Recenzja rozprawy doktorskiej mgr Kamila Dyby

pt. "Zastosowanie wielowymiarowych funkcji kwantylowych do badania wielowymiarowych porzadków stochastycznych"

Wiesław Dziubdziela

W recenzowanej rozprawie Autor podejmuje próbę wykorzystania uogólnionych funkcji kwantylowych zdefiniowanych w 1992 roku przez Einmahla i Masona w pracy [1] do opisu wielowymiarowych porządków stochastycznych. W tym celu definiuje tzw. "wielowymiarowe kwantyle oparte o parę wektorów losowych".

Einmahl i Mason wprowadzili uogólniony kwantyl w następujący sposób. Oznaczmy przez $\mathbb{B}$ zbiory borelowskie przestrzeni $\mathbb{R}^d$. Niech P będzie miarą na $\left(\mathbb{R}^d, \mathbb{B}\right)$. Niech λ będzie funkcją rzeczywistą zbioru określoną na podzbiorze $\mathcal{C} \subset \mathbb{B}$. Uogólnioną funkcję kwantylową definiujemy przez

$$U\left(t; P, \mathcal{C}\right) = \inf\left\{\lambda\left(A\right) : \ P\left(A\right) \geq t, A \in \mathcal{C}\right\}, \quad 0 < t < 1.$$

Dla dowolnego rozkładu P na $\mathbb{R}^d$, funkcją głębokości (depth lub depth function) jest każde nieujemne rzeczywiste przekształcenie $D(\mathbf{x}, P)$, $\mathbf{x} \in \mathbb{R}^d$, porządkujące obserwacje $\mathbf{x}$ w zależnosci od odległości od centrum wielowymiarowego rozkładu P (zob. np. [4],[6]).

Serfling ([5],[6]), specyfikując rodzinę zbiorów $\mathcal{C}$ w definicji Einmahla i Masona, definiuje tzw. uogólnioną funkcję kwantylową opartą na funkcji głąbi (generalized depth-based quantile function) $D(\mathbf{x}, P)$ przyjmując, dla wszystkich $t \in (0,1)$,

$$U\left(t; P, \mathcal{C}_D\right) = \inf\left\{\lambda\left(A\right) : \ P\left(A\right) \geq t, A \in \mathcal{C}_D\right\}, \quad 0 < t < 1,$$

gdzie

$$\mathcal{C}_D = \left\{I\left(t, D, P\right) : 0 < t < \sup_{\mathbf{x}\in\mathbb{R}^d} D\left(\mathbf{x}, P\right)\right\}$$

i

$$I(t, D, P) = \left\{ \mathbf{x} \in \mathbb{R}^d : D(\mathbf{x}, P) \geq t \right\}.$$

Istnieje znaczna liczba różnych podejśc do definicji wielowymiarowych kwantyli, uważa się jednak ([2]), że podejście oparte o funkcje głębokości jest najważniejsze.

W rozprawie doktorskiej przedstawiona jest następująca konstrukcja wielowymiarowego "kwantyla". Przedstawimy tu jedną z sześciu definicji wprowadzanych w podobny sposób.

Niech $\mathbf{X}$ i $\mathbf{Y}$ będą wektorami losowymi o wartościach w $\mathbb{R}^d$. Rodzinę podzbiorów $\mathcal{C}$ określa się wzorem

$$\mathcal{C} = \left\{ [-\infty, \mathbf{x}] : \mathbf{x} \in \bar{\mathbb{R}}^d = [-\infty, +\infty]^d \right\}.$$

Funkcję zbioru λ określamy jako dystrybuantę wektora $\mathbf{Y}$, czyli $\lambda([-\infty, \mathbf{x}]) = G(\mathbf{x})$. Rozkład P jest rozkładem wektora $\mathbf{X}$, zatem $P(A) = P([-\infty, \mathbf{x}]) = F(\mathbf{x})$.

Otrzymujemy "kwantyl", oznaczony przez przez Autora $GF^{-1}(t)$ zdefiniowany wzorem

$$GF^{-1}(t) = \inf \left\{ G(\mathbf{x}) : F(\mathbf{x}) \geq t, \mathbf{x} \in \bar{\mathbb{R}}^d \right\}.$$

Na stronie 37 Rozprawy można znaleźć tą definicję i pięć pozostałych . Autor nazywa je "wielowymiarowymi funkcjami kwantylowymi".

Wydaje się jednak, że fukcja kwantylowa powinna być związana z pojedyńczym wektorem losowym a nie parą wektorów. Na stronie 43 Rozprawy Autor pisze o "funkcjach kwantylowych wyznaczonych w oparciu o parę wektorów losowych", takie stwierdzenie wydaje się dziwne.

Ograniczmy uwagę do funkcji $GF^{-1}(t), t \in (0,1)$. Przyjmijmy dla prostoty, ze rozważane dystrybuanty są ciagłe. W przypadku $d = 1$, mamy

$$GF^{-1}(t) = \inf \left\{ G(x) : F(x) \geq t, x \in \bar{\mathbb{R}} \right\} = G\left(F^{-1}(x)\right).$$

Zatem funkcja $GF^{-1}(t) = G \circ F^{-1}(t)$ jest złożeniem funkcji G i F^{-1}. Złożenie to występuje w charakteryzacjach porzadków stochastycznych w jednym wymiarze (zob. str. 21-22 Rozprawy).

Na przykład, warunek $Y \leq_{st} X$ zachodzi wtedy i tylko wtedy, gdy $G \circ F^{-1}(t) \geq t$ dla każdego $t \in (0,1)$.

Dla $d \geq 1$, Autor w Twierdzeniu 3.1 charakteryzuje porzadek dolny ortantowy. Pokazuje On, że warunek $Y \leq_{lo} X$ zachodzi wtedy i tylko wtedy, gdy $GF^{-1}(t) \geq t$ dla każdego $t \in (0,1)$.

Dla $d = 1$ złożenie $GF^{-1}(t) = G \circ F^{-1}(t)$ nosi nazwę *krzywej porządkowej dominacji (ordinal dominance curve)* (zob. str. 139. recenzowanej Rozprawy) i można przyjąć, że dla $d > 1$, funkcja $GF^{-1}(t)$ jest jej rozszeszeniem na wiecej wymiarów.

Na podstawie powyższych uwag recenzent uważa, że tytuł rozprawy: "*Zastosowanie wielowymiarowych funkcji kwantylowych do badania wielowymiarowych porządków stochastycznych*", nie odpowiada jej zawartości. Autor nie bada wielowymiarowych kwantyli, ale raczej pewne odpowiedniki złożeń wielowymiarowych dystrybuant. Jednakże, bez względu na to, jak nazywa się rozważane pojęcia, Autor otrzymuje nowe wyniki charakteryzujące wielowymiarowe stochastycze porządki.

Rozprawa doktorska składa się z wstępu, sześciu rozdziałów i bibliografii, liczącej 41 pozycji.

Rozdział 1. *Kwantyle i porządki stochastyczne w jednym i wielu wymiarach.*

W rozdziale tym przedstawiono i usystematyzowano dobrze znane wyniki dotyczące funkcji kwantylowych w jednym wymiarze i wybranych jednowymiarowych porządków stochastycznych (porządek stochastyczny, odwrotny porządek hazardowy, porządek hazardowy i porządek ilorazu wiarogodności). Omówiono także wielowymiarowe porządki takie jak: porządek dolny ortandowy, porządek górny ortandowy, porządek stochastyczny słaby odwrotny porządek hazardowy, słaby porządek hazardowy i porządek ilorazu wiarygodności. Ważnym paragrafem rozdziału jest paragraf poświecony koncepcji *wielowymiarowych funkcji kwantylowych.* We wcześniejszym fragmencie recenzji koncepcja ta została przedstawiona i skrytykowana.

Rozdział 2. *Własności wielowymiarowych funkcji kwantylowych.*

Autor przygotowuje tu aparat potrzebny mu w dalszej części pracy. Badane są własności sześciu funkcji zdefiniowanych w Rozdziale 1, które Autor nazywa wielowymiarowymi funkcjami kwantylowymi. Własnościami tymi są monotoniczność, ciągłość, odwracalność, granice w $\mathbf{0}$ i $\mathbf{1}$. W ostatnim obszernym paragrafie 2.4 Rozdziału 2 przedstawione są własności funkcji kwantylowych rozkładów zniekształconych.

Rozdział zawiera wyniki własne Autora.

Na stronie 43 znajduje sie sformułowanie Autora: "*...funkcjami kwantylowymi wyznaczonymi w oparciu parę wektorów losowych...*", które podkreśla, według recenzenta, nienaturalność przyjętej prze Autora definicji wielowymiarowych funkcji kwantylowych.

Rozdział 3. *Wielowymiarowe funkcje kwantylowe a porządki stochastyczne.*

Badane są związki pomiędzy wielowymiarowymi funkcjami kwantylowymi a porządkami stochastycznymi. Rozważone są porządki ortantowe, porządek stochastyczny, słabe porządki hazardowe i porządek ilorazu wiarogodności.

Rozdział zawiera wyniki własne Autora.

Rozdział 4. *Uporządkowanie rozkładów zadanych przez archimedesowe funkcje łącznikowe.*

W rozdziale rozważane są wektory losowe o archimedesowych funkcjach łącznikowych (kopulach). W tym przypadku funkcje kwantylowe można zapisać w jawnej postaci. Korzystając z tego faktu, scharakteryzowano uporządkowanie archimedesowych funkcjach łącznikowych. Podano ilustrujące przykłady.

Rozdział zawiera wyniki własne Autora.

Rozdział 5. *Porządkowanie dwóch par rozkładów i metryki niezmiennicze względem wielowymiarowych porządków stochastycznych.*

Autor podejmuje próbę uogólnienia wybranych wyników dla $d = 1$ z pracy Lehmanna i Rojo [3] na przypadek dowolnego $d > 1$. Są to uogólnienia stwierdzenia : "*rozkład zmiennej losowej* Y_2 *jest bardziej na prawo od* X_2 *niż* Y_1 *od* X_1" oraz pojęcia metryk niezmienniczych względem porządków stochastycznych.

Rozdział zawiera wyniki własne Autora.

Rozdział 6. *Empiryczne wielowymiarowe funkcje kwantylowe.*

W rozdziele tym zdefiniowano tzw. *empiryczne wielowymiarowe funkcje kwantylowe.* Zbadano ich monotoniczność i jednostronną ciągłość. Omówiono praktyczne (jawne) wyznaczanie empirycznycy funkcji kwantylowych. Rozdział kończy paragraf zatytułowany "*Twierdzenia graniczne.*". Zawiera on interesujace wyniki asymptotyczne dotyczące jednostajnej zbieżności z prawdopodobieństwem 1.

Rozdział zawiera wyniki własne Autora.

Recenzent wyraża rozczarowanie, że Autor nie publikował żadnych wyników otrzymanych w pracy i przez to nie podał ich ocenie przez specjalistów. W pracy można znaleźć fragmenty, które po odpowiednim przeredagowaniu można publikować, np. z Rozdziału 6, paragraf " *Twierdzenia graniczne.*".

Otrzymane wyniki nie są, w ocenie recenzenta, dobrze osadzone w literaturze przedmiotu. Brak jest wyraźnych odniesień do znanych już rezultatów. Autor nie uwypukla co jest jego osobistym wkładem.

Praca jest pisana w sposób przejrzysty i staranny. Przedstawiane dowody są kompletne. Można na ich postawie wnioskować o dobrej znajomosci przez Autora rozważanej teorii i stosowanych w niej metod.

Podsumowując, mimo wskazanych wyżej niedociągnięć i przedstawionych krytyczych uwag, recenzowana Rozprawa Doktorska zawiera na tyle interesujący i oryginalny materiał, że recenzent wnioskuje o dopuszczenie Pana magistra Kamila Dybę do dalszych etapów przewodu doktorskiego.

Literatura

[1] Einmahl, J.H.J., Mason, D.M. (1992). Generalized quantile processes, *Ann. Statist.*, **20**, 1062-1078.

[2] Hamel, A.H., Kostner, D. (2018). Cone distribution functions and quantiles for multivariate random variables, *J. Multivariate Anal.*, **167**, 97-113.

[3] Einmahl, J.H.J., Mason, D.M. (1992). Generalized quantile processes, *Ann. Statist.*, **20**, 1062-1078.

[4] Lehmann, E.L., Rojo, J. (1992). Invariant directional orderings, *Ann. Statist.*, **20**, 2100-2110.

[5] Serfling, R. (2002). Quantile functions for multivariate analysis: approaches and applications, *Statist. Neerlandica*, **56**, 214-232.

[6] Serfling, R. (2002). Generalized quantile processes based on multivariate depth functions, with applications in nonparametric multivariate analysis, *J. Multivariate Anal.*, **83**, 232-247.

[7] Serfling, R., Zuo, Y. (2010). Discussion, *Ann. Statist.*, **38**, 676-684.

25.04.2019 Wiesław Dziubdziel