ZASTOSOWANIA MATEMATYKI
VI (1961)

J. MIKIEWICZ (Wrocław)

OCENA WSPÓLNEJ PRZECIĘTNEJ NA PODSTAWIE PRÓB Z POPULACYJ NORMALNYCH O RÓŻNYCH WARIANCJACH

1. W niniejszej pracy rozpatrzymy następujący model statystyczny. Istnieje k populacyj o rozkładach normalnych ze wspólną przeciętną μ i wariancjami: $\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2$. Z każdej z tych populacyj wybieramy losowo n_i elementów tak, że w sumie otrzymujemy $n = \sum_{i=1}^{k} n_i$ elementów. Zadaniem statystyka jest najlepsza ocena wspólnego parametru μ na podstawie statystyki, czyli funkcji próbki

$$(x_{11}, \ldots, x_{1n_1}, x_{21}, \ldots, x_{2n_2}, \ldots, x_{k1}, \ldots, x_{kn_k}). \tag{1}$$

Zagadnienie to ma znaczenie praktyczne. Wyobraźmy sobie bowiem, iż posiadamy pewien materiał obserwacyjny, dotyczący pomiarów jakiejś wielkości, ale otrzymany za pomocą przyrządów o różnej dokładności (np. pomiary odległości między dwoma punktami w terenie lub pomiary astronomiczne). Powstaje wtedy pytanie, jak najlepiej wykorzystać ten posiadany materiał? Pytanie to wiąże się z przedstawionym wyżej problematem statystycznym.

W dalszym ciągu podam najpierw rozwiązanie metodą „klasyczną”, a następnie pokażę, że rozwiązanie to jest także rozwiązaniem minimaksowym ze stanowiska teorii gier.

2. Mamy próbkę n-elementową, podzieloną na grupy złożone z n_i elementów w i-tej grupie, czyli złożone z niezależnych zmiennych losowych X_{ij}, które dla $i = \text{const}$ mają ten sam rozkład $N(\mu, \sigma_i^2)$ $(1 \leqslant j \leqslant n_i)$, to jest rozkład normalny o wspólnej przeciętnej μ oraz grupowej wariancji σ_i^2. Oznaczmy

$$Y_i = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij}. \tag{2}$$

Zmienna losowa Y_i ma rozkład normalny $N(\mu, \sigma_i^2/n_i)$. Utwórzmy statystykę $\sum_{i=1}^{k} a_i Y_i$, gdzie Y_i są to niezależne zmienne losowe o rozkładach $N(\mu, \sigma_i^2/n_i)$.

Zadanie polega na zminimizowaniu wariancji nieobciążonej sumy takich zmiennych losowych, a więc na znalezieniu współczynników $a_1^0, a_2^0, \ldots, a_k^0$ takich, żeby w klasie estymatorów postaci $\sum_{i=1}^{k} a_i Y_i$, spełniających warunek

$$E\Big(\sum_{i=1}^{k} a_i Y_i\Big) = \mu, \tag{3}$$

zachodziła równość

$$E\Big(\sum_{i=1}^{k} a_i^0 Y_i - \mu\Big)^2 = \min E\Big(\sum_{i=1}^{k} a_i Y_i - \mu\Big)^2. \tag{4}$$

Lecz $E\Big(\sum_{i=1}^{k} a_i Y_i\Big) = \sum_{i=1}^{k} a_i E(Y_i) = \sum_{i=1}^{k} a_i \mu$, a więc, na mocy warunku (3),

$$\sum_{i=1}^{k} a_i = 1. \tag{5}$$

Ponadto, wobec niezależności zmiennych X, jest

$$E\Big(\sum_{i=1}^{k} a_i Y_i - \mu\Big)^2 = E\Big(\sum_{i=1}^{k} a_i Y_i - E\Big(\sum_{i=1}^{k} a_i Y_i\Big)\Big)^2 = \sum_{i=1}^{k} a_i^2 E\big(Y_i - E(Y_i)\big)^2 =$$

$$= \sum_{i=1}^{k} \frac{a_i^2}{n_i^2} E\Big(\sum_{j=1}^{n_i} X_{ij} - E\Big(\sum_{j=1}^{n_i} X_{ij}\Big)\Big)^2 = \sum_{i=1}^{k} \frac{a_i^2}{n_i} \sigma_i^2.$$

Trzeba więc zminimizować wyrażenie $\sum_{i=1}^{k} \frac{a_i^2}{n_i} \sigma_i^2$ z warunkiem ubocznym (5). Łatwy rachunek prowadzi do wyrażenia

$$a_i = \frac{n_i}{\sigma_i^2} \Big(\sum_{j=1}^{k} \frac{n_j}{\sigma_j^2}\Big)^{-1}. \tag{6}$$

3. Założymy teraz, że z każdą oceną μ' parametru μ związana jest *strata*

$$L(\mu', \mu) = (\mu' - \mu)^2.$$

Nazwijmy *ryzykiem* oczekiwaną wartość straty związaną z danym estymatorem φ parametru μ. Mamy

$$R(\varphi, \mu) = E\{L[\varphi(z), \mu] \mid \mu\} \overset{\text{def}}{=} \int_Z L[\varphi(z), \mu] dF(z \mid \mu),$$

gdzie $z = (x_{11}, \ldots, x_{1n_1}, \ldots, x_{k1}, \ldots, x_{kn_k})$. Z oznacza tu przestrzeń próbek, tj. zbiór wszystkich możliwych wartości z. F jest dystrybuantą zmiennej losowej $Z = (X_{11}, \ldots, X_{1n_1}, \ldots, X_{k1}, \ldots, X_{kn_k})$.

Niniejszy ustęp poświęcony jest znalezieniu estymatora $\bar{\varphi}$, który minimizuje $\sup_{\mu} R(\varphi, \mu)$.

Wykażemy, że *estymator* $\bar{\varphi}$ *jest postaci*

$$\bar{\varphi}(z) = \sum_{i=1}^{k} a_i y_i, \quad gdzie \quad y_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij},$$

a współczynniki $a_1, \ldots, a_k$ *dane są wzorem* (6).

Niech Θ będzie klasą wszystkich możliwych dystrybuant a priori parametru μ, tj. niemalejących ciągłych lewostronnie funkcji $\theta(\mu)$, spełniających warunki: $\theta(-\infty) = 0, \theta(+\infty) = 1$ (¹).

Określmy oczekiwane ryzyko $r(\varphi, \theta)$:

$$(7) \qquad r(\varphi, \theta) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} R(\varphi, \mu) d\theta(\mu) = \int_{-\infty}^{\infty} \int_{Z} [\varphi(z) - \mu]^2 dF(z \mid \mu) d\theta(\mu).$$

Ponieważ $\sup_{\mu} R(\varphi, \mu) = \sup_{\theta \in \Theta} r(\varphi, \theta) \leqslant R(\bar{\varphi}, \mu)$ (²), jeśli $R(\bar{\varphi}, \mu)$ jest stałe przy wszelkim μ, zbadajmy wyrażenie $R(\bar{\varphi}, \mu)$:

$$R(\bar{\varphi}, \mu) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \left(\sum_i \frac{n_i}{\sigma_i^2} y_i \Big/ \sum_j \frac{n_j}{\sigma_j^2} - \mu \right)^2 \prod_{i=1}^{k} dF_i(y_i) =$$

$$= E_F \left(\sum_i \frac{n_i}{\sigma_i^2} y_i \Big/ \sum_j \frac{n_j}{\sigma_j^2} - \mu \right)^2 =$$

$$= E_F \left[\mu^2 - 2\mu \left(\sum_i \frac{n_i}{\sigma_i^2} y_i \Big/ \sum_j \frac{n_j}{\sigma_j^2} \right) + \left(\sum_i \frac{n_i}{\sigma_i^2} y_i \Big/ \sum_j \frac{n_j}{\sigma_j^2} \right)^2 \right] =$$

$$= \left[\mu^2 - 2\mu^2 \left(\sum_i \frac{n_i}{\sigma_i^2} \Big/ \sum_j \frac{n_j}{\sigma_j^2} \right) + \mu^2 \left(\sum_i \frac{n_i}{\sigma_i^2} \Big/ \sum_j \frac{n_j}{\sigma_j^2} \right)^2 + \right.$$

$$\left. + \sum_i \left(\frac{n_i}{\sigma_i^2} \right)^2 \frac{\sigma_i^2}{n_i} \Big/ \left(\sum_j \frac{n_j}{\sigma_j} \right)^2 \right] = \left(\sum_j \frac{n_j}{\sigma_j^2} \right)^{-1},$$

gdyż suma trzech pierwszych składników znika. Określmy

$$(8) \qquad W \stackrel{\text{def}}{=} \left(\sum_{j=1}^{k} \frac{n_j}{\sigma_j^2} \right)^{-1} = R(\bar{\varphi}, \mu).$$

(¹) Szersze omówienie użytych tu pojęć można znaleźć np. w pracy [3].

(²) Równość $\sup_{\mu} R(\varphi, \mu) = \sup_{\theta \in \Theta} r(\varphi, \theta)$ wynika z warunku $\int_{-\infty}^{\infty} d\theta(\mu) = 1$.

Wykazaliśmy, że $\sup_\mu R(\bar{\varphi}, \mu) \leqslant W$. Ponieważ dla dowolnego $\theta \in \Theta$ zachodzi podwójna nierówność

$$\inf_\varphi \sup_\mu R(\varphi, \mu) \geqslant \sup_\mu \inf_\varphi R(\varphi, \mu) \geqslant \inf_\varphi r(\varphi, \theta),$$

wystarczy wykazać, że do każdego $\varepsilon > 0$ istnieje taka dystrybuanta a priori θ_ε, że $\inf_\varphi r(\varphi, \theta_\varepsilon) > W - \varepsilon$.

Jak wiadomo (patrz [2]), przy ustalonej funkcji $\theta(\mu)$ funkcjonał (7) osiąga minimum dla

$$\varphi(z) = \frac{\int\limits_{-\infty}^{\infty} \mu f(z \mid \mu) d\theta(\mu)}{\int\limits_{-\infty}^{\infty} f(z \mid \mu) d\theta(\mu)}.$$

Oznaczmy przez $\varphi_a(z)$ estymator, który minimizuje oczekiwane ryzyko $r(\varphi, \theta_a)$, jeśli

$$\theta_a(\mu) = \begin{cases} 0 & \text{dla} \quad \mu < -a, \\ \dfrac{\mu}{2a} & \text{dla} \quad -a \leqslant \mu \leqslant a, \\ 1 & \text{dla} \quad a < \mu. \end{cases} \tag{9}$$

Mamy więc

$$\varphi_a(z) = \frac{\int\limits_{-a}^{a} \mu f(z \mid \mu) d\mu}{\int\limits_{-a}^{a} f(z \mid \mu) d\mu} = \frac{\int\limits_{-a}^{a} \mu \prod\limits_{i=1}^{k} \exp\left[-\frac{1}{2} \cdot \frac{n_i (y_i - \mu)^2}{\sigma_i^2}\right] d\mu}{\int\limits_{-a}^{a} \prod\limits_{i=1}^{k} \exp\left[-\frac{1}{2} \cdot \frac{n_i (y_i - \mu)^2}{\sigma_i^2}\right] d\mu} =$$

$$= \frac{\int\limits_{-a}^{a} \mu \exp\left[-\frac{1}{2}\left(\sum \frac{n_i}{\sigma_i^2} y_i^2 - 2\mu \sum \frac{n_i}{\sigma_i^2} y_i + \mu^2 \sum \frac{n_i}{\sigma_i^2}\right)\right] d\mu}{\int\limits_{-a}^{a} \exp\left[-\frac{1}{2}\left(\sum \frac{n_i}{\sigma_i^2} y_i^2 - 2\mu \sum \frac{n_i}{\sigma_i^2} y_i + \mu^2 \sum \frac{n_i}{\sigma_i^2}\right)\right] d\mu} =$$

$$= \frac{\int\limits_{-a}^{a} \mu \exp\left\{-\frac{1}{2}\left[\sum \frac{n_i}{\sigma_i^2} \left(\frac{\sum \frac{n_i}{\sigma_i^2} y_i}{\sum \frac{n_i}{\sigma_i^2}} - \mu\right)^2 + \sum \frac{n_i}{\sigma_i^2} y_i^2 - \frac{\left(\sum \frac{n_i}{\sigma_i^2} y_i\right)^2}{\sum \frac{n_i}{\sigma_i^2}}\right]\right\} d\mu}{\int\limits_{-a}^{a} \exp\left\{-\frac{1}{2}\left[\sum \frac{n_i}{\sigma_i^2} \left(\frac{\sum \frac{n_i}{\sigma_i^2} y_i}{\sum \frac{n_i}{\sigma_i^2}} - \mu\right)^2 + \sum \frac{n_i}{\sigma_i^2} y_i^2 - \frac{\left(\sum \frac{n_i}{\sigma_i^2} y_i\right)^2}{\sum \frac{n_i}{\sigma_i^2}}\right]\right\} d\mu} =$$

$$= \frac{\int\limits_{-a}^{a} \mu \exp\left[-\frac{1}{2}\sum \frac{n_i}{\sigma_i^2}\left(\frac{\sum \frac{n_i}{\sigma_i^2} y_i}{\sum \frac{n_i}{\sigma_i^2}} - \mu\right)^2\right] d\mu}{\int\limits_{-a}^{a} \exp\left[-\frac{1}{2}\sum \frac{n_i}{\sigma_i^2}\left(\frac{\sum \frac{n_i}{\sigma_i^2} y_i}{\sum \frac{n_i}{\sigma_i^2}} - \mu\right)^2\right] d\mu}.$$

Z wzoru (9) wynika, że

$$\lim_{a\to\infty} \varphi_a = \sum_i \frac{n_i}{\sigma_i^2} y_i \Big/ \sum_i \frac{n_i}{\sigma_i^2} = \bar{\varphi}.$$

Ponadto

$$r(\varphi_a, \theta_a) = \int\limits_{-a}^{a} E_F(\varphi_a - \mu)^2 d\theta(\mu) = \frac{1}{2a} \int\limits_{-a}^{a} (E_F\varphi_a^2 - 2\mu E_F\varphi_a + \mu^2) d\mu.$$

W dalszym ciągu skorzystamy z tożsamości, którą wyrazimy symbolicznie

(10) $$\varphi_a = J(\bar{\varphi}, a)[\zeta(\bar{\varphi}, a) - \zeta(\bar{\varphi}, -a)] + \bar{\varphi},$$

przy czym oznaczyliśmy

$$\zeta(\bar{\varphi}, a) = \omega \exp\left[-\frac{1}{2\omega}(\bar{\varphi}+a)^2\right], \quad \text{gdzie} \quad \frac{1}{\omega} = \sum_{i=1}^{k} \frac{n_i}{\sigma_i^2},$$

$$J(\bar{\varphi}, a) = \left\{\int\limits_{-a}^{a} \exp\left[-\frac{1}{2\omega}(\bar{\varphi}+u)^2\right] du\right\}^{-1}.$$

Tożsamość tę otrzymujemy z postaci (9) estymatora.

Korzystając z tożsamości (10) otrzymujemy dalej:

$$r(\varphi_a, \theta_a) = \frac{1}{2a} \int\limits_{-a}^{a} \Big\{E_F\bar{\varphi}^2 - 2\mu E_F\bar{\varphi} + \mu^2 + 2E_F\big[\bar{\varphi} J(\bar{\varphi}, a) \times$$

$$\times\big(\zeta(\bar{\varphi}, a) - \zeta(\bar{\varphi}, -a)\big)\big] - 2\mu E_F\big[J(\bar{\varphi}, a)\big(\zeta(\bar{\varphi}, a) -$$

$$- \zeta(\bar{\varphi}, -a)\big)\big] + E_F\big[J(\bar{\varphi}, a)\big(\zeta(\bar{\varphi}, a) - \zeta(\bar{\varphi}, -a)\big)\big]^2\Big\} d\mu.$$

Porównując w tym wyrażeniu składniki pod całką z przekształceniem doprowadzającym do (8), stwierdzamy, że

$$r(\varphi_a, \theta_a) = \frac{1}{2a}\int_{-a}^{a}(W+D_1+D_2)d\mu = W+\frac{1}{2a}\int_{-a}^{a}D_1 d\mu+\frac{1}{2a}\int_{-a}^{a}D_2 d\mu,$$

gdzie

$$D_1 = 2E_F\big[(\bar{\varphi}-\mu)\big(\zeta(\bar{\varphi}, a)-\zeta(\bar{\varphi}, -a)\big)J(\bar{\varphi}, a)\big],$$
$$D_2 = E_F\big[\big(\zeta(\bar{\varphi}, a)-\zeta(\bar{\varphi}, -a)\big)J(\bar{\varphi}, a)\big]^2.$$

Aby dowieść naszego twierdzenia wystarczy wykazać, że

$$\lim_{a\to\infty}\frac{1}{2a}\int_{-a}^{a}D_1 d\mu = \lim_{a\to\infty}\frac{1}{2a}\int_{-a}^{a}D_2 d\mu = 0. \tag{11}$$

Wyrażenie D_2 jest nieujemne; ponieważ jednak, jak wiemy z (8), zachodzi nierówność $W \geqslant r(\varphi_a, \theta_a)$, więc zbieżność wyrażenia $\frac{1}{2a}\int_{-a}^{a}D_1 d\mu$ implikuje zbieżność wyrażenia na D_2 ze wzoru (11).

Zbadajmy więc wyrażenie D_1. Wyrażenie to jest ograniczone, gdyż $|\zeta(\bar{\varphi}, a)-\zeta(\bar{\varphi}, -a)| < \omega$, przy czym dla $\bar{\varphi} < 0$, $\zeta(\bar{\varphi}, a)-\zeta(\bar{\varphi}, -a) > 0$, a dla $\bar{\varphi} > 0$, $\zeta(\bar{\varphi}, -a)-\zeta(\bar{\varphi}, a) > 0$; ponadto z określenia funkcji $J(\bar{\varphi}, a)$ wynika, że przy $a > 0$ $J(\bar{\varphi}, a) < \infty$ oraz dla $a \to \infty$, $J(\bar{\varphi}, a) \to (2\pi\omega)^{-1/2}$ dla każdego $\bar{\varphi}$.

Ograniczoność i zbieżność funkcji $|\zeta(\bar{\varphi}, a)-\zeta(\bar{\varphi}, -a)|J(\bar{\varphi}, a)$, przy $a \to \infty$, daje już łatwo dowód naszego twierdzenia. Jeśli przez dz oznaczymy element przestrzeni Z, to otrzymamy

$$\frac{1}{2a}\int_{-a}^{a}D_1 d\mu = \frac{1}{a}\int_{-a}^{a}\Big\{\int_{\bar{\varphi}<0}(\bar{\varphi}-\mu)|\zeta(\bar{\varphi}, a)-\zeta(\bar{\varphi}, -a)|J(\bar{\varphi}, a)f(z\mid\mu)dz -$$
$$\int_{\bar{\varphi}>0}(\bar{\varphi}-\mu)\,|\zeta(\bar{\varphi}, a)-\zeta(\bar{\varphi}, -a)|J(\bar{\varphi}, a)f(z\mid\mu)dz\Big\}d\mu.$$

Przekształćmy pierwszy składnik, majoryzując i korzystając z twierdzenia Fubiniego:

$$0 \leqslant \frac{1}{a}\int_{-a}^{a}\int_{\bar{\varphi}<0}(\bar{\varphi}-\mu)[\zeta(\bar{\varphi}, a)-\zeta(\bar{\varphi}, -a)]J(\bar{\varphi}, a)f(z\mid\mu)dzd\mu <$$
$$< \frac{\omega}{a}\int_{-a}^{a}\int_{\bar{\varphi}<0}(\bar{\varphi}-\mu)J(\bar{\varphi}, a)f(z\mid\mu)dzd\mu =$$
$$= \frac{\omega}{a}\int_{\bar{\varphi}<0}J(\bar{\varphi}, a)\Big[\bar{\varphi}\int_{-a}^{a}f(z\mid\mu)d\mu - \int_{-a}^{a}\mu f(z\mid\mu)d\mu\Big]dz.$$

Stąd, przechodząc do granicy, otrzymujemy

$$\lim_{a\to\infty}\frac{1}{a}\int\limits_{-a}^{a}\int\limits_{\bar{\varphi}<0}(\bar{\varphi}-\mu)[\zeta(\bar{\varphi},a)-\zeta(\bar{\varphi},-a)]J(\bar{\varphi},a)f(z\mid\mu)dzd\mu = 0,$$

bo

$$\lim_{a\to\infty}\int\limits_{-a}^{a}\int\limits_{\bar{\varphi}<0} = \omega\int\limits_{\bar{\varphi}<0}(2\pi\omega)^{-1/2}[\bar{\varphi}-\bar{\varphi}].$$

Analogicznie uzyskujemy zbieżność dla drugiej półprzestrzeni w Z. Stąd, tym bardziej zachodzi zbieżność dla całej przestrzeni, czyli dla $\frac{1}{2a}\int\limits_{-a}^{a}D_1 d\mu$.

Analogicznie dowodzimy zbieżności całki $\frac{1}{2a}\int\limits_{-a}^{a}D_2 d\mu$.

To kończy dowód twierdzenia.

Warto jeszcze zwrócić uwagę na fakt, że dla estymatora $\bar{\varphi}$ średni błąd kwadratowy $R(\bar{\varphi},\mu)$ jest niezależny od ocenianego parametru μ, a więc, jeśli wariancje $\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2$ są znane, stosując estymator $\bar{\varphi}$ możemy zawsze określić, ile średnio pomyliliśmy się w ocenie.

Prace cytowane

[1] M. Fisz, *Rachunek prawdopodobieństwa i statystyka matematyczna*, Warszawa 1958.

[2] J. L. Hodges Jr. and E. L. Lehmann, *Some problems in minimax point estimation*, Ann. Math. Statistics 21 (1950).

[3] S. Trybuła, *O minimaksowej estymacji parametrów w rozkładzie wielomianowym*, Zastosowania Matematyki 3 (1958), str. 307-322.

Praca wpłynęła 30. 10. 1960

ЯН МИКЕВИЧ (Вроцлав)

ОЦЕНКА ОБЩЕЙ СРЕДНЕЙ ПРИ ПОМОЩИ ВЫБОРОК ИЗ НОРМАЛЬНЫХ СОВОКУПНОСТЕЙ С РАЗНЫМИ ДИСПЕРСИЯМИ

РЕЗЮМЕ

Если из k совокупностей с той же средней μ, но с разными дисперсиями $\sigma_1^2, \ldots, \sigma_k^2$, выберем по n_i элементов так, чтобы $\sum_{i=1}^{k} n_i = n$, то при помощи такой выборки можно оценивать значение параметра μ.

В статье дается сперва „классическое" решение путем минимизации дисперсии взвешенной суммы случайных величин, а затем — минимаксовое решение при предположении квадратичной потери.

Оказывается, что оба решения тождественны, т.е. наилучшая оценка имеет вид

$$\mu' = \sum_{i=1}^{k} \frac{n_i}{\sigma_i^2} y_i \Big/ \sum_{i=1}^{k} \frac{n_i}{\sigma_i^2},$$

где y_i есть средняя выборки из i-ой совокупности.

J. MIKIEWICZ (Wrocław)

THE ESTIMATE OF THE COMMON MEAN FROM SAMPLES FROM NORMAL POPULATIONS WITH DIFFERENT VARIANCES

SUMMARY

Suppose we are given k normal populations with the same mean μ and different variances $\sigma_1^2, \ldots, \sigma_k^2$. If we choose samples of sizes n_i from these populations, where $\sum_{i=1}^{k} n_i = n$, we can estimate then mean μ with this sample. The paper presents at first the "classical" solution obtained by minimizing the variance of the weighted sum of averages of the chosen random variables, and then presents the minimax solution under the assumption of quadratic loss function. It turns out that these solutions are identical, i.e. the best estimate is of the form

$$\mu' = \sum_{i=1}^{k} \frac{n_i}{\sigma_i^2} y_i \Big/ \sum_{i=1}^{k} \frac{n_i}{\sigma_i^2},$$

where y_i is the sample mean from the sample from the i-th population.
