

*Dariusz Parys**

BOOTSTRAPOWY TEST WIELOKROTNY

Streszczenie. W artykule podjęto próbę zastosowania techniki bootstrapowej w testowaniu wielokrotnym. Przy testowaniu wielu hipotez jednocześnie należy zwrócić szczególną uwagę na ogólny poziom istotności definiowany jako prawdopodobieństwo odrzucenia co najmniej jednej prawdziwej hipotezy zerowej. Zaproponowano sekwencyjną procedurę testową służącą do weryfikacji hipotezy o odpowiednim uporządkowaniu średnich w populacjach reprezentowanych przez próby losowe.

W procedurze tej zastosowano technikę bootstrapową czyli taką, w której nieznaną rozkład zostaje zastąpiony jego empirycznym odpowiednikiem.

Słowa kluczowe: test wielokrotny, test sekwencyjnego odrzucania, technika bootstrapowa.

I. SEKWENCYJNY TEST WIELOKROTNY

Rozważmy sytuację, w której pobrano L prób o liczebnościach odpowiednio $n_1, n_2, \dots, n_L$ z L populacji. Porównujemy wartości oczekiwane $\mu_1, \dots, \mu_L$, w celu stwierdzenia, które z tych parametrów są różne oraz ustalenia ich porządku w kolejności malejącej.

Jeżeli $L > 2$, wtedy mamy do czynienia z problemem porównań wielokrotnych (por. Miller 1980). Testujemy hipotezę zerową $H_0: \mu_1 = \dots = \mu_L$. Wydaje się jednak ona mniej odpowiednia dla testowania porównań parami, stąd hipotezami zerowymi będą hipotezy postaci:

$$H_{0ij}: \mu_i \leq \mu_j, \quad i, j = 1, 2, \dots, L; \quad i \neq j \quad (1)$$

natomiast alternatywne hipotezy będą postaci:

$$H_{1ij}: \mu_i > \mu_j, \quad i, j = 1, 2, \dots, L; \quad i \neq j \quad (2)$$

* Dr, Instytut Ekonometrii i Statystyki, Uniwersytet Łódzki.

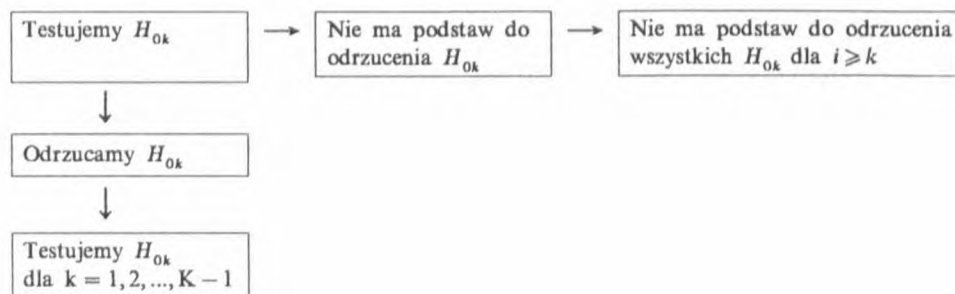
Niech $D_{ij} = \mu_i - \mu_j$ oraz $\hat{D}_{ij} = \bar{y}_i - y_j$ będzie estymatorem D_{ij} . Ustawiamy $\hat{D}_{ij}$ w kolejności malejącej, oznaczając $\hat{D}_1$ jako największą różnicę, aż do najmniejszej $\hat{D}_K$, gdzie $K = L(L-1)$.

Hipotezy (1) i (2) porządkujemy w ten sam sposób. Biorąc pod uwagę oczywisty fakt, że $\mu_i \leq \mu_j \Leftrightarrow \hat{D}_{ij} \leq 0$, możemy zapisać hipotezy w następujący sposób:

$$H_{0k}: D_k \leq 0; k = 1, 2, \dots, K \quad (3)$$

$$H_{1k}: D_k > 0; k = 1, 2, \dots, K \quad (4)$$

Hipotezy (3) testujemy w sposób sekwencyjny (por. Holm 1977):



II. METODA BOOTSTRAPOWA

Bootstrap (Efron 1982) jest metodą pobierania prób w celu estymacji rozkładów z próby bez posiadania informacji o rozkładzie populacji.

Rozważmy próbę o liczebności n z populacji o nieznanym dystrybucie rozkładu F .

$$x_1, \dots, x_n \sim F \quad (5)$$

Niech $\hat{\theta}$ będzie funkcją z tej próby

$$\hat{\theta} = \hat{\theta}(x_1, \dots, x_n) \quad (6)$$

i niech $\hat{F}$ oznacza empiryczny, bootstrapowy rozkład prawdopodobieństwa.

Losujemy próbę $x_1^*, \dots, x_2^*, \dots, x_n^*$ i obliczamy wartości statystyki

$$\hat{\theta}^* = \hat{\theta}(x_1^*, x_2^*, \dots, x_n^*), x_i^* \sim \hat{F} \quad (7)$$

z próby bootstrapowej.

Taka procedura jest powtarzana B razy, dając powtórzenia estymatorów $\hat{\theta}_1^*, \hat{\theta}_2^*, \dots, \hat{\theta}_B^*$. Otrzymujemy więc rozkład z próby estymatora $\hat{\theta}$ oraz estymator wariancji $\hat{\theta}$:

$$D^2(\hat{\theta}) = \frac{1}{B-1} \sum_{i=1}^B (\hat{\theta}_i^* - \hat{\theta}^*)^2 \quad (8)$$

$$\text{gdzie } \hat{\theta} = \frac{1}{B} \sum_{i=1}^B \hat{\theta}_i^*$$

III. BOOTSTRAPOWY TEST WIELOKROTNY

Idea bootstrapowej procedury testu wielokrotnego opiera się na utworzeniu wszystkich możliwych porównań parami spośród L średnich z prób $\bar{y}_1, \bar{y}_2, \dots, \bar{y}_L$ oraz stwierdzeniu, czy obserwowane różnice między średnimi w próbach bootstrapowych mogą pojawiać się losowo i czy można na ich podstawie wnioskować o różnicach między odpowiednimi parametrami.

Od każdej obserwacji odejmujemy jej średnią $\bar{y}_i$. Przekształcamy w ten sposób L prób tak, aby w każdej z nich średnia wynosiła zero.

Z tak przekształconych rozkładów generujemy próbę bootstrapową oraz testując hipotezę (3) obliczamy różnicę bootstrapową $\hat{D}_k^*$. Niech $\bar{y}_i^*$ oznacza średnią z próby bootstrapowej dla $i = 1, 2, \dots, L$. Ponieważ $E(\bar{y}_i^*) = 0$, stąd $E(\hat{D}_k^*) = 0$ dla $k = 1, \dots, K$. Wtedy każde odchylenie od zera dla $\hat{D}_k^*$ jest losowe.

Porównujemy prawdziwe różnice z prób z bootstrapowymi i mamy podstawę do stwierdzenia, czy te pierwsze pojawiają się przypadkowo, czy wskazują na różnice pomiędzy parametrami.

Wprowadźmy teraz zmienne I_k , $k = 1, \dots, K$, które będą nam wskazywać na powyższe różnice. Wartości I_k przyjmowane są w następujący sposób: jeżeli $\hat{D}_k^* < \hat{D}_k$, $k = 1, \dots, K$ oraz $k' = k, k+1, \dots, K$, wtedy $I_{k'} = I_k + 1$. Wiemy więc, ile jest przypadków w próbie bootstrapowej takich, że $\hat{D}_k^* < \hat{D}_k$. Dla największej różnicy $\hat{D}_1$ stwierdzamy, która z różnic bootstrapowych jest tą największą. W przeciwnym przypadku $\hat{D}_1 > 0$ i dla D_2 stwierdzamy, która z różnic bootstrapowych (oprócz $\hat{D}_1^*$) jest tą największą itd. dla K różnic z próby.

Cała procedura jest powtarzana B -razy, gdzie B jest liczbą z przedziału (1000, 10 000). Oczywiście $0 \leq I_k \leq B$; $k = 1, 2, \dots, k$ i łatwo zauważyć, że jeżeli liczba wskázówek I_k dla hipotezy H_{0k} jest dostatecznie duża, wtedy H_{0k} może być odrzucona. Prawdopodobieństwo odrzucenia H_{0k} na podstawie wyników z próby, jeżeli w rzeczywistości jest ona prawdziwa, jest równe

I_k/B . Jeżeli wartość krytyczna byłaby równa $\hat{D}_k$, to poziom istotności byłby równy I_k/B . Decyzja przejścia lub odrzucenia ma charakter sekwencyjny. Jeżeli ogólny poziom istotności jest ustalony jako α , to wtedy dla $k = 1$:

- jeżeli $I_k/B \leq \alpha$ to odrzucamy H_{0k} i testujemy hipotezę H_{0k+1}
- jeżeli $I_k/B > \alpha$, to nie ma podstaw do odrzucenia H_{0i} dla $i \geq k$.

Można zauważyć, że hipoteza H_{0k} nie może być odrzucona tak długo, dopóki nie odrzucimy wszystkich hipotez H_{0i} dla $i < k$.

Rozważmy przypadek dla $L = 3$. Mamy do czynienia z następującymi możliwymi zależnościami między badanymi średnimi:

$$\mu_1 = \mu_2 = \mu_3 \quad (9)$$

$$\mu_1 = \mu_2 \neq \mu_3 \quad (10)$$

$$\mu_1 \neq \mu_2 = \mu_3 \quad (11)$$

$$\mu_1 = \mu_2 \neq \mu_3 \quad (12)$$

$$\mu_1 \neq \mu_2 \neq \mu_3 \wedge \mu_1 \neq \mu_3 \quad (13)$$

Zależność (9) odpowiada hipotezie zerowej (1). Przypuśćmy, że pierwsza hipoteza $\mu_i < \mu_j$, $i, j = 1, 2, \dots, L$ jest odrzucona, gdyż $\bar{y}_i - \bar{y}_j$ jest największą różnicą z próby. Stwierdzamy więc, że $\mu_i > \mu_j$. W następnym kroku nie musimy rozważać żadnej hipotezy zawierającej $\mu_i = \mu_j$. Niech na przykład $i = 1$ i $j = 3$. Po stwierdzeniu, że $\mu_1 > \mu_3$ pozostają nam tylko trzy możliwości: (10), (11) i (13), ponieważ zależności między średnimi (9) i (12) zawierają $\mu_1 = \mu_3$. Wynika stąd, że testowanie następnej hipotezy zerowej, powiedzmy $\mu_1 \leq \mu_2$, przeprowadzamy zgodnie z jednym założeń (10), (11) lub (13).

Ponieważ nie wiemy, które z nich jest prawdziwe, testowanie przeprowadzamy osobno dla (10), (11) i (13), a największą liczbę wskazówek I_k wykorzystujemy do weryfikacji hipotezy. Wzrost mocy jest oczywisty. Liczone są różnice bootstrapowe między pierwszą i drugą średnią oraz między drugą i trzecią (11), gdy testowanie bez rozważania możliwości (9)–(13) oznacza testowanie wszystkich różnic jednocześnie. Jeżeli $\mu_1 \neq \mu_3$, to w tym samym czasie jest możliwe, aby $\mu_1 = \mu_2$ i $\mu_2 = \mu_3$. Procedura testowa może zakończyć się w każdym jej kroku. Wynika stąd, że w przypadku odrzucenia pierwszej hipotezy, dalsze wnioski nie są już konieczne. Pojawia się tutaj pewna sugestia, aby odrzucać tak dużo hipotez, jak tylko jest to możliwe bez zakończenia procedury oraz włączenia tego kroku od pierwszej hipotezy, która nie była odrzucona. Cała procedura testowa zatrzymuje się, gdy nawet końcowy krok nie powoduje odrzucenia pewnej hipotezy.

Powyższa metoda wydaje się bardziej ogólna niż porównania wielokrotne. Procedura bootstrapowa zastępuje teoretyczne rozkłady ich empirycznymi odpowiednikami. Nie występuje tutaj jakiegokolwiek ograniczenie liczby testowanych hipotez, liczby obserwacji lub różnych rozmiarów prób.

LITERATURA

- Efron B. (1982), *The Jackknife, the Bootstrap and Other Resampling Plans*, Society for Industrial and Applied Mathematics, Philadelphia.
- Holm S. (1977), *Sequentially Rejective Multiple Procedures*, Statistical Research Report 1977-1, University of Umea, Institute of Mathematics and Statistics.
- Miller R. G. Jr (1980), *Simulations Statistical Inference*, 2nd edn., Springer Verlag, New York.

Dariusz Parys

THE BOOTSTRAP MULTIPLE TEST

(Summary)

In this paper a new multiple test procedure is presented. This procedure is based on the bootstrap technique.

We consider a situation where L samples, of sizes $n_1, n_2, \dots, n_L$, from L distributors, with expected values $\mu_1, \dots, \mu_L$, are to be compared. The null and alternative hypotheses are the following

$$H_{0ij}: \mu_i \leq \mu_j; \quad i, j = 1, 2, \dots, L \text{ and } i \neq j$$

$$H_{1ij}: \mu_i > \mu_j; \quad i, j = 1, 2, \dots, L \text{ and } i \neq j$$

These hypotheses are tested in sequentially rejective manner and the test in each step is performed by means of the bootstrap technique. Compared to other methods this procedure is very general according to number of hypotheses, sample distribution and number of hypotheses.