

Tomasz Rychlik
Instytut Matematyczny PAN
Warszawa

Warszawa, 13 czerwca 2023 r.

Recenzja poprawionej wersji rozprawy doktorskiej
pani mgr Krystyny Maciąg
pt. *Charakteryzacje rozkładów prawdopodobieństwa przez regresje
niekolejnych uogólnionych statystyk porządkowych
oraz dyskretnych wartości rekordowych*

Rozprawa pani mgr Krystyny Maciąg dotyczy charakteryzacji rozkładu brzegowego wektora (lub ciągu) niezależnych zmiennych losowych o jednakowym rozkładzie na podstawie pojedynczych regresji funkcji niekolejnych elementów wybranych uporządkowanych wektorów funkcji opartych na zmiennych bazowych, w sytuacji, gdy elementy te tworzą łańcuch Markowa. Dokładniej, rozważane są następujące modele. W pierwszym jest to wektor uogólnionych statystyk pozycyjnych $X_*^{(1)}, \dots, X_*^{(r)}, \dots, X_*^{(r+\ell)}$ o zadanych parametrach $\gamma_1, \dots, \gamma_r, \dots, \gamma_{r+\ell} > 0$ i charakteryzowanej ciągłej dystrybucji bazowej F . Zadanie polega na jednoznacznym wyznaczeniu F w sytuacji, gdy znana jest pojedyncza zależność regresyjna $\xi(x) = \mathbb{E}(h(X_*^{(r+\ell)}) | X_*^{(r)} = x)$ dla zadanych ściśle rosnących funkcji ciągłych ξ i h na przedziałowym nośniku (α, β) dystrybuanty F oraz liczb naturalnych $r \geq 1$ i $\ell \geq 2$. W drugim analogiczne zadanie rozważano dla regresji $\xi(x) = \mathbb{E}(h(W_{r+\ell}) | W_r = x)$ w ciągu słabych rekordów $W_1, \dots, W_r, \dots, W_{r+\ell}$ zdefiniowanych na bazie dyskretnych zmiennych losowych o wartościach całkowitych nieujemnych i skończonym bądź nieskończonym nośniku. Rozważana jest też modyfikacja tego problemu polegająca na charakteryzacji ogona rozkładu bazowego na podstawie regresji $\xi(x) = \mathbb{E}(h(R_{r+\ell}) | R_r = x)$, gdy $R_1, \dots, R_r, \dots, R_{r+\ell}$ są klasycznymi rekordami w ciągu zmiennych losowych o wartościach całkowitych nieujemnych.

Dla uogólnionych statystyk pozycyjnych zadanie to było rozwiązane w przypadku $\ell = 1$ (p. pozycja [19] bibliografii rozprawy), a dla niekolejnych $\ell \geq 2$ tylko przy ograniczeniu do liniowych funkcji ξ i h (p. również [16]). Ponadto Bieniek [12] pokazał, że pełną charakteryzację rozkładu można uzyskać przy znajomości drugiej funkcji regresji $\xi_1(x) = \mathbb{E}(h(X_*^{(r+\ell)}) | X_*^{(r+1)} = x)$. W przypadku słabych rekordów znane są wyniki charakteryzacyjne dla $\ell = 1$ (p. [54]) oraz dla $\ell = 2$, gdy

h jest identycznością, a nośnik obserwacji jest nieskończony (por. [3]). Charakteryzacje ogonów rozkładów dyskretnych za pomocą regresji kolejnych zwykłych rekordów były rozważane w pracach [40] i [53]. Żadna z metod dowodowych w tych pracach nie da się rozszerzyć na przypadek większych odstępów ℓ między uporządkowanymi obserwacjami i ogólnych funkcji regresji. Niniejsza rozprawa doktorska stanowi próbę takiego rozszerzenia.

Rozprawa liczy 104 strony, składa się ze wstępu z krótkim omówieniem jej zawartości, 4 rozdziałów, podsumowania wskazującego najistotniejsze osiągnięcia i problemy nie rozwiązane, oraz bibliografii liczącej 57 pozycji literaturowych cytowanych w pracy. Rozdział 1 zawiera wiadomości wstępne wykorzystane dalej w dowodach głównych wyników. Są tam opisane modele ciągów uporządkowanych zmiennych losowych analizowanych w pracy, własności Markowa tych ciągów, oraz podstawowe własności funkcji regresji w rozważanych modelach. Szczególnie użyteczne są tu rekurencyjna zależność kolejnych funkcji regresji opisana w Lemacie 1.1, uporządkowanie $\xi > h$ oraz zbieżność ich wartości do wspólnej granicy na prawym końcu nośnika β (p. Lematy 1.3 i 1.5). Rozdział ten zawiera też krótki opis dotychczas uzyskanych wyników w dziedzinie charakteryzacji rozkładów prawdopodobieństwa przez regresje uporządkowanych zmiennych losowych.

Kolejne trzy rozdziały obejmują zasadnicze wyniki rozprawy. Dotyczą one odpowiednio regresji dla uogólnionych statystyk pozycyjnych z bazową ciągłą gęstością i dystrybuantą oraz dla dyskretnych wartości rekordowych. Wyniki te są one bardziej skomplikowane i mniej spektakularne niż w innych klasycznych problemach charakteryzacji. Np., główny wynik Rozdziału 3 i chyba też całej rozprawy (p. Twierdzenie 3.1) brzmi następująco.: Jeśli dla ustalonych parametrów $0 < \gamma_{r+1} \leq \dots \leq \gamma_{r+\ell}$ (konkretne wartości $\gamma_1, \dots, \gamma_r$ są tu nieistotne), $\ell \geq 2$, i rosnących funkcji ciągłych $\xi < h$ na ustalonym przedziale (α, β) układ równań całkowych

$$y_i(x) - y_i(\alpha) = \frac{\gamma_{r+i+1}}{\gamma_{r+i}} \int_{\alpha}^x \frac{y_i(t) - y_{i+1}(t)}{\xi(t) - y_1(t)} d\xi(t), \quad i = 1, \dots, \ell - 1,$$

ma jedyne rozwiązanie $(\varphi_i, \dots, \varphi_{\ell-1})$ spełniające warunki

$$h(x) = \varphi_{\ell}(x) < \varphi_{\ell-1}(x) < \dots < \varphi_1(x) < \xi(x), \quad \alpha < x < \beta,$$

$$\int_{\alpha}^x \frac{d\xi(t)}{\xi(t) - \varphi_1(t)} \begin{cases} < \infty, & \alpha < x < \beta, \\ = \infty, & x = \beta, \end{cases} \quad (1)$$

oraz

$$\lim_{x \nearrow \beta} \xi(x) \exp \left(- \int_{\alpha}^x \frac{d\xi(t)}{\xi(t) - \varphi_1(t)} \right) = 0, \quad (2)$$

to równanie regresji

$$\xi(x) = \mathbb{E} \left(h(X_*^{(r+\ell)}) | X_*^{(r)} = x \right), \quad \alpha < x < \beta, \quad (3)$$

jednoznacznie charakteryzuje dystrybuantę bazową ciągu uogólnionych statystyk pozycyjnych $X_*^1, \dots, X_*^{(r+\ell)}$, która jest postaci

$$F(x) = 1 - \exp \left(- \int_{\alpha}^x \frac{1}{\xi(t) - h(t)} d \left(\frac{\xi}{\gamma_{r+1}} + \sum_{i=1}^{\ell-1} \frac{\varphi_i}{\gamma_{r+i+1}} \right) (t) \right), \quad \alpha < x < \beta. \quad (4)$$

Metody wyznaczania tak postawionych problemów istnienia i jednoznaczności rozwiązań układów równań całkowych (czy nawet ich uproszczonych wersji różniczkowych w Twierdzeniu 2.2) nie są znane w literaturze. Autorce udało się jedynie uzyskać istnienie jednoznacznych rozwiązań, gdy odstęp $\ell = 2$, a funkcja h jest ograniczona lub gdy ξ jest asymptotycznie równoważna pewnej liniowej transformacji funkcji h .

Przedstawione w pracy przykłady wskazują na istotną praktyczną ograniczoność przedstawianych tam metod charakteryzacyjnych. Np., dla danego rozkładu liczy się często skomplikowaną funkcję regresji i potem wykazuje się, że ona charakteryzuje jednoznacznie rozkład (p. Przykłady 2.1, 3.3) lub dla danego rozkładu F dowodzi się, że funkcja regresji charakteryzuje jednoznacznie ten rozkład, ale nie podaje się analitycznej postaci tej regresji (p. Przykład 2.2).

W poprawionej wersji pracy Twierdzenie 3.1 zawiera dwa dodatkowe warunki opisujące charakteryzowaną dystrybuantę: górną część wzoru (1) oraz wzór (2). Jak pokazuje dodany Lemat 3.5, są one wymagane, aby funkcja określona wzorem (4) była w istocie właściwą dystrybuantą miary probabilistycznej o nośniku zawartym w przedziale (α, β) . Analogiczne uproszczone warunki pojawiły się w Lematach 2.1, 2.2, oraz Twierdzeniach 2.1, 2.2 charakteryzujących rozkłady o ciągłych gęstościach na podstawie regresji (3).

W przypadku charakteryzacji wynikających z regresji słabych wartości rekordowych

$$\xi(x) = \mathbb{E} (h(W_{r+\ell}) | W_r = x), \quad \alpha < x < \beta,$$

z nieskończonym nośnikiem uzupełnienia wymagał jedynie warunek

$$\lim_{n \rightarrow \infty} \xi(n) \prod_{j=0}^{n-1} \frac{\xi(j) - \varphi_1(j)}{\xi(j+1) - \varphi_1(j)} = 0.$$

Omówione wyżej poprawki są kluczowe dla poprawności głównych tez rozprawy.

Inne jej obszary zostały również opisane z większą skrupulatnością. Przede wszystkim, precyzyjnie został przeanalizowany przypadek, w którym funkcja regresji w lewym końcu nośnika może zbiegać do minus nieskończoności, który w pierwotnej wersji potraktowano dość zdawkowo. Starannie wykluczono możliwość regresji liniowej dla rozkładów startujących z minus nieskończoności. Doprecyzowano argumentację w dowodzie Twierdzenia 3.3 opisującego regresję z odstępem dwa dla funkcji regresji asymptotycznie liniowej. Poprawiono zakres argumentów w przypadku zwykłych rekordów w ciągach dyskretnych.

Aktualna wersja pracy nie zawiera moim zdaniem istotnych luk, a prezentowane w niej wyniki są poprawnie dowiedzione. Jej temat jest trudny i nie zawiera wielu spektakularnych finalnych rezultatów. Stanowi jednak dobry punkt wyjścia do dalszych badań w tym kierunku. Uważam, że jest to dobra rozprawa doktorska i wnioskuję o dopuszczenie jej Autorki do dalszych etapów postępowania doktorskiego.

Tomasz Rychlik

Tomasz Rychlik