

WYKORZYSTANIE ZANURZANIA OBSERWACJI W PRÓBIE DO KONSTRUKCJI KART KONTROLNYCH

Małgorzata Kobylińska
Katedra Metod Ilościowych
Uniwersytet Warmińsko-Mazurski w Olsztynie
angosiak@poczta.onet.pl

Streszczenie: Sterowanie procesem produkcyjnym wymaga monitorowania obserwowanych odchyłeń zmiennych diagnostycznych od wartości celowych. Podstawową rolę w działaniach sterowania jakością odgrywiają karty kontrolne, które są jednym z podstawowych narzędzi statystycznej kontroli jakości. Należą do bieżących metod kontroli, jakości, a jednocześnie przy właściwym stosowaniu mają duże znaczenie przy poprawie jakości produkcji. W 1975 roku zostało wprowadzone, przez Tukey'a pojęcie zanurzania obserwacji w próbie wielowymiarowej. Miara zanurzania może być wykorzystywana do klasyfikacji obserwacji wielowymiarowych względem ich oddalenia od centralnego skupienia oraz do wykrywania i eliminacji obserwacji nietypowych. W pracy zaprezentowano wykorzystanie miary zanurzania obserwacji w próbie do konstrukcji kart kontrolnych.

Słowa kluczowe: statystyczne sterowanie jakością procesów, karty kontrolne, zanurzania obserwacji w próbie

WPROWADZENIE

Nowe technologie umożliwiają wytwarzanie produktów charakteryzujących się coraz wyższym poziomem użyteczności. Ważną rolę w funkcjonowaniu nowoczesnego przedsiębiorstwa odgrywa sprawnie działający system zarządzania jakością usług i wyrobów. Przedsiębiorstwo produkcyjne dąży do tego, żeby jak najlepszej jakości były produkty, które wytwarza. Jakość produktu decyduje o jego przydatności dla klienta oraz satysfakcji klienta z produktu, który został przez niego nabyty. Dobra jakość wytwarzanych przez przedsiębiorstwo produktów lub świadczonych przez nie usług pozwala na utrzymanie się na rynku, na którym konkurencja jest bardzo duża. Na proces produkcyjny oddziałują pewne czynniki,

które są przyczyną jego zmienności. Na podstawie kontroli przebiegu charakterystyk wartości zaobserwowanych parametrów danego procesu można dokonać analizy jego zmienności.

Chcąc skutecznie zarządzać jakością przedsiębiorstwo powinno wyznaczyć stan ilościowy jakości produktu. W tym przypadku konieczne jest ustalenie zbioru cech (właściwości), według których będzie przeprowadzana ocena produktu lub usługi. Właściwości te mogą być wyrażone za pomocą wartości liczbowych, lub mogą odpowiadać im pewne kategorie obrazujące dany stan rzeczy [Hamrol 2008].

Sterowanie jakością procesów produkcyjnych obejmuje między innymi: opracowanie planów jakości procesów i wyrobów, zarządzanie konfiguracją i ryzykiem w procesach produkcyjnych, nadzorowanie jakości oraz ocenę zdolności procesów produkcyjnych [Szkoda 2004].

Ważną rolę w sterowaniu jakością procesów produkcyjnych odgrywają metody statystyczne, wśród nich metody statystyki opisowej i matematycznej. Statystyczne sterowanie procesem produkcyjnym (SPC) ma na celu poprawę jego jakości poprzez realizowaną w rzeczywistym czasie kontrolę jego przebiegu, w celu wykrycia ewentualnych rozregulowań. Zmienność procesu może być wywołana przez dwa rodzaje przyczyn: losowe lub specjalne. Pierwsze z nich są stale obecne w procesie, nieuniknione w danych warunkach oraz nie są łatwo rozpoznawalne, kolejne – pojawiają się nieregularnie i trudno przewidzieć ich skutki, mogą być stałym elementem procesu lub pojawiać się w nim przypadkowo [Wiśniewska, Malinowska 2011].

Celem statystycznego testowania procesem produkcyjnym jest doprowadzenie go do stabilnego poziomu, utrzymanie go na tym poziomie oraz zapewnienie spełnienia założonych wymagań przez wyroby i usługi. Proces produkcyjny nazywamy stabilnym, jeżeli czynniki wpływające na jego przebieg mają charakter wyłącznie losowy [Sałaciński 2009].

KARTY KONTROLNE JEDNOWYMIAROWE I WIELOWYMIAROWE

Karty kontrolne zostały wprowadzone przez W. A. Shewharta w 1924 roku i stały się niezwykle użyteczne zarówno w produkcji jak i usługach. Ich wykorzystanie do monitorowania procesów produkcyjnych umożliwia znaczną redukcję zmienności oraz prowadzi do wzrostu jakości produkcji. Zadaniem kart kontrolnych jest dostarczenie informacji, wykorzystując metody graficzne, o stabilności procesu, czyli czy wybrane do kontroli danych cech parametry statystyczne „nie zmieniają się” w czasie [Hamrol 2008].

W kartach kontrolnych Shewharta wykorzystywane są prawa rachunku prawdopodobieństwa i statystyki matematycznej. Jeżeli jakiegokolwiek sygnały świadczą o rozregulowaniu procesu produkcji, można podjąć decyzje o jego przerwaniu i skorygowaniu jego przebiegu.

Karty kontrolne konstruowane dla zmiennych ilościowych mogą monitorować informacje o pojedynczych właściwościach procesu (karty kontrolne

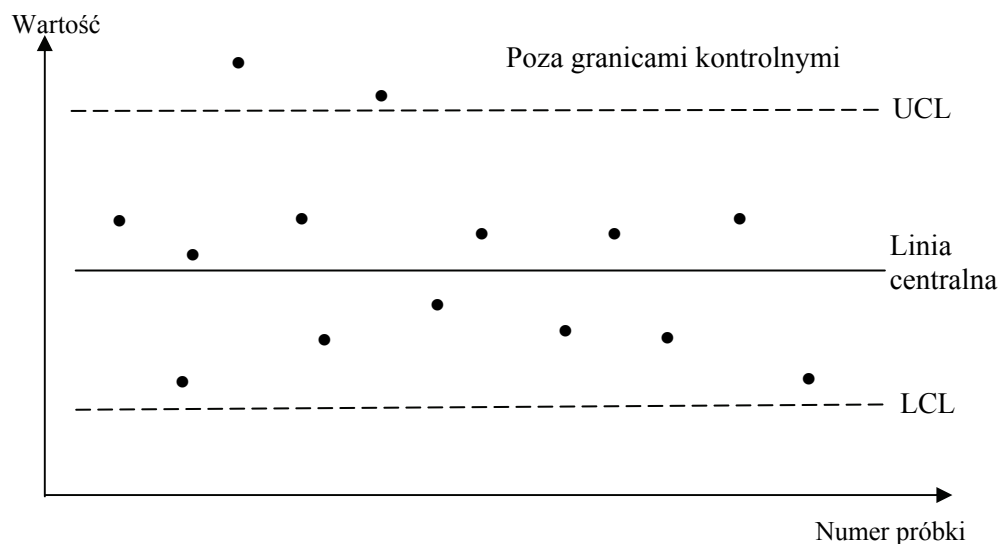
jednowymiarowe) lub wielu jego właściwościach (karty kontrolne wielowymiarowe). Umożliwiają monitorowanie zmian charakterystyk procesu dotyczących położenia, rozproszenia lub jego stabilności [Płaska 2000].

Klasyfikacji kart kontrolnych dokonać można ze względu na liczbę rozpatrywanych cech, rodzaj rozkładu analizowanej cechy lub jej charakter. Wśród klasycznych kart kontrolnych wyróżniamy karty kontrolne dla cech dyskretnych (zliczeniowych) oraz cech ciągłych (mierzalnych). Karty sekwencyjne (karta MA, EWMA, IX-MR, CUSUM) wykorzystywane są do wykrywania pojawiających się trendów oraz niewielkich przesunięć wartości średniej procesu produkcyjnego. Szczegółowo typy kart kontrolnych zostały omówione w pracy Sałacińskiego [Sałaciński 2009].

Karta kontrolna jako narzędzie wspomagające regulację procesów produkcji jest wykresem przedstawiającym statystyki np. wartość średnią, odchylenie standardowe w czasie lub według numeru kolejnej próbki pobranej do badania. Głównym elementem każdej karty jest diagram kontrolny, na którym nanoszone są:

- Linia centralna (CL) – jest zwykle średnią z rozważanych danych procesu,
- Linie kontrolne - określane jako górna granica kontrolna (UCL) oraz dolna granica kontrolna (LCL) (w zależności od potrzeb badania stosuje się jednostronny lub dwustronny obszar krytyczny). W wielu przypadkach wyznacza się także granice ostrzegawcze.

Rysunek 1. Karta kontrolna



Źródło: Aczel [2000]

Linie kontrolne wyznaczają zakres zmienności monitorowanej statystyki procesu, tzn. zakres, której przyczyną są tylko czynniki losowe. Jeżeli badana

statystyka przekracza granice kontrolne lub gdy jej wykres wykazuje pewne stałe zmiany, proces może być niestabilizowany i powinien być skorygowany [Aczel 2000].

W rzeczywistości często mamy do czynienia z procesami produkcji, do opisu których wykorzystuje się kilka właściwości, które mogą być od siebie niezależne lub powiązane ze sobą liniowo. W przypadku monitorowania dwóch właściwości procesu wykorzystać można elipsy kontrolne. Jednak na ich wykresie nie ma możliwości zaobserwowania trendów, wzorców, tendencji procesu oraz jego zmienności w czasie.

W celu monitorowania w czasie parametrów statystycznych wybranych do kontroli wielu cech procesu można wykorzystać kilka kart kontrolnych. Niesie to za sobą kilka wad np. konieczność konstrukcji i analizowania znacznej liczby kart kontrolnych lub pomijanie zależności pomiędzy cechami. Ocena procesu ze względu na każdą z właściwości oddzielnie może być myląca, w związku z wzajemnymi powiązania między nimi [Demski 2010].

Pomocnym narzędzie w celu monitorowania procesów wielowymiarowych jest karta T^2 Hotellinga, na której zamiast zaobserwowanych w czasie wartości przedstawiana jest przekształcona wartość, tzw. odległość T^2 Holettinga. Jest ona wyznaczona, jako różnica pomiędzy średnią wartością w próbce i średnią w populacji.

$$T^2(p) = n[(\bar{x}(p) - \bar{\bar{x}})^T s^{-1}(\bar{x}(p) - \bar{\bar{x}})] \quad (1)$$

gdzie

n - liczba jednostek w próbce,

$\bar{x}(p)$ - wektor średnich w próbce,

$\bar{\bar{x}}$ - wektor średnich uregulowanego procesu,

s^{-1} - macierz odwrotna do macierzy kowariancji s [patrz np. Dembski 2010].

Wartości tej odległości są nanoszone na karcie kontrolnej dla poszczególnych grup pomiarowych odnotowanych na osi poziomej karty. Należy zaznaczyć, że na kartach kontrolnych Hotellinga dolna granica kontrolna nie występuje. W przypadku zaobserwowania rozregulowania procesu konieczne jest zidentyfikowanie cechy, która była tego przyczyną. W tym celu sporządza się wykresy jednowymiarowych kart kontrolnych np. $\bar{x} - R$. Jak już wcześniej wspomniano jest to działanie dość pracochłonne, szczególnie gdy badanych jest kilkanaście lub kilkadziesiąt właściwości procesu.

Użytecznym narzędziem kontrolowania procesu o wielu właściwościach może być analiza głównych składowych. Umożliwia ona ze zbioru wszystkich zmiennych charakteryzujących dany proces uzyskanie zestawu nowych zmiennych, nazywanych składowymi głównymi. Stosunkowo niewielka liczba składowych głównych przenosi większość informacji zawartych w dużej liczbie zmiennych

wejściowych. W przypadku nadzorowania procesu z wykorzystaniem tej metody w pierwszym kroku przeprowadza się analizę składowych głównych, a w kolejnym konstruuje się karty kontrolne Hotellinga dla wyodrębnionych najważniejszych składowych.

W większości wykorzystywanych kart kontrolnych zakłada się, że zebrane wyniki pomiarów mają rozkład normalny lub zbliżony do normalnego. Założenie to jest wymagane ze względu na postać wzorów, zgodnie z którymi wyznacza się granice kontrolne. W przypadku niespełnienia założenia dotyczącego zgodności z rozkładem normalnym na karcie mogą pojawiać się fałszywe sygnały o rozregulowaniu procesu produkcji lub nie będą odnotowane wartości świadczące o jego niestabilności [Greber 2002].

Jednym z głównych założeń w metodach konstrukcji kart kontrolnych Hotellinga oraz tych, w których wykorzystuje się składowe główne jest, żeby zależność pomiędzy właściwościami procesu była liniowa [Demski 2010].

ZANURZANIE OBSERWACJI W PRÓBIE W KARTACH KONTROLNYCH

Dokonując analizy statystycznej zjawisk empirycznych często napotyka się na utrudnienia wynikające z niespełnienia założeń stawianych wykorzystywanym metodom statystycznym. Wiele z tych metod wymaga spełnienia założeń dotyczących zgodności z rozkładem normalnym lub co najmniej przybliżonego rozkładu normalnego. W związku z powyższym zachodzi konieczność stosowania metod statystycznych, które pomijają klasyczne założenia stochastyczne.

W 1975 roku zostało wprowadzone przez Tukey'a pojęcie zanurzania obserwacji w zbiorze danych. Pojęcie to było rozwijane przez wielu badaczy z punktu widzenia jego przydatności w statystycznej analizie danych, szczególnie w analizie danych z obserwacjami nietypowymi. W literaturze przedmiotu szeroko omówione zostały różne kryteria wyznaczania miary zanurzenia obserwacji w próbie m. in. kryterium Mahalanobisa, kryterium zanurzania sympleksowego Liu lub kryterium Barnett'a [Liu i in. 1999, Rousseeuw i Ruts 1996].

Niech $P_n^p = \{x_1, x_2, \dots, x_n\}$ będzie układem wektorów wyrażających próbę p -wymiarową o liczebności n pochodzącą z pewnego p -wymiarowego rozkładu określonego dystrybuantą F_p oraz niech $\theta \in R^p$ będzie pewnym punktem z przestrzeni rzeczywistej R^p . Punkt θ może być punktem z próby P_n^p . Koncepcja zanurzania obserwacji w próbie Tukey'a oparta jest na następującej definicji.

Definicja 1. Zanurzeniem punktu θ w próbie P_n^p nazywamy funkcję próby $zan_p(\theta, P_n^p)$ o następującej postaci

$$zan_p(\theta, P_n^p) = \min_{\|u\|=1} \#\{i : u^T \theta \geq u^T x_i\}, \quad (2)$$

gdzie u jest wektorem kolumnowym z przestrzeni R^p o długości 1, czyli o normie euklidesowej $\|u\|=1$, natomiast $\#\{\cdot\}$ oznacza liczebność rozważanego zbioru.

W pracy do konstrukcji kart kontrolnych wykorzystane zostało kryterium zanurzania Mahalanobisa. Jest ono określone według następującej definicji.

Definicja 2. Miarą zanurzania Mahalanobisa ($Mzan_p$) punktu θ w próbie P_n^p nazywamy funkcję

$$Mzan_p(\theta; P_n^p) = [1 + Q(\theta, P_n^p)]^{-1}, \quad (3)$$

gdzie $Q(\theta, P_n^p) = (\theta - \bar{x})^T S^{-1}(\theta - \bar{x})$ jest odległością Mahalanobisa

wektora θ od wektora średnich \bar{x} , przy czym $\theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \dots \\ \theta_p \end{bmatrix}$, $\bar{x} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \dots \\ \bar{x}_p \end{bmatrix}$,

$\bar{x} = \frac{1}{n} \sum_{j=1}^n x_j$, S jest macierzą kowariancji między rozważanymi

p wektorami, natomiast S^{-1} jest jej macierzą odwrotną.

W sensie matematycznym miara zanurzania $Mzan_p$ wyznaczona na podstawie wzoru (3) bazuje na odległości między punktem przestrzeni p wymiarowej i wektorem średnich \bar{x} , według metryki wyznaczonej odwrotnością macierzy S . Własności miary zanurzania $Mzan_p$ wynikają bezpośrednio z podstawowych własności odległości Mahalanobisa [patrz np. Mahalanobis 1936, Bartkowiak 1976].

Założmy, że proces produkcyjny charakteryzowany jest za pomocą p właściwości ($p > 1$). Niech G będzie pewnym p -wymiarowym rozkładem teoretycznym określonym dystrybuantą F_G oraz niech y_1, y_2, \dots, y_m będzie układem p -wymiarowych wektorów pochodzących z rozkładu G . $P_{m(G)}^p = \{y_1, y_2, \dots, y_m\}$ o liczebności m jest próbą odniesienia

w kontekście monitorowanego procesu produkcji. Niech $P_{n(F)}^p = \{x_1, x_2, \dots, x_n\}$ będzie układem obserwowalnych wektorów procesu produkcji wyrażającym próbę p -wymiarową o liczebności n pochodzącą w pewnego p -wymiarowego rozkładu określonego dystrybuantą F_F . Zbadanie, czy proces produkcji uległ rozregulowaniu sprowadza się do sprawdzenia, czy obserwacje próby $P_{n(F)}^p$ należą do rozkładu G. Karty kontrolne, które są wykorzystywane w tym celu oparte są na mierze zanurzania obserwacji w próbie. Miary zanurzania wyznaczone dla obserwacji próby umożliwiają ich uporządkowanie względem oddalenia od centralnego skupienia, które jest wyznaczone przez obserwacje z najwyższymi wartościami miary zanurzania. Punkty, którym odpowiadają najniższe wartości tej miary leżą na zewnątrz „chmury danych” i mogą być traktowane jako odstające, czyli nie należące do danego rozkładu.

Procedura sprawdzania czy dany proces uległ rozregulowaniu sprowadza się do konstrukcji kart kontrolnych z wykorzystaniem miary zanurzania obserwacji oraz próbie [Liu 1995]. W tym celu wyznaczane są miary zanurzania obserwacji prób $P_{m(G)}^p$ i $P_{n(F)}^p$. Karta kontrolna oparta na miarach zanurzania jest wykresem wartości $r_{G_m}(x_i)$ w czasie z wartością $CL=0,5$ i linią kontrolną α . Jeżeli wartości $r_{G_m}(x_i)$ są mniejsze od zadanej wartości α , to wystąpiło rozregulowanie procesu. Wartości statystyki $r_{G_m}(x_i)$ wyznaczamy jako

$$r_{G_m}(x_i) = \# \{y_j / \text{zan}_G(y_j) \leq \text{zan}_G(x_i), i = 1, 2, \dots, n, j = 1, 2, \dots, m\} / m \quad (4)$$

gdzie $\#\{\cdot\}$ oznacza liczebność rozważanego zbioru, $\text{zan}_G(\cdot)$ jest wartością miary zanurzania obserwacji w próbie wyznaczoną na podstawie wzoru 3.

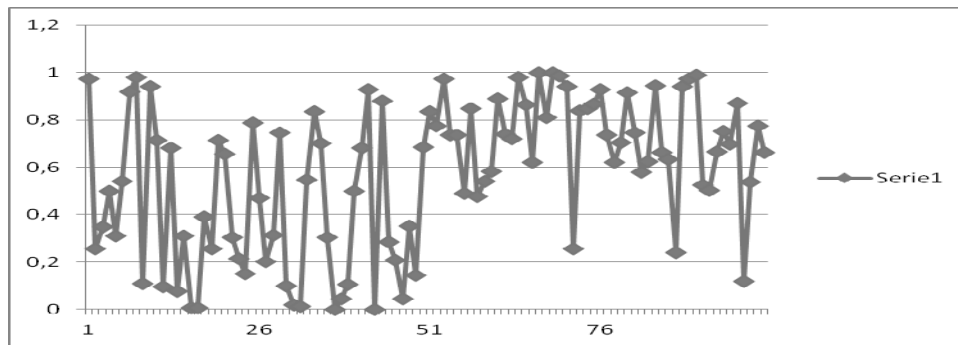
W celu zilustrowania konstrukcji kart kontrolnych posłużono się próbami dwuwymiarowymi, przyjmując za każdym razem inną konfigurację badanych prób:

- Próba PD1 pochodzi ze standardowego dwuwymiarowego rozkładu normalnego,
- Próba PD2 – pierwszych 250 obserwacji pochodzi ze standardowego dwuwymiarowego rozkładu normalnego, kolejnych 100 z dwuwymiarowego rozkładu normalnego z wektorem wartości oczekiwanych $[3,3]$ i macierzą kowariancji $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$,
- Próba PD8 – pierwszych 50 obserwacji pochodzi ze standardowego dwuwymiarowego rozkładu normalnego, kolejnych 50 pochodzi

z dwuwymiarowego rozkładu normalnego z zerowym wektorem wartości oczekiwanych oraz macierzą kowariancji $\begin{bmatrix} 0,5 & 0 \\ 0 & 0,5 \end{bmatrix}$.

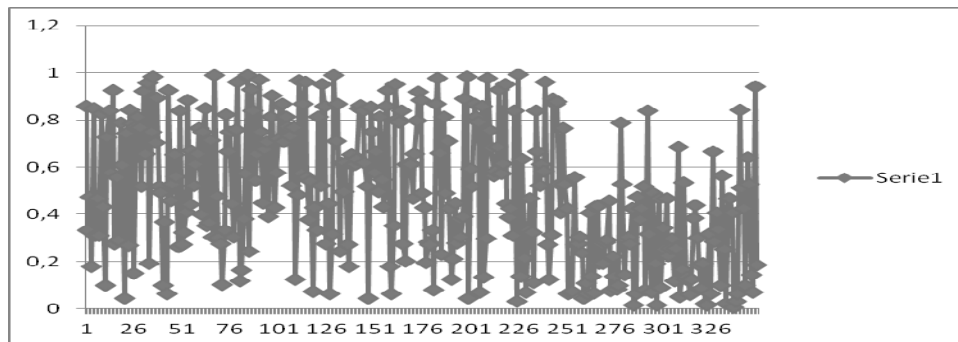
PD1 jest próbą odniesienia w kontekście monitorowanego procesu.

Rysunek 1. Karta kontrolna PD8



Źródło: opracowanie własne

Rysunek 2. Karta kontrolna PD2



Źródło: opracowanie własne

Na rysunkach przedstawione zostały procesy, w których widoczne są rozregulowania. Na pierwszym nich (rys.1) od 51 próbki wzrosła wartość statystyki $r_{G_m}(x_i)$. Wysoka jej wartość świadczy o tym, że większej części obserwacji próby PD1 odpowiadają niższe wartości miary zanurzenia w porównaniu z obserwacjami próby PD8. Oznacza to, że w próbie PD8 występuje większe skupienie obserwacji wokół wektora wartości średnich w porównaniu z próbą PD1. Na rysunku 2 zauważyć można, że od 251 próbki zmalały wartości statystyki $r_{G_m}(x_i)$. Może to być uwarunkowane przesunięciem rozkładu lub zwiększeniem jego zróżnicowania.

PODSUMOWANIE

W pracy zaprezentowano wykorzystanie miary zanurzania obserwacji w próbie do konstrukcji kart kontrolnych. Klasyczne wielowymiarowe karty kontrolne oparte na statystyce χ^2 lub na odległości T^2 Hotellinga zakładają żeby dane pochodzące z pomiarów dokonywanych przy sterowaniu procesem miały rozkład zgodny z rozkładem normalnym. Założenie to jest również wymagane przy konstrukcji elipsy kontrolnej dla przypadku dwuwymiarowego.

Konstrukcja kart kontrolnych opartych na miarach zanurzania obserwacji w próbie dla przypadku wielowymiarowego nie wymaga złożonych przekształceń danych oraz zaawansowanych metod analizy statystycznej. Nie jest wymagane również spełnienia założeń dotyczących zgodności z rozkładem normalnym.

W literaturze dotyczącej zarządzania jakością procesów produkcji szeroko omówione zostały różne rodzaje kart kontrolnych. Ich wybór zależy od specyfiki procesu, który ma być monitorowany. W pracy omówiona została karta kontrolna oparta na statystyce r , która podobnie jak karta \bar{x} analizuje zmiany miar położenia oraz zmienności. W literaturze przedmiotu omówione zostały inne karty kontrolne, w których wykorzystuje się zanurzanie obserwacji w próbie [Liu 1995, Liu i in. 2004]. Będą one przedmiotem dalszych rozważań.

BIBLIOGRAFIA

- Aczel A.D (2000) Statystyka w zarządzaniu, PWN, Warszawa.
- Bartkowiak A., (1976), An Abbreviated Method of Calculating the Mahalanobis Distance or Residua Sum of Squares in a Linear Regression Model, *Applicationes Mathematicae*, 15: 215-222.
- Demski T. (2010) Sterowanie jakością procesu o wielu właściwościach: wielowymiarowe karty kontrolne i inne narzędzia, www.statsoft.pl/czytelnia.html
- Greber T. (2002) Karty kontrolne przy ocenie liczbowej sterowania procesami- założenia i analiza, www.statsoft.pl/czytelnia.html
- Hamrol A. (2008) Zarządzanie jakością z przykładami, PWN, Warszawa.
- Liu R.Y, Parelius J.M., Singh K. (1999) Multivariate Analysis by Data Depth: Descriptive Statistics, Graphics and Inference, *The Annals of Statistics*, 27, 783-858.
- Liu R. (1995), Control Chart for Multivariate Processes, *Journal of the American Statistical Association* December, Vol. 90, No. 432.
- Liu, R.Y., Singh, K. and Teng, J.H. (2004) DDMA-charts: nonparametric multivariate moving average controlcharts based on data depth. *Allgemeines Statistisches Archiv*, 88, 235-258.
- Mahalanobis P. C., (1936), On the Generalized Distance in Statistics, *Proc. Nat. Inst. Sci. India*, 2, 49-55.
- Płaska S. (2000) Wprowadzenie do statystycznego sterowania procesami technologicznymi, Wydawnictwo Politechniki Lubelskiej, Lublin.
- Malinowska M., Wiśniewska E. (2011) Zarządzanie jakością żywności. Systemy, koncepcje, instrumenty, Difin.

- Rousseeuw P.J. , Ruts I. (1996), Bivariate Location Depth, *Applied Statistics* , 45, 516-526.
- Salaciński T. (2009) SPC-statystyczne sterowanie procesami produkcji, Oficyna Wydawnicza Politechniki Warszawskiej.
- Szkoda J. (2004) Sterowanie jakością procesów produkcyjnych. Teoria i praktyka, Wydawnictwo Uniwersytetu Warmińsko-Mazurskiego w Olsztynie.
- Tukey J.W., (1975) Mathematics and the Picturing of Data, *Proceedings of the International Congress of Mathematicians*, 523-531.

USING OBSERVATION DEPTH IN A SAMPLE TO CONSTRUCT CONTROL CARDS

Abstract: Controlling a production process requires monitoring of the observed deviations of diagnostic variables from target values. The fundamental role in quality control is played by control cards, which are one of the main tools of statistical quality control. They are one of ongoing quality control methods and – when used properly – are very important in improving production quality. In 1975, Tukey introduced the concept of observation depth in a multi-dimensional sample. The depth measure can be used to classify multi-dimensional observations in regard to the distance from the central cluster and to detect and eliminate atypical observations. This article presents the application of a measure of observation depth in a sample in the construction of control cards.

Keywords: observation depth measure in a sample, control cards