{p_L}{p_M}(1-p_M)^N + \sum_{k=1}^{N} \frac{1}{k} k \frac{p_L}{p_M} \binom{N}{k} p_M^k (1-p_M)^{N-k} =$$

$$= \frac{p_L}{p_M} \left[\sum_{k=0}^{N} \binom{N}{k} p_M^k (1-p_M)^{N-k} \right] = \frac{p_L}{p_M} = w'.$$

$$D^2\left(\frac{L}{M}\right) = \frac{p_L^2}{p_M^2}(1-p_M)^N + \sum_{k=1}^{N}\sum_{j=1}^{k}\frac{j^2}{k^2}P(L=j,\ M=k) - \frac{p_L^2}{p_M^2} = \tag{3}$$

$$= \frac{p_L^2}{p_M^2}[(1-p_M)^N-1] + \sum_{k=1}^{N}\frac{1}{k^2}\left[\sum_{j=1}^{k} j^2\binom{k}{j}\left(\frac{p_L}{p_M}\right)^j\left(1-\frac{p_L}{p_M}\right)^{k-j}\right]\times$$

$$\times\binom{N}{k}p_M^k(1-p_M)^{N-k} =$$

$$= \frac{p_L^2}{p_M^2}[(1-p_M)^N-1] + \sum_{k=1}^{N}\frac{1}{k^2}\left(k\frac{p_L}{p_M} + k(k-1)\frac{p_L^2}{p_M^2}\right)\times$$

$$\times\binom{N}{k}p_M^k(1-p_M)^{N-k} =$$

$$= \frac{p_L^2}{p_M^2}[(1-p_M)^N-1] + \frac{p_L^2}{p_M^2}[1-(1-p_M)^N] + \sum_{k=1}^{N}\left(\frac{p_L}{p_M} - \frac{p_L^2}{p_M^2}\right)\frac{1}{k}\times$$

$$\times\binom{N}{k}p_M^k(1-p_M)^{N-k} =$$

$$= \frac{p_L}{p_M}\left(1-\frac{p_L}{p_M}\right)\sum_{k=1}^{N}\frac{1}{k}\cdot\frac{N!}{k!(N-k)!}p_M^k(1-p_M)^{N-k} \leqslant$$

$$\leqslant \frac{p_L}{p_M}\left(1-\frac{p_L}{p_M}\right)\sum_{k=1}^{N}\frac{2}{k+1}\cdot\frac{N!}{k!(N-k)!}p_M^k(1-p_M)^{N-k} =$$

$$= \frac{p_L}{p_M}\left(1-\frac{p_L}{p_M}\right)\frac{2}{(N+1)p_M}\sum_{k+1=2}^{N+1}\frac{(N+1)!}{(k+1)!(N-k)!}p_M^{k+1}(1-p_M)^{N-k} <$$

$$< \frac{p_L}{p_M}\left(1-\frac{p_L}{p_M}\right)\frac{2}{(N+1)p_M}\xrightarrow[N\to\infty]{} 0.$$

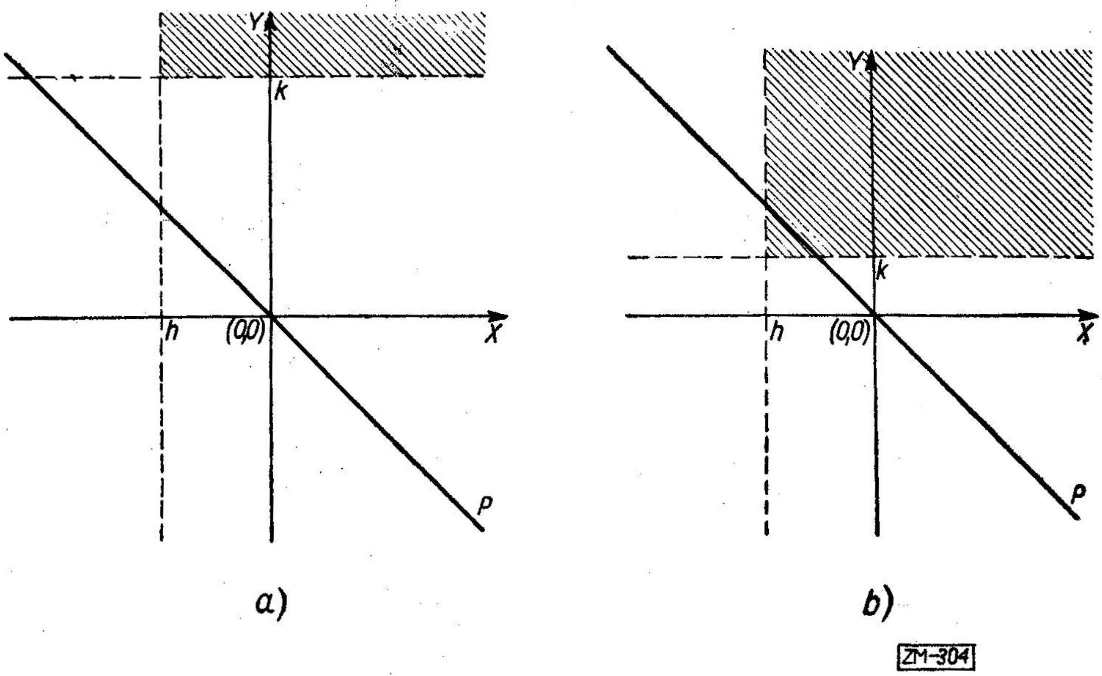
Zmienna losowa L/M jest zatem nieobciążonym estymatorem w' z wariancją bliską zeru dla dużych partii, a więc zaproponowane odsiewanie powoduje średnio zmniejszenie wadliwości partii.

Zależność między w a w' można wyrazić dokładnymi wzorami. Z definicji $w' = P(Y > y_0, X > \xi)/P(X > \xi)$ zależy od $h = (\xi-\mu_x)/\sigma_x$, $k = (y_0-\mu_y)/\sigma_y$ oraz ϱ. Ponieważ $w = P(Y > y_0)$, więc $1-w = G(k)$, gdzie G jest dystrybuantą standaryzowanej zmiennej normalnej.

Dla $\varrho = 0$ $\quad w' = w.$

$$\text{Dla } \varrho = -1 \quad w' = \begin{cases} 0, & \text{gdy } w \leqslant G(h), \\ \dfrac{w-G(h)}{1-G(h)}, & \text{gdy } w > G(h), \end{cases}$$

co wynika z rysunku 1: przy $\varrho = -1$ łączny rozkład (X, Y) skupia się na prostej P. W przypadku a) mamy $w < G(h)$, prosta P nie przecina obszaru zakreskowanego i całka po tym obszarze równa się zero, więc $w' = 0$. W przypadku b) mamy $w > G(h)$, całka po obszarze zakreskowanym równa się $w-G(h)$, więc $w' = \big(w-G(h)\big)/\big(1-G(h)\big)$.



Rys. 1. Zbiór wartości zmiennej losowej (X, Y): a) przy mniejszej wadliwości; b) przy większej wadliwości

Dla $-1 < \varrho < 0$ można obliczyć w' za pomocą tablic Owena [2]; według przyjętych tam oznaczeń

$$w' = \frac{1-G(h)-G(h)+B(h, k, \varrho)}{1-G(h)}.$$

Wykres na rysunku 2 przedstawia dwuparametrową rodzinę krzywych $w' = w'(w; \varrho, h)$. Wykres ten pozwala odczytać średnią wadliwość po odsianiu partii o wadliwości w, przedstawia więc korzyści takiego odsiewania. Natomiast wykres na rysunku 3 przedstawia ujemną stronę odsiewania, która polega na tym, że także i dobre sztuki zostają odrzucone. Na osi rzędnych odczytuje się prawdopodobieństwo odrzucenia dobrej sztuki przy odsiewaniu partii o wadliwości w (inaczej średni procent sztuk dobrych odrzuconych przy odsiewaniu). Oznaczono je przez m (od słowa „marnotrawstwo''):

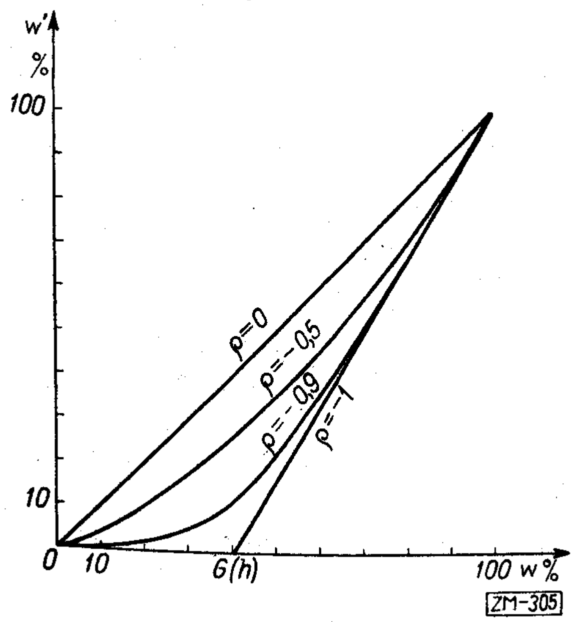
$$m = P(X < \xi,\ Y < y_0) = m(w;\, h,\, \varrho);$$

$$\text{dla } \varrho = 0 \qquad m = G(h)[1-w],$$

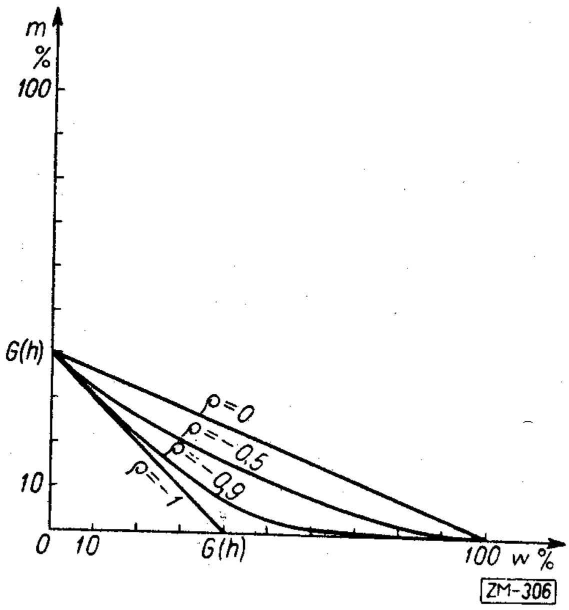
$$\text{dla } \varrho = -1 \qquad m = \begin{cases} G(h)-w, & \text{gdy } w \leqslant G(h), \\ 0, & \text{gdy } w > G(h), \end{cases}$$

dla $-1 < \varrho < 0$ według oznaczeń w tablicach Owena $m = B(h,\, k,\, \varrho)$.

Dla większej przejrzystości rysunki 2 i 3 zostały zrobione przy przesadnie dużym $G(h)$ (równym $40^0/_0$).



Rys. 2. Dwuparametrowa rodzina krzywych $w' = w'(w;\, h,\, \varrho)$



Rys. 3. Dwuparametrowa rodzina krzywych $m = m(w;\, h,\, \varrho)$

Za pomocą obu wykresów (rys. 2 i 3) można zawsze opisać konsekwencje stosowania każdego testu odbiorczego z odsiewaniem. Takich testów, dostosowanych do różnych sytuacji, może być bardzo wiele. Gdy na przykład produkcja jest tania, można zwracać mniejszą uwagę na wykres m; z drugiej strony, w rozmaitym stopniu może nam zależeć na uszlachetnieniu partii.

2. Przedstawię teraz dwa różne warianty postępowania odbiorczego z odsiewaniem.

Jądrem obu wariantów jest test z § 2 pracy [1], który będę odtąd nazywała *testem pierwotnym*. Przypomnę go w skrócie: całą partię o liczności N sprawdzamy na cechę łatwą X, obliczamy wartość przeciętną $\bar{x}_N$, małą próbkę o liczności n sprawdzamy na cechę trudną Y, obliczamy $\bar{x}_n$ i $\bar{y}_n$, a następnie

$$\bar{Y}_N^{\bullet} = \bar{y}_n - \varrho \frac{\sigma_y}{\sigma_x} (\bar{x}_n - \bar{x}_N).$$

Partię uznajemy za dobrą, jeśli $\bar{Y}_N < y_k$, za niedobrą, jeśli $\bar{Y}_N > y_k$.

Funkcja mocy tego testu ma postać

$$y_0 - y_k = \sigma_y \left[T(w) - T(Q) \frac{\sqrt{1-\varrho^2}}{\sqrt{n}} \right], \tag{4}$$

gdzie w jest średnią wadliwością partii, Q prawdopodobieństwem odrzucenia partii, $T(p)$ funkcją odwrotną dystrybuanty standaryzowanej zmiennej normalnej ze znakiem minus. Po nałożeniu na funkcję mocy warunków

$$\text{jeśli } w = w_i, \text{ to } Q = Q_i \ (i = 1, 2),$$

obliczamy parametry testu pierwotnego według wzorów

$$n = (1 - \varrho^2) \left[\frac{T(Q_1) - T(Q_2)}{T(w_1) - T(w_2)} \right]^2, \tag{5}$$

$$y_k = y_0 - \sigma_y \frac{T(w_2) T(Q_1) - T(w_1) T(Q_2)}{T(Q_1) - T(Q_2)}. \tag{6}$$

W moim pierwszym wariancie z góry wiadomo o partiach przedstawianych do odbioru, że ich wadliwość zwykle jest bardzo wysoka — na przykład wynosi około $4^0/_0$, podczas gdy nie budzi sprzeciwów wadliwość około $2^0/_0$, a nie ma możliwości uregulowania produkcji.

Stosujemy test, w którym z każdej partii odrzucamy przy odsiewaniu średnio $3^0/_0$ elementów. Sam test jest niemal identyczny z pierwotnym: obliczamy $\bar{x}_N$ i $\bar{Y}_N$ i jeśli $\bar{Y}_N < y_k$, odrzucamy z partii sztuki o $X < \xi$, a resztę partii uznajemy za dobrą; w przeciwnym razie partię odrzucamy. Parametrami testu są n, y_k i ξ. Pierwsze dwa parametry uzależniamy od mocy testu, a parametr ξ obliczamy w następujący sposób.

Wiadomo, że $G(h) = 3^0/_0$, gdzie $h = (\xi - \mu_x)/\sigma_x$; σ_x jest znane, za μ_x przyjmujemy $\bar{x}_N$. Stąd

$$\xi = h\sigma_x + \bar{x}_N = -1{,}88\sigma_x + \bar{x}_N.$$

Funkcja mocy tego testu będzie miała taką samą postać jak w teście pierwotnym:

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

w inny sposób trzeba jednak nałożyć na nią warunki. Muszą one bowiem dotyczyć wadliwości partii po odsianiu:

$$\text{dla } w'_i = w_i \quad Q = Q_i \ (i = 1,2).$$

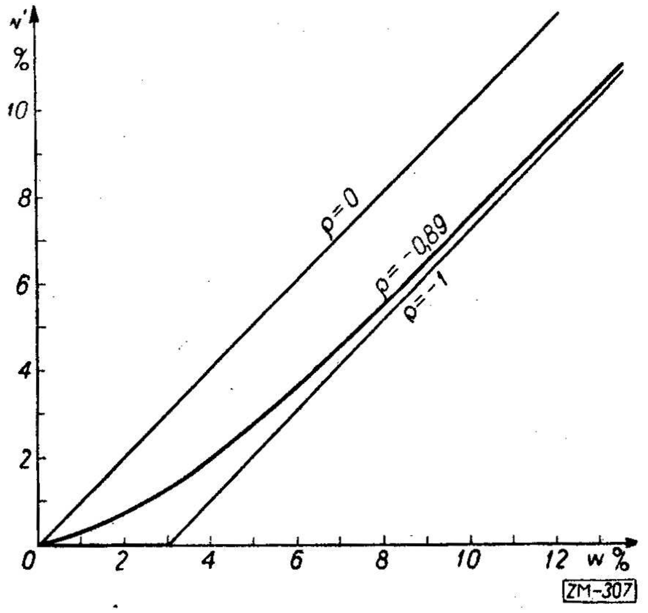
Z krzywej $w' = w'(w; h, \varrho)$ można odczytać, jakim to odpowiada wadliwościom partii przed odsiewaniem, a wtedy warunki będą już miały postać taką, jak w teście pierwotnym, co pozwoli obliczyć n i y_k. Z krzywej $m = m(w; h, \varrho)$ można odczytać, jaki procent dobrych sztuk jest średnio odrzucany przy najczęstszej wadliwości partii 4%.

PRZYKŁAD 1. Wszystkie dane liczbowe przyjmuję takie same, jak w przykładzie na test pierwotny:

$$\varrho = -0{,}89, \quad \sigma_x = 8,$$
$$\sigma_y = 30; \quad y_0 = 200.$$

Żądamy, żeby

$$(7) \quad \begin{array}{ll} \text{dla } w' = 1\%, & Q = 5\%, \\ \text{dla } w' = 7{,}5\%, & Q = 95\%. \end{array}$$



Rys. 4. Średnia wadliwość partii po sortowaniu trzyprocentowym

Rysunek 4 przedstawia krzywą $w'(w)$ dla $G(h) = 3\%$ i $\varrho = -0{,}89$. Za jej pomocą zastępujemy warunki (6) przez warunki równoważne

$$(7') \quad \begin{array}{ll} w = 2{,}65\%, & Q = 5\%, \\ w = 10{,}2\%, & Q = 95\% \end{array}$$

i obliczamy na podstawie (5) i (6) $n \approx 5$, $y_k = 152$.

Jeśli w praktyce otrzymano $\bar{x}_N = 71$, $\bar{Y}_N = 148$, to partię uznajemy za dobrą po odrzuceniu sztuk, dla których $X < \xi$, gdzie $\xi = -1{,}88 \cdot 8 + +71 = 56$.

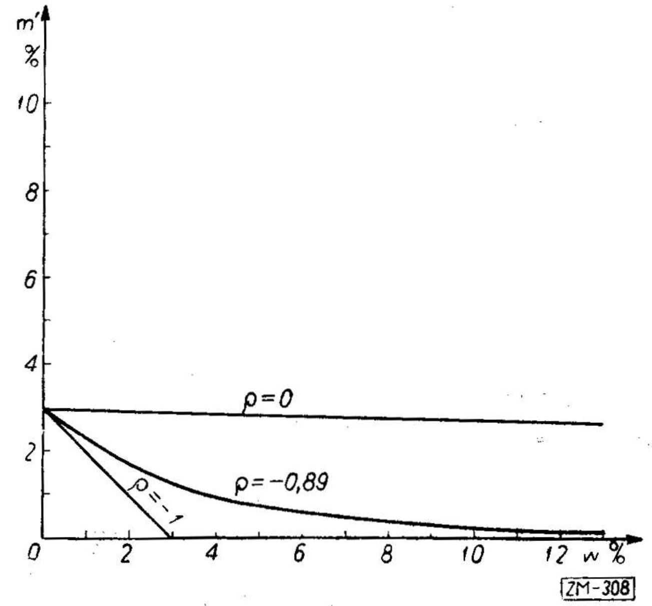
Moc testu przedstawia tablica 1.

TABLICA 1

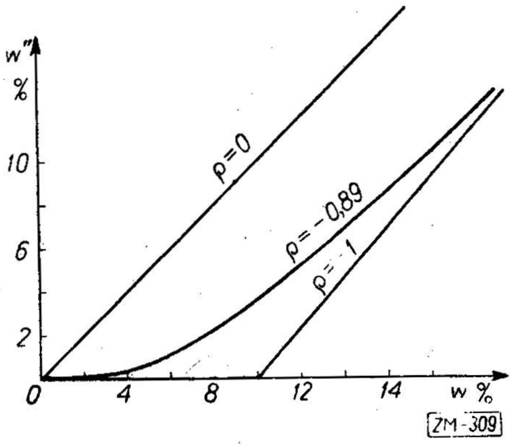
Średnia wadliwość partii w (w %)	0	2,7	4,1	5,3	6,5	8,7	9,6
Średnia wadliwość partii po odsiewaniu w' (w %)	0	1	2	3	4	6	7,5
Prawdopodobieństwo odrzucenia partii Q (w %)	0	5	25	48	66	88	95

Rysunek 5 przedstawia krzywą $m'(w)$ dla $G(h) = 3\%$ i $\varrho = -0{,}89$. Przy najczęstszej wadliwości partii przed odsiewaniem, wynoszącej 4%, odrzuca się średnio $0{,}9\%$ elementów.

Warunki (7) zostały tak dobrane, żeby moc testu była niemal dokładnie taka, jak w przykładzie na test pierwotny i żeby nie uległa zmianie wielkość próbki. Oznacza to, że „dostawca" w obu testach ma identyczną sytuację i że koszt badania próbki na cechę trudną jest ten sam.



Rys. 5. Średni procent sztuk dobrych odrzuconych przy sortowaniu trzyprocentowym



Rys. 6. Średnia wadliwość partii po sortowaniu dziesięcioprocentowym

Po odsiewaniu średnia wadliwość się zmniejsza — „odbiorca" jest tu bardziej uprzywilejowany niż w teście pierwotnym. Dzieje się to za cenę dodatkowej pracy przy odsiewaniu oraz za cenę odrzucenia pewnej liczby sztuk dobrych.

W drugim wariancie nie wiemy nic a priori o wadliwości partii i stosujemy postępowanie kilkustopniowe: partia nieprzyjęta w pewnym stopniu może zostać uszlachetniona przez odsiewanie (przy którym oprócz sztuk niedobrych dyskwalifikuje się na ogół pewną liczbę sztuk dobrych) i przyjęta w stopniu następnym. Jako przykład podaję następujący test o 4 stopniach (decyzjach).

Postępując jak w teście pierwotnym, obliczamy $\bar{x}_N$ i $\bar{Y}_N$.

Jeśli $\bar{Y}_N < y_k$, partię przyjmujemy;
jeśli $y_k < \bar{Y}_N < y'_k$, partię przyjmujemy po odrzuceniu sztuk, dla których $X < \xi_1$;
jeśli $y'_k < \bar{Y}_N < y''_k$, partię przyjmujemy po odrzuceniu sztuk, dla których $X < \xi_2$;
jeśli $\bar{Y}_N > y''_k$, partię odrzucamy.

Do konstrukcji takiego testu potrzebne są wykresy przedstawione na rysunkach 4, 5, 6 i 7. Parametrami testu są n, y_k, y_k', y_k'', ξ_1 i ξ_2. Musi zachodzić $y_k < y_k' < y_k''$, $\xi_1 < \xi_2$.

Przyjmijmy, że jedno odsiewanie jest trzyprocentowe, drugie dziesięcioprocentowe. Wtedy

$$\xi_1 = h_1\sigma_x + \bar{x}_N, \quad \text{gdzie} \quad G(h_1) = 3\%, \quad \text{czyli} \quad h_1 = -1{,}88,$$

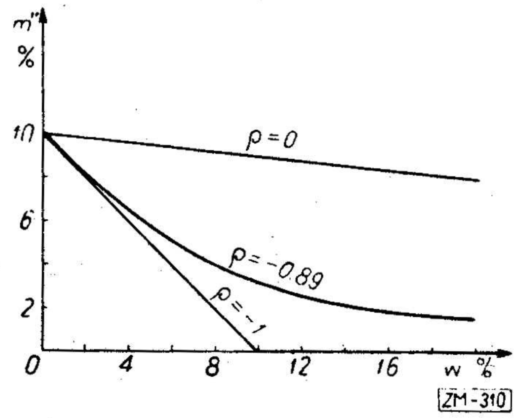
$$\xi_2 = h_2\sigma_x + \bar{x}_N, \quad \text{gdzie} \quad G(h_2) = 10\%, \quad \text{czyli} \quad h_2 = -1{,}28.$$

Funkcja mocy testu ma postać

$$y_0 - y_k'' = \sigma_y\left[T(w) - T(Q'')\frac{\sqrt{1-\varrho^2}}{\sqrt{n}}\right].$$

Jest to po prostu funkcja mocy w teście pierwotnym z parametrami y_k'' i n; Q'' oznacza prawdopodobieństwo odrzucenia partii. Nakładamy na nią dwa łagodne warunki, zgodnie z postulatem, że chcemy partie często przyjmować i obliczamy n i y_k'' jak w teście pierwotnym. Następnie obliczamy y_k'. Jeśli przez Q' oznaczyć prawdopodobieństwo odrzucenia partii lub odsiewania 10-procentowego, to y_k' jest jedynym parametrem krzywej

$$y_0 - y_k' = \sigma_y\left[T(w) - T(Q')\frac{\sqrt{1-\varrho^2}}{\sqrt{n}}\right]$$



Rys. 7. Średni procent sztuk dobrych odrzuconych przy sortowaniu dziesięcioprocentowym

(n już ustalono poprzednio). Nakładamy więc jeden warunek na tę krzywą, ostrzejszy niż poprzednio, i obliczamy y_k'. W ten sam sposób możemy znaleźć y_k: przez Q oznaczamy prawdopodobieństwo odrzucenia lub jakiegokolwiek odsiewania partii i nakładamy warunek ostrzejszy niż poprzednio na krzywą

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

Przebieg trzech krzywych Q'', Q' i Q w zależności od w zawiera wszystkie informacje o teście, interesujące dostawcę (partia przyjęta po odsiewaniu może być na przykład niżej oceniona). Aby pokazać odbiorcy, czego się może po tym teście spodziewać, zbadamy AOQ (average outgoing quality):

$$AOQ = (1-Q)w + (Q'-Q)w' + (Q''-Q')w'' + Q''\cdot 0,$$

gdzie w' jest wadliwością po odsiewaniu 3-procentowym partii o wadliwości w, w'' jest wadliwością po odsiewaniu 10-procentowym partii o wadliwości w.

Do pełnej charakterystyki testu trzeba jeszcze obliczyć średni procent odrzuconych sztuk dobrych przy przyjęciu partii o wadliwości w:

$$m = \frac{(1-Q)\cdot 0 + (Q-Q')m' + (Q'-Q'')m''}{1-Q''},$$

gdzie m' jest średnim procentem sztuk dobrych odrzuconych przy odsiewaniu 3-procentowym, m'' jest średnim procentem sztuk dobrych odrzuconych przy odsiewaniu 10-procentowym.

Przykład 2. W dalszym ciągu używam tych samych danych, które przyjęto w przykładzie na test pierwotny:

$$\varrho = -0{,}89, \quad \sigma_x = 8, \quad \sigma_y = 30, \quad y_0 = 200.$$

Nakładam warunki na Q'':

$$\text{dla } w_2 = 2\% \quad Q_2'' = 0{,}2\%, \qquad \text{dla } w_1 = 10\% \quad Q_1'' = 80\%.$$

Stąd

$$n = (1-\varrho^2)\left[\frac{T(Q_1'')-T(Q_2'')}{T(w_1)-T(w_2)}\right]^2 = 4{,}85 \approx 5,$$

$$y_k'' = y_0 - \sigma_y \frac{T(w_2)T(Q_1'')-T(w_1)T(Q_2'')}{T(Q_1'')-T(Q_2'')} = 156{,}3.$$

Jako następny warunek (na Q') przyjmuję

$$w_1 = 10\%, \quad Q_1' = 93\%.$$

Stąd

$$y_k' = y_0 - \sigma_y\left[T(w_1) - T(Q_1')\frac{\sqrt{1-\varrho^2}}{\sqrt{n}}\right] = 152{,}6.$$

Z tego obliczam, że gdy $w_2 = 2\%$, to

$$T(Q_k') = \sqrt{\frac{n}{1-\varrho^2}}\left[T(w_2) - \frac{y_0 - y_k'}{\sigma_y}\right] = 2{,}32, \qquad Q_2' = 1\%.$$

Wreszcie warunek na Q,

$$w_1 = 10\%, \quad Q = 97\%,$$

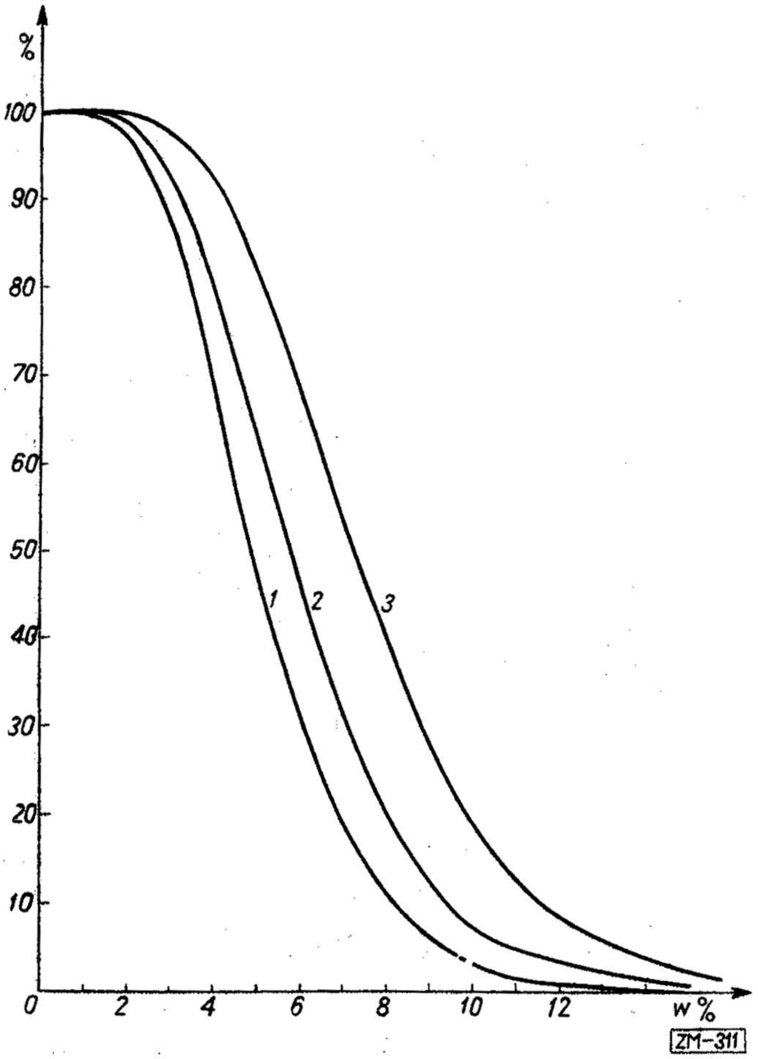
pozwala znaleźć

$$y_k = 150{,}2.$$

Z tego obliczam, że gdy $w_2 = 2\%$, to

$$T(Q_2) = 1{,}93, \quad Q_2 = 2{,}72.$$

A więc partię przyjmujemy bez odsiewania, gdy $\overline{Y}_N < 150{,}2$, przyjmujemy po 3-procentowym odsiewaniu, gdy $150{,}2 < \overline{Y}_N < 152{,}6$, po 10-procentowym odsiewaniu, gdy $152{,}6 < \overline{Y}_N < 156{,}3$, i odrzucamy, gdy $\overline{Y}_N > 156{,}3$.



Rys. 8. Krzywa 1 — prawdopodobieństwo przyjęcia partii boz sortowania; krzywa 2 — prawdopodobieństwo przyjęcia partii bez sortowania lub z sortowaniem trzyprocentowym; krzywa 3 — prawdopodobieństwo przyjęcia partii

Konsekwencje stosowania tego testu są zestawione w tablicy 2, prawdopodobieństwa poszczególnych decyzji testu przy średniej wadliwości partii w można odczytać z rysunku 8: krzywa 1 przedstawia prawdopodobieństwo I decyzji, krzywa 2 — I lub II itd.

TABLICA 2

Średnia wadliwość partii zgłoszonej do odbioru, w (w %)	1	2	4	6	8	10	12
średnia wadliwość partii po odsiewaniu 3-procent., w' (w %)	0,25	0,7	2	3,7	5,5	7,5	9,4
średnia wadliwość partii po odsiewaniu 10-procent., w'' (w %)	0,05	0,1	0,4	1,2	2,2	3,6	5,1
prawdopodobieństwo odrzucenia partii (moc testu), Q'' (w %)	0,001	0,2	7,7	31	60	81	91,5
prawdopodobieństwo odrzucenia partii lub odsiewania 10-procentowego, Q' (w %)	0,01	1,1	20	54	80	93	97,5
prawdopodobieństwo odrzucenia lub jakiegokolwiek odsiewania partii, Q (w %)	0,05	3	33	69	89	97	99
przeciętna wadliwość wychodząca, AOQ (w %)	1	1,95	2,99	2,69	1,82	1,07	0,57
przeciętna wadliwość wychodząca pod warunkiem, że partia jest przyjęta, $AOQ/(1-Q'')$ (w %)	1	1,95	3,25	3,90	4,55	5,63	6,70
średni procent sztuk dobrych odrzuconych przy przyjęciu partii, m (w %)	0	0,11	0,97	1,83	2,10	2,26	1,65

Także i w tym przykładzie wielkość próbki n jest równa wielkości próbki w przykładzie na test pierwotny. W ten sposób zasadnicze koszty badania obu testów są jednakowe; oczywiście wielkość próbki nie musi pozostawać stała.

Opisany wyżej test jest dużo słabszy niż test pierwotny, a więc „dostawca" jest tu bardziej uprzywilejowany niż w teście pierwotnym. Jednocześnie po łącznym przeanalizowaniu czwartego i ósmego wiersza tablicy 2, widzimy, że także „odbiorca" jest bardziej uprzywilejowany w opisanym tu teście niż w teście pierwotnym. Dzieje się tak za cenę pracy przy odsiewaniu i przy dość uciążliwych obliczeniach oraz za cenę odrzucenia pewnej liczby sztuk dobrych.

Prace cytowane

[1] J. Oderfeld, *Statystyczne badanie na cechy skorelowane*, Zastosowania Matematyki 4 (1959), str. 255–264.

[2] D. B. Owen, *Tables for computing bivariate normal probabilities*, Annals of Mathematical Statistics (1956), str. 1075.

INSTYTUT MATEMATYCZNY POLSKIEJ AKADEMII NAUK

Praca wpłynęła 15. 11. 1958

Э. ПЛЕЩИНСКАЯ (Варшава)

СОРТИРОВКА В СТАТИСТИЧЕСКОМ ИССЛЕДОВАНИИ КОРРЕЛИРОВАНЫХ ХАРАКТЕРИСТИК

РЕЗЮМЕ

В партиях товара, подвергнутых статистической проверке, нас интересует характеристика предметов Y (трудно исследуемая), о которой известно, что она скоррелирована с характеристикой X (легко исследуемой).

Автор работы [1] привел способ получения оценки значения средней характеристики Y, основанный на исследовании целой партии на характеристику X и только малой выборки на характеристику Y, и на этой основе составил критерий приема.

В настоящей работе предлагается связать исследование целой партии на легкую характеристику X с сортировкой партии, т. е. с отбрасыванием тех штук, в которых значения X дают возможность надеяться на удовлетворительное значение Y.

Первая часть работы содержит описание сортировки и доказательство того, что такая сортировка приводит в среднем к уменьшению плохого качества партии.

Вторая часть работы содержит примеры двух критериев приема с сортировкой и описание результатов применения этих критериев.

E. PLESZCZYŃSKA (Warszawa)

SCREENING IN STATISTICAL TESTING FOR CORRELATED CHARACTERISTICS

SUMMARY

In lots of goods which are subjected to statistical testing we are interested in a characteristic Y (difficult to test) which is known to be correlated with a characteristic X (easy to test).

The author of paper [1] gives a method of finding the estimator of the average value of characteristic Y based on the testing of the whole lot for the characteristic X and only a small sample for the characteristic Y: he works out an acceptance rule on this basis.

The author of the present paper suggests that the testing of the whole lot for the easy characteristic X should be combined with screening the lot, i. e. ~~with~~ rejecting pieces in which the value of X gives reason to expect unsatisfactory values of Y.

The first part of the paper contains a description of screening and a proof that screening of this kind brings about a mean reduction of the fraction defective.

The second part contains examples of two acceptance tests with screening, and a description of the consequences of using those tests.
