Józef Gajowski

Zygmunt Przybycin

WŁASNOŚCI ESTYMATORÓW CHARAKTERYSTYK FUNKCYJNYCH POLA LOSOWEGO

Wprowadzenie

Pole losowe jest probabilistycznym modelem zjawisk losowych zachodzących w przestrzeni wielowymiarowej. W warunkach rzeczywistych modelu tego na ogół nie znamy, możemy jedynie założyć znajomość realizacji pola losowego, które otrzymujemy przez obserwację zjawiska bądź za pomocą eksperymentu. Identyfikację modelu probabilistycznego (wyznaczenie charakterystyk pola losowego na podstawie zebranych danych) nazywamy wnioskowaniem statystycznym. Podstawą wnioskowania statystycznego jest model statystyczny (przestrzeń statystyczna).

Aby zdefiniować przestrzeń statystyczną, indukowaną przez pola losowe, wprowadzimy następujące oznaczenia i pojęcia:

Niech C jest zbiorem wartości, jakie przyjmuje zmienna losowa $Y(x)$, $x \in T \subset R^k$, $k \geq 2$ ($Y(\cdot)$ jest polem losowym), zaś $\mathcal{B} = B(C)$ jest sigma ciałem generowanym przez zbiór C. Wówczas $P(B_i) = P(Y(x_i) \in B_i$, $B_i \in \mathcal{B}(C)$, $i = 1,2,...,l$ jest skończenie wymiarowym rozkładem prawdopodobieństwa pola losowego. Rozkład ten jest zazwyczaj nieznany, jednak wiedza statystyczna dotycząca badanego zjawiska pozwala na określenie rodziny rozkładów $\mathrm{P} = \{P_q(B_i) : q \in Q,\ \ i = 1,..,l\}$ indeksowanej nieznanym parametrem q. System $< C, \mathcal{B}(C), \mathcal{P} >$ będziemy nazywać przestrzenią statystyczną indukowaną przez pole losowe $Y(\cdot)$.

Model statystyczny będący sformalizowanym opisem badanego zjawiska losowego powinien dostarczyć niezbędnej wiedzy o mechanizmach i procesach nim rządzących; jest on przeważnie kompromisem pomiędzy prostotą modelu a jego adekwatnością.

Nośnikiem informacji, na podstawie których prowadzi się wnioskowanie statystyczne, jest próba pola losowego.

Załóżmy, że w przestrzeni $T \subset R^k$ jest określona rodzina podzbiorów $\{\Delta_i\}$, $i=1,...,n$ (Δ_i może być punktem) i niech na każdym zbiorze Δ_i tej rodziny jest określone pole losowe $(Y_i(x),\ x \in \Delta_i\ )$, $i=1,...,n$. Zakładamy ponadto, że pole losowe $(Y_i(x),\ x \in \Delta_i)$ jest równoważne stochastycznie polu losowemu $(Y(x),\ x \in T)$, tzn. skończenie wymiarowe rozkłady tego pola należą do rodziny P dla $i=1,...,n$.

Ciąg niezależnych pól losowych $\{(\ Y_i(x),\ x \in \Delta_i\ )\}$, $i=1,...,n$ nazywamy n-elementową próbą losową pola losowego $Y(\cdot)$. Zauważmy, że ciąg pól losowych $\{(Y_i(x),\ x \in \Delta_i)\}$, $i=1,...,n$ może być ciągiem zmiennych losowych; będzie tak wówczas, gdy zbiory rodziny $\{\Delta_i\}$, $i=1,...,n$ zredukują się do punktów obszaru T.

Próba pola losowego generuje sigma ciało będące iloczynem kartezjańskim sigma ciał generowanych przez poszczególne pola losowe należące do próby. To sigma ciało będziemy nazywać przestrzenią prób i oznaczać symbolem W .

Dokonując pomiaru obserwacji pola losowego na zbiorach rodziny $\{\Delta_i\}$, $i=1,...,n$ otrzymujemy ciąg $\{y(\Delta_i\ )\}$, $i=1,...,n$, który jest realizacją próby pola losowego oraz elementem przestrzeni prób.

1. Estymacja charakterystyk funkcyjnych pola losowego

Rodzina skończenie wymiarowych rozkładów daje pełny probabilistyczny opis pola losowego, jednak w praktyce często zastępuje się go uproszczonym opisem, wykorzystując do tego odpowiednio dobrane charakterystyki funkcyjne. Charakterystykami funkcyjnymi pola losowego są:

– funkcja wartości oczekiwanej,
– funkcja korelacyjna.

Funkcje te są zwykle nieznane, dlatego należy je oszacować na podstawie próby pola losowego.

Rozważmy dwa rodzaje próby pola losowego:
– próba ciągła,
– próba dyskretna.

Przejdziemy najpierw do konstrukcji próby ciągłej pola losowego.

Przypuśćmy, że zbiory rodziny $\{\Delta_i\}$, $i=1,\dots,n$ należą do przestrzeni B^k, $k \geq 2$ i są liniowo uporządkowane relacją zawierania się zbiorów, a niezależne pola losowe $(Y_i(x),\ x \in \Delta_i)$ są równoważne stochastycznie polu losowemu $(Y(x),\ x \in T \cap \Delta_i)$. Ciąg $\{(Y_i(x),\ x \in \Delta_i)\}$, $i=1,\dots,n$ będziemy nazywać próbą ciągłą pola losowego. Do wyznaczania charakterystyk funkcyjnych pola losowego, przy założeniu znajomości próby ciągłej, stosuje się rezultaty teorii pól losowych omówione w pracy Gajowskiego i Przybycina (2002).

Estymator funkcji wartości oczekiwanej jednorodnego pola losowego przyjmuje postać:

$$\hat{m} = \frac{1}{|\Delta|} \int_{\Delta} Y(x)\, dx, \quad \Delta \in \{\Delta_i\},\ i = 1,\dots,n \tag{1}$$

gdzie $|\bullet|$ oznacza miarę zbioru.

Można udowodnić (Lieonienko, Iwanow, 1980), że estymator (1) funkcji wartości oczekiwanej jednorodnego pola losowego jest estymatorem zgodnym i nieobciążonym.

Dowodzi się ponadto, że warunkiem dostatecznym asymptotycznej zbieżności rozkładu estymatora (1) do rozkładu normalnego jest to, aby funkcja gęstości spektralnej jednorodnego pola losowego reprezentującego szum spełniała warunek f (0) > 0. Przy spełnieniu postawionego warunku, rozkład asymptotyczny estymatora funkcji wartości oczekiwanej jednorodnego pola losowego przyjmuje postać:

$$N\left(m_0,\ V\overline{(2\pi)^n f(0)}/|\Delta|\right) \text{ dla } \mathrm{m}_0 = \mathrm{E}(\mathrm{Y}(\cdot))$$

W pracy Przybycina (1995) określono warunki asymptotycznej efektywności estymatora (1). Okazuje się, że warunkiem koniecznym i dostatecznym asymptotycznej efektywności rozważanego estymatora jest to, aby funkcja gęstości spektralnej pola losowego reprezentującego szum była funkcją stałą. Oznacza to, że jeżeli w świetle przeprowadzonych badań statystycznych lub innych przesłanek, założenie o stałości funkcji gęstości spektralnej może być przyjęte, to otrzymany estymator wartości oczekiwanej jednorodnego pola losowego jest najlepszym estymatorem tego parametru w zbiorze estymatorów dopuszczalnych. Fakt ten, jak również prostota estymatora, przemawiają za jego praktycznym zastosowaniem.

Drugą charakterystyką funkcyjną pola losowego jest funkcja korelacyjna. Funkcję tę można wyznaczyć na podstawie realizacji badanego pola przy uwzględnieniu dodatkowych informacji, co do możliwych klas tej funkcji. Jednym z estymatorów funkcji korelacji jednorodnego pola losowego o zerowej wartości oczekiwanej jest wyrażenie:

$$\hat{R}(h) = \frac{1}{|\Delta|} \int_{\Delta} Y(x) \cdot Y(x+h)\, dx, \qquad x, h \in \Delta \tag{2}$$

W pracy Lieonienki i Iwanowa (1980) wykazano, że estymator (2) jest zgodnym estymatorem funkcji korelacji jednorodnego pola losowego, co w praktyce oznacza, że prawdopodobieństwo popełnienia błędu w szacunku funkcji korelacji w estymacji punktowej za pomocą tego estymatora można uczynić dowolnie małym poprzez odpowiedni dobór obszaru Δ.

W literaturze sformułowano warunki asymptotycznej efektywności rozważanego estymatora. Warunki te, z uwagi na ich złożoność, mają jedynie znaczenie teoretyczne, gdyż w praktyce są one trudne, a w niektórych przypadkach wręcz niemożliwe do zweryfikowania. Okazuje się jednak, że jeżeli rozważane pole losowe jest gaussowskim polem, to można sformułować stosunkowo proste warunki asymptotycznej efektywności estymatora (2).

Wprowadźmy następujące oznaczenie:

$$U_h(x) = Y(x+h) \cdot Y(x), \qquad x, h \in T \tag{3}$$

Pole $\mathrm{U_h}(\mathrm{x})$ jest iloczynem gaussowskiego pola losowego $Y(\cdot)$ w punktach x oraz x + h. W pracy Przybycina (1995) wykazano, że rozważane pole losowe $U_h(x)$ jest jednorodnym polem losowym. Wystarczy teraz zauważyć, że estymator (3) wartości oczekiwanej pola losowego $U_h(\cdot)$ jest równocześnie estymatorem funkcji korelacji pola losowego $Y(x)$. Uwzględniając własność estymatora wartości oczekiwanej jednorodnego pola losowego, można sformułować warunki asymptotycznej normalności tego estymatora oraz warunki asymptotycznej efektywności. Oznacza to, że jeżeli funkcja gęstości pola losowego reprezentująca szum jest dodatnia w punkcie zero, to rozkład estymatora (3) jest asymptotycznie normalny; jeżeli ponadto funkcja ta jest stała, to estymator ten jest również asymptotycznie efektywny.

Uwzględniając fakt asymptotycznej zbieżności estymatora wartości oczekiwanej jednorodnego pola losowego do rozkładu normalnego można wyznaczyć przedział ufności tego parametru. Ponieważ:

$$(\hat{m} - m_0) \xrightarrow{D} N_1\left(O, \sqrt{\frac{(2\Pi)^n f(0)}{|\Delta|}}\right)^*$$

więc można napisać:

$$P\{-t_\alpha < \hat{m} < t_\alpha\} = 1 - a \tag{4}$$

gdzie:

$$\hat{m} = \frac{\hat{m} - m_0}{\sqrt{(2\Pi)^n f(0)}} \cdot \sqrt{|\Delta|}$$

Dla ustalonego współczynnika ufności α z tablic rozkładu normalnego odczytujemy liczbę t_α. Po przekształceniu zależności (4) otrzymujemy przedział ufności dla wartości oczekiwanej jednorodnego pola losowego:

$$P\left\{\hat{m} - t_\alpha\sqrt{\frac{(2\Pi)^n f(0)}{|\Delta|}} < m_0 < \hat{m} + t_\alpha\sqrt{\frac{(2\Pi)^n f(0)}{|\Delta|}}\right\} = 1 - a \tag{5}$$

Wprowadzając do wzoru (5) w miejsce $\hat{m}$ ocenę wartości oczekiwanej jednorodnego pola losowego $\tilde{m}$, otrzymujemy liczbowy przedział ufności w postaci:

$$\tilde{m} - t_\alpha\sqrt{\frac{(2\Pi)^n f(0)}{|\Delta|}}, \tilde{m} + t_\alpha\sqrt{\frac{(2\Pi)^n f(0)}{|\Delta|}}$$

Analogicznie wyznacza się przedział ufności dla funkcji korelacyjnej gaussowskiego pola losowego.

Drugim rodzajem próby pola losowego jest próba punktowa. Jeżeli zbiory rodziny $\{\Delta_i\}$, i=1,...,n redukują się do punktów przestrzeni $T \subset R^k$, $k \geq 2$, to $Y(x_i), \{x_i\} = \Delta_i$ jest zmienną losową.

Ciąg zmiennych losowych $\{(Y(x_i), x_i \in T)\}$, i=1,...,n nazywamy próbą dyskretną (punktową) pola losowego *Y(·)*. Dla próby dyskretnej estymatory parametrów funkcyjnych pola losowego są analogiczne do estymatorów parametrów zmiennej losowej.

---

* Rozważa się tutaj tzw. słabą zbieżność. Szerzej na ten temat w pracy Lieonienki i Iwanowa (1980).

Estymator funkcji wartości oczekiwanej jednorodnego pola losowego przyjmuje postać:

$$\overline{m} = \frac{1}{n}\sum_{i=1}^{n} Y(x_i) \qquad (6)$$

Dla jednorodnego gaussowskiego pola losowego estymator (6) ma rozkład normalny z wartością oczekiwaną m_0 oraz wariancją $\sigma_y^2 = \frac{R(0)}{n}$, gdzie $R(x),\ x \in T$ jest funkcją korelacyjną pola losowego $Y(\cdot)$.

Estymator (6) jest estymatorem nieobciążonym; ponadto jeżeli funkcja korelacyjna $R(\cdot)$ spełnia warunek ergodyczności, to jest on również estymatorem zgodnym.

Estymatorem funkcji korelacji jednorodnego pola losowego jest wyrażenie:

$$\hat{R}(h) = \frac{1}{|N(h)|}\sum_{N(h)}\left(Y(x_i) - \overline{Y}\right)\left(Y(x_j) - \overline{Y}\right) \qquad (7)$$

gdzie:

$$N(h) = \{(i,j) : x_i - x_j = h\},\ x_i, x_j \in T$$

$|N(h)|$ oznacza liczebność zbioru $N(h)$

oraz:

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y(x_i)$$

Wartość oczekiwana estymatora (7) (Cressie, 1993) jest równa:

$$\mathrm{E}\left[\hat{R}(h)\right] = R(h) - \frac{\left[R(0) + 2\sum_{i=1}^{r} R(i)\right]}{n} + o\left(\frac{1}{n^2}\right)$$

gdzie r oznacza stopień skorelowania pola losowego z wartością średnią z próby losowej.

Z ostatniej równości wynika, że estymator (7) funkcji korelacyjnej jest estymatorem asymptotycznie nieobciążonym.

2. Schematy losowania próby dyskretnej

Własności asymptotyczne estymatora współczynników regresji liniowej pola losowego przedstawione w pracy Lieonienki i Iwanowa (1980) zostały otrzymane przy założeniu zmierzania rodziny podzbiorów $\{\Delta_i\}, i = 1,...,n$ do zbioru granicznego według Van Hoffa. Oznacza to, iż sposób wyboru punktów pomiarowych w próbie dyskretnej, ich rozmieszczenie w obszarze Δ istotnie wpływa na „dobroć" estymatora parametrów pola losowego. Wpływ lokalizacji punktów pomiarowych na „dobroć" estymacji prześledzimy na wybranych schematach losowania, ograniczając się do estymatorów wartości średniej pola losowego.

Załóżmy, że ciągłe pole losowe ($\mathrm{Y}(\mathrm{x}),\ \mathrm{x} \in \mathrm{T}$) jest rozważane na pewnym obszarze $\Delta \subset \mathrm{T}$; wartość średnią pola na tym obszarze określamy następująco:

$$\hat{Y}(\Delta) = \frac{1}{|\Delta|}\int_{\Delta} Y(x)dx \tag{8}$$

Warto zauważyć, że wartość średnia określona na ciągłej próbie losowej wzorem (8) jest jednocześnie estymatorem funkcji wartości oczekiwanej jednorodnego pola losowego.

Przypuśćmy, że w obszarze Δ została wyznaczona próba punktowa pola losowego $Y(\cdot)$, tzn. dany jest ciąg $\{Y(x_i)\},\ x_i \in \Delta$, *i=1,...,n*. Wartość średnią (8) będziemy aproksymować średnią arytmetyczną z próby punktowej, tj. wyrażeniem:

$$\overline{\mathrm{Y}} = \frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{Y}(\mathrm{x}_{\mathrm{i}}) \tag{9}$$

Oczywiście średnia z próby jest równocześnie estymatorem funkcji wartości oczekiwanej jednorodnego pola losowego dla próby dyskretnej.

Standardowymi schematami losowania próby dyskretnej pola losowego są:

- losowanie proste nieograniczone,
- losowanie warstwowe,
- losowanie systematyczne.

Załóżmy, że dane jest pole losowe ($Y(x),\ \mathrm{x} \in \mathrm{T}$). W przypadku losowania prostego nieograniczonego próbę otrzymujemy w następujący sposób:

- z obszaru $\Delta \subset T$ wybieramy losowo n punktów pomiarowych x_i; każdy punkt pomiaru jest wybierany niezależnie od pozostałych i od pola losowego z tym samym dodatnim prawdopodobieństwem, natomiast punktowi x_i jest przyporządkowana zmienna losowa $Y(x_i)$,
- ciągowi punktów $\{x_i\}$ odpowiada ciąg zmiennych losowych $\{Y(x_i)\}$, i=1,...,n, który stanowi próbę; próbę tę dalej będziemy nazywać próbą losową prostą; dla ustalonego zdarzenia $\omega \in \Omega$ otrzymujemy ciąg $\{y(x_i)\}$, i=1,...,n, tj. realizację próby prostej.

W omówionym schemacie losowania będziemy zakładali, że punkt pomiaru już raz wylosowany nie bierze udziału w dalszym losowaniu. Jest to tzw. wariant losowania bezzwrotnego, w odróżnieniu od losowania zwrotnego, tj. takiego, w którym punkt pomiaru już raz wylosowany bierze udział w dalszym losowaniu. Można pokazać, że oba warianty losowania, przy dostatecznie dużej próbie, osiągają rezultaty różniące się nieistotnie.

Losowanie warstwowe polega na tym, że:

- obszar Δ dzielimy na rozłączne warstwy $\Delta_1, \Delta_2, ..., \Delta_k$; z każdej warstwy Δ_j, j=1,...,k w sposób niezależny wybieramy n_j punktów pomiaru $x_i^j, i = 1,...,n_j$,
- ciągowi $\{x_i^j\}$ odpowiada ciąg zmiennych losowych $\{Y_j(x_i)\}$, $j = 1,...,k$, $i = 1,...,n_j$; wylosowane próby z poszczególnych warstw tworzą n-elementową próbę warstwową, przy czym $n = \sum_{j=1}^{k} n_j$.

Podziału obszaru Δ na warstwy powinno dokonywać się według kryterium, które można łatwo zastosować. Ponadto warstwowanie powinno być przeprowadzone w ten sposób, aby każdy element zbioru Δ należał tylko do jednej warstwy. Sposób podziału zbiorowości na warstwy zależy od konkretnych warunków, jakie występują w danym badaniu, również liczba warstw jest z reguły sprawą otwartą.

W schemacie losowania systematycznego próbę otrzymujemy następująco:

– obszar Δ dzielimy na rozłączne, przystające* zbiory $\Delta_1, \Delta_2, ..., \Delta_n$; z każdego z podzbiorów Δ_i, $i = 1,...,n$ wybieramy według pewnej reguły punkt pomiaru x_i,
– ciągowi $\{x_i\}$ odpowiada ciąg zmiennych losowych $\{Y(x_i)\}$ – próba systematyczna.

Sposób wyboru punktu pomiarowego $x_i \in \Delta_i$ może być dowolny. Najczęściej jest stosowany tzw. wybór centryczny – punkt x_i jest wtedy środkiem symetrii zbioru Δ_i.

Zubrzycki (1957) zaproponował następujący sposób wyboru punktów pomiarowych x_i:

– ze zbioru Δ_{i-1} w sposób losowy wybieramy punkt startowy x_{i-1}, $i \geq 2$,
– punkt $x_i \in \Delta_i$ wybieramy w ten sposób, że po dokonaniu przesunięcia zbioru Δ_{i-1} o wektor $[x_{i-1}, x_i]$, zbiór Δ_{i-1} pokryje się ze zbiorem Δ_i.

Losowanie próby według schematu systematycznego jest narażone na niebezpieczeństwo otrzymania próby niereprezentatywnej. Sytuacja taka może wystąpić wówczas, gdy badane pole losowe wykazuje tendencje do wahań periodycznych. Może się tak zdarzyć, że punkt startowy losowania pokryje się z punktem ekstremalnym badanego pola; jeżeli ponadto długość cyklu pokrywa się z długością wektora $[x_{i-1}, x_i]$, to dokonane na podstawie takiej próby charakterystyki będą istotnie różniły się od ich wartości rzeczywistych.

Planując badanie statystyczne należy bardzo ostrożnie podejmować decyzje co do ewentualnego systematycznego pobierania próby, gdyż większość procesów występujących w przyrodzie oraz życiu gospodarczym wykazuje wahania periodyczne będące efektem zjawisk astronomicznych bądź konsekwencją określonych zwyczajów czy norm prawnych. Jeżeli dodatkowe informacje o badanym polu losowym przemawiają za systematycznym pobraniem próby, to decyzja o takim wyborze musi być poprzedzona wnikliwą analizą wahań periodycznych; chodzi o to, aby nie wpaść w pułapkę występujących cykli.

Dla omówionych schematów losowania Ripley (1981) wyznaczył wariancję różnicy $\overline{Y} - \hat{Y}(\Delta)$, którą będziemy nazywać błędem estymacji. Dla próby prostej otrzymał on następującą zależność:

---

* Zbiory Δ_i, Δ_j nazywamy przystającymi, jeżeli istnieje translacja, która przeprowadza jeden zbiór w drugi.

$$\operatorname{var}_p\left\{\overline{Y}-\hat{Y}(\Delta)\right\}=\frac{1}{n}\left[\sigma^2-E\big(R(x,y)\big)\right] \tag{10}$$

gdzie $\sigma^2=\operatorname{var}\{Y(x)\}$, $R(x,y)$ jest funkcją korelacyjną pola losowego $Y(x)$, natomiast x, y są dowolnymi punktami obszaru Δ.

Wobec właściwości funkcji korelacyjnej drugi składnik wzoru (10) (przy dostatecznej odległości pomiędzy punktami x, y) zmierza do zera i może zostać pominięty.

Dla próby warstwowej wariancja różnicy $\overline{Y}-\hat{Y}(\Delta)$, przy założeniu, że warstwy są zbiorami pokrywającymi się, jest wyrażona wzorem:

$$\operatorname{var}_w\left\{\overline{Y}-\hat{Y}(\Delta)\right\}=\frac{1}{n}\left[\sigma^2-E\big(R(x_j,y_j)\big)\right] \tag{11}$$

gdzie x_j, y_j są dowolnymi punktami j-tej warstwy.

Ponieważ funkcja korelacyjna jest funkcją monotoniczną malejącą, więc przy odpowiednim rozwarstwieniu obszaru Δ będzie zachodziła nierówność:

$$E\big(R(x_j,y_j)\big)\geq E\big(R(x,y)\big)$$

co wobec (10) i (11) oznacza, że:

$$\operatorname{var}_p\left\{\overline{Y}-\hat{Y}(\Delta)\right\}\geq\operatorname{var}_w\left\{\overline{Y}-\hat{Y}(\Delta)\right\}$$

Losowanie warstwowe może być więc skuteczniejsze od losowania prostego. Ideą losowania warstwowego jest taki wybór warstw, aby $E\big(R(x_j,y_j)\big)$ było maksymalnie duże, co w konsekwencji oznacza, że próba warstwowa powinna być efektywniejsza od próby prostej.

Losowanie proste jest niejednokrotnie mało efektywne, dlatego powinno być stosowane jedynie w tych przypadkach, gdy jest to uzasadnione. Decyzja o wyborze próby prostej nieograniczonej jest zasadna tylko wówczas, gdy dysponujemy próbą o niewielkiej liczebności lub gdy posiadane informacje dodatkowe o badanym polu losowym nie wskazują na inny alternatywny schemat losowania.

Dla izotropowego pola losowego zachodzi następująca równość:

$$E(R(x,y)) = \int_0^\infty R(r) b_\Delta(r) dr$$

gdzie $b_\Delta(r)$ jest funkcją wagową, która zależy od kształtu obszaru Δ; można udowodnić, że przyjmuje ona największe wartości dla obszarów zwartych, takich jak kula czy kwadrat, natomiast r oznacza odległość pomiędzy punktami x i y.

Wstawiając poprzednią zależność do wzoru (11) i zakładając, że wszystkie warstwy są przystające do pewnej kuli K, otrzymujemy*:

$$\mathrm{var}\left\{\overline{Y} - \hat{Y}(\Delta)\right\} = \frac{\sigma^2}{n}\left[1 - \int_0^\infty R(r) b_K(r) dr\right] \tag{12}$$

Z równości (12) wynika, że efektywność warstwowania zależy od funkcji korelacji oraz że powinna ona przyjmować możliwie duże wartości, co oznacza, że:

– warstwy powinny być możliwie małe,
– warstwy powinny być jednorodne ze względu na cechę reprezentowaną przez badane pole.

W celu przeprowadzenia dyskusji dotyczącej efektywności losowania systematycznego skorzystamy z wyniku otrzymanego przez Zubrzyckiego (1957). Wykazał on, że dla próby systematycznej prawdziwa jest równość:

$$\mathrm{var}_s\left\{\overline{Y} - \hat{Y}(\Delta)\right\} = \frac{\sigma^2}{n}\sum_{i,j} R(x_i, x_j) - \mathrm{var}\left(\hat{Y}(\Delta)\right) \tag{13}$$

gdzie x_i jest środkiem ciężkości zbioru Δ_i.

Analizując wzór (13) zauważamy, że wariancja różnicy $\overline{Y} - \hat{Y}(\Delta)$ może być dostatecznie mała przy odpowiednim podziale obszaru Δ.

Ponieważ losowanie systematyczne jest szczególnym przypadkiem losowania warstwowego $(n_j = 1, j = 1,...,k)$, więc wcześniejsze uwagi uwzględnione przy próbie warstwowej przenoszą się na próbę systematyczną.

Poza tym Zubrzycki udowodnił, że jeżeli dla dowolnych $i, j \in \{1,...,n\}$ jest spełniony warunek:

* Dowód równości (12) można znaleźć w pracy Ripley'a (1981).

$$R(x_i, y_j) \leq \frac{1}{|\Delta_1|^2} \int_{\Delta_i} \int_{\Delta_j} R(x, y)\,dxdy$$

to:

$$\mathrm{var}_s \left\{ \overline{Y} - \hat{Y}(\Delta) \right\} \leq \mathrm{var}_w \left\{ \overline{Y} - \hat{Y}(\Delta) \right\}$$

Ostatni warunek wskazuje na to, iż losowanie systematyczne może być efektywniejsze od losowania warstwowego. Podsumowując należy stwierdzić, że:

– próba warstwowa może być efektywniejsza od próby prostej,
– próba systematyczna może być efektywniejsza od próby warstwowej.

Zakończenie

Przeprowadzona dyskusja dotycząca efektywności poszczególnych schematów losowania próby dyskretnej pola losowego została ograniczona do wybranego parametru, co nie powinno wpływać na ogólność następującego wniosku: efektywność estymatora parametru pola losowego jest funkcją schematu losowania próby.

Sformułowany wniosek upoważnia do postawienia zadania polegającego na wyznaczeniu optymalnej próby służącej do oszacowania wybranego parametru pola losowego w określonej klasie estymatorów.

Zadanie wyboru optymalnej próby pola losowego, według autorów opracowania, jest zadaniem otwartym. Zaprezentowane wyniki mają jedynie charakter ogólny i mogą stanowić podstawę do rozwiązania postawionego zadania.

Literatura

Chow G.C.: *Ekonometria.* PWN, Warszawa 1995.

Cressie N.A.C.: *Statistics for Spatial Data.* John Wiley, New York 1993.

Gajowski J., Przybycin Z.: *Własności estymatorów wybranej klasy modeli przestrzennych.* Zeszyty Naukowe Uniwersytetu Opolskiego, Opole 2002.

Lieonienko H.H., Iwanow A.W.: *Staticzjeskij analiz słuczajnych poliej.* Wyższa Szkoła, Kijów 1980.

Matheror G.: *La Theorie des Variables Regionaliseer et ses Applications.* Masson, Paris 1965.

Przybycin Z.: *Estymatory f współczynników regresji liniowej stacjonarnego pola losowego.* Zeszyty Naukowe AE, Katowice 1995.

Ripley B.D.: *Spatial Statistics.* John Wiley, New York 1981.

Zubrzycki S.: *O szacowaniu parametrów złóż geologicznych.* „Zastosowania Matematyki" 1957.

THE PROPERTIES OF THE ESTIMATORS OF THE FUNCTIONAL CHARACTERISTICS OF A RANDOM FIELD

Summary

The paper deals with estimator properties of the functional characteristics of the random fields for a continuous and discrete test. Due to the fact that in empirical examination we generally deal with the discrete test, the basic schemes of a discrete test sampling have been taken into account in the aspect of their effectiveness.