UNIWERSYTET WROCŁAWSKI

Prof. Ryszard Szekli

Instytut Matematyczny
Pl. Grunwaldzki 2/4
50-384 Wrocław
szekli@math.uni.wroc.pl

Recenzja pracy doktorskiej magistra Pawła Skolińskiego pt. *Estymacja parametrów niejednorodnych procesów gamma.*

Pan mgr Paweł Skoliński jest studentem studiów doktoranckich w dyscyplinie matematyka na Wydziale Matematyki Politechniki Wrocławskiej. Zgodnie z MathSciNet jest on autorem jednej publikacji w J. Stat. Theory Pract. [21], wspólnej z promotorką jego rozprawy profesor Politechniki Wrocławskiej Alicją Jokiel-Rokitą. Rozprawa doktorska magistra Skolińskiego powstała w oparciu o wspomnianą publikację i składa się z 6 rozdziałów opisanych na 112 stronach tekstu w języku polskim, zakończonych listą 49 referencji.

Tematyka pracy dotyczy estymacji parametrów pewnej klasy procesów punktowych na R_+, o punktach $0 = T_0 < T_1 < T_2 \ldots$, dla których stochastyczna intensywność względem naturalnej filtracji jest dana prostym wzorem (stosując oznaczenia z rozprawy):

$$g(t) = z(\Lambda(t) - \Lambda(T_{N(t-)}))\lambda(t),$$

gdzie $\Lambda(t) = \int_0^\infty \lambda(u)du$, $\lambda : R_+ \to R_+ : \int_0^\infty \lambda(u)du = \infty$ oraz $z(t)$ jest funkcją hazardową rozkładu $Gamma(\kappa, 1)$ (κ to parametr kształtu, parametr skali równy 1), $N(t) = \sum_{n=0}^\infty I_{\{T_n \leq t\}}$.
Tego typu procesy mają łączny rozkład zmiennych $T_1, \ldots, T_n$ o gęstości

$$f_n(t_1, \ldots, t_n) = \prod_{i=1}^n \lambda(t_i)[\Lambda(t_i) - \Lambda(t_{i-1})]^{\kappa-1} \exp[-\Lambda(t_n)]/\Gamma(\kappa)^n,$$

co daje możliwość estymacji parametrycznej modelu jeśli za funkcję $\lambda(t)$ przyjmiemy konkretną funkcję zależną od pewnych parametrów. W rozprawie rozpatruje się 2 przypadki:

$$\lambda(t) = \rho e^{\beta t}, \ \rho > 0, \ \beta \in R,$$

oraz

$$\lambda(t) = \alpha\beta t^{\beta-1}, \ \alpha > 0, \beta > 0,$$

nazywając te modele $MGP(\rho, \beta, \kappa)$ oraz $MPLP(\alpha, \beta, \kappa)$, odpowiednio.

W szerszej perspektywie, badane przypadki wpisują się w tematykę statystycznych modeli opartych o procesy punktowe, opisanych obszernie w książce Andersen et al. [2], lecz bezpośrednia motywacja badań ujętych w rozprawie doktorskiej nie polegała na badaniu teorii z punktu widzenia procesów punktowych, lecz pochodzi od pracy Bermana [5] stanowiąc kontynuację badań z prac [19] i [20] poświęconych zagadnieniom estymacji parametrów modeli metodą największej wiarogodności.

W przypadku, gdy $\kappa = 1$, to badany proces punktowy jest niejednorodnym procesem Poissona. Gdy $\lambda(t) = 1$, to proces punktowy jest procesem odnowy o odstępach o rozkładzie *Gamma*. Te dwa przypadki interpretowane są w literaturze jako modele minimalnej naprawy (minimal repair) i perfekcyjnej naprawy (perfect repair), odpowiednio. Badane modele przy dowolnych parametrach, stanowią propozycję odzwierciedlania tak zwanej częściowej naprawy (imperfect repair).

Najważniejsze rezultaty rozprawy zawarte są w rozdziale 2 (Tw. 2.2.2 - 2.2.5) i dotyczą własności asymptotycznych estymatorów największej wiarygodności w modelu MGP, w szczególności podana jest wielowymiarowe centralne twierdzenie graniczne dla odpowiednio standaryzowanych estymatorów największej wiarygodności modelu $MGP(\rho, \beta, \kappa)$ z jawną postacią macierzy kowariancji, która jest osobliwa. Rezultaty te oparte są o wyniki pracy [21] wspomnianej we wstępie recenzji.

Przejdę teraz do bardziej detalicznego opisu rozprawy. Rozdział 1 stanowi wprowadzenie do rozprawy i w dużej części obejmuje on fakty dotyczące teorii procesów punktowych na R_+. Na końcu tego rozdziału podane są wzory na łączne rozkłady zmiennych $T_1, \ldots, T_n$, z wyodrębnieniem przypadków $MPLP$ oraz MGP. Zarys teorii procesów punktowych przedstawiony w tym rozdziale, przy użyciu Definicji 1.1.8 i argumenty użyte w Twierdzeniu 1.2.3 mają wartość intuicyjną, lecz klasyczna, formalna teoria stochastycznych intensywności jest bardziej subtelna, bazując na teorii martyngałów, czasach zatrzymania, przewidywalnych projekcjach, tak jak przedstawiono ją na przykład w rozdziale II. 4.1 cytowanej książki [2]. Część dotycząca ogólnej teorii procesów punktowych mogłaby być pominięta, bo do dalszych rozważań w rozprawie potrzebne są jedynie łączne rozkłady $T_1, \ldots, T_n$, które są znane z literatury (np. [5]) i łatwe do otrzymania bez użycia teorii procesów punktowych.

Rozdział 2, w swej pierwszej części, zawiera wyliczenia estymatorów parametrów modelu $MPLP$ metodą największej wiarygodności poprzez wyznaczenie

$$\tilde{\Theta}_n = arg \max_{\Theta} \log L_n(\Theta; \mathbf{t})$$

gdzie $\Theta = (\alpha, \beta, \kappa)$, $\alpha > 0, \beta > 0$
oraz $L_n(\Theta; \mathbf{t})$ jest funkcją wiarygodności modelu $MPLP(\alpha, \beta, \kappa)$ przy $\mathbf{t} = (t_1, \ldots, t_n)$. Istnienie dopuszczalnych rozwiązań zależące od realizacji $\mathbf{t}$ jest opisane jako Fakt 2.1.1 oraz Lemat 2.1.1 , które są oparte o wyniki pracy [20].

Własności asymptotyczne estymatorów największej wiarygodności parametrów procesu $MPLP$ zostały udowodnione w pracy [4] i są przedstawione w rozprawie w Twierdzeniach 2.1.1. i 2.1.2. W szczególności Twierdzenie 2.1.2 zawiera wielowymiarowe centralne twierdzenie graniczne dla odpowiednio standaryzowanych estymatorów $(\tilde{\alpha}, \tilde{\beta}, \tilde{\kappa})$ parametrów (α, β, κ) przy realizacjach $(t_1, \ldots, t_n)$ w modelu $MPLP$. Postać (osobliwej) macierzy kowariancji w tym twierdzeniu jest jawna. Opis estymatorów w modelu $MPLP$, gdy $\kappa = 1$ oraz odpowiednie własności asymptotyczne zawarte są w Fakcie 2.1.2 i Twierdzeniu 2.1.3, wziętych z prac [36] i [39]. Inny szczególny przypadek MPLP, gdy $\beta = 1$ opisany jest w Fakcie 2.1.3 i Twierdzeniu 2.1.4, w oparciu o [6]. Część pierwsza rozdziału 2, która jest opracowaniem wyników teoretycznych z literatury, dowodzi posiadania ogólnej wiedzy teoretycznej doktoranta w zakresie zagadnień estymacji parametrycznej procesów punktowych tworzonych z procesów odnowy poprzez pewne przekształcenia osi czasowej (w przypadku, gdy odstępy między punktami w procesie odnowy mają rozkłady gamma, takie procesy autor nazywa niejednorodnymi procesami gamma).

W drugiej części rozdziału 2 znajdują się analogiczne badania dla modelu MGP. Postać estymatorów oraz pewne warunki istnienia rozwiązań podane są w Fakcie 2.2.1 i Twierdzeniu 2.2.1 w oparciu o wyniki z pracy [19]. Własności asymptotyczne estymatorów największej wiarygodności modelu MGP opisane w rozprawie opierają się o wspomnianą we wstępie recenzji pracę [21], wspólną z promotorką rozprawy doktorskiej. Są to nowe wyniki, które wymagały niestandardowych metod ze względu na osobliwość macierzy kowariancji pojawiających się w badanym modelu. Interesujące jest w szczególności Twierdzenie 2.2.4, które zawiera wielowymiarowe centralne twierdzenie graniczne dla odpowiednio standaryzowanych estymatorów $(\tilde{\rho}_n, \tilde{\beta}_n, \tilde{\kappa}_n)$ parametrów (ρ, β, κ) w modelu $MGP(\rho, \beta, \kappa)$. Postać osobliwej macierzy kowariancji jest jawna, i co ciekawe, ma podobną strukturę do macierzy kowariancji w analogicznym twierdzeniu dla modelu $MPLP$. Osobno badany jest przypadek $\kappa = 1$ w Twierdzeniu 2.2.5, które jest udowodnione poprzez adaptację metody dowodu z przypadku ogólnego. Twierdzenia 2.2.2 -2.2.4 są, z dokładnością do zmiany kilku oznaczeń, wzięte wprost z pracy [21]. Dowody tych twierdzeń umieszczone w rozdziale 6 stanowią przetłumaczenie dowodów z języka angielskiego z pracy [21], z wyjątkiem Twierdzenia 2.2.4, gdzie autor rozprawy dowód tego twierdzenia rozpisał dokładniej, przy okazji poprawiając błędne $C^*(\Theta_0)$ z pracy [21], tam w dowodzie Twierdzenia 3, na $C^*_n(\Theta_0)$ w dowodzie Tw. 2.2.4 swojej rozprawy. Niestety, w dowodzie swojego Twierdzenia 2.2.5 na stronie 107 rozprawy znajduje się podobny błąd, gdzie powinno być $d^*_n(\beta_0)$ zamiast $d^*(\Theta_0)$. W Lemacie 6.2.2 w (ii) powinna być zbieżność według prawdopodobieństwa, zamiast zbieżności według rozkładu. Oceniając umiejętność samodzielnego prowadzenia pracy naukowej doktoranta należy zwrócić uwagę na fakt, że więszość oryginalnych wyników dotyczących rozwiązania problemu parametrycznej estymacji w badanej klasie modeli pochodzi ze wspólnej pracy. Zakładam, że wkład doktoranta w tej pracy był decydujący w dowodzeniu prezentowanych twierdzeń, a rola promotora polegała na postawieniu problemu i na współpracy przy przygotowaniu publikacji. Dowody twierdzeń oparte są o klasyczne centralne twierdzenia graniczne, ale sprawdzenie odpowiednich warunków, umiejętne grupowanie, oznaczanie i szacowanie zmiennych wymagało sprawności technicznej oraz znajomości wielu faktów.

Rozdział 3 jest poświęcony zbadaniu alternatywnych metod estymacji parametrów modeli $MPLP$ oraz MGP i przedstawione rozważania podane są dla kompletności prezentacji, jak stwierdza autor rozprawy już na wstępie tego rozdziału. Stwierdza on, że w przypadku rozważanych modeli te inne metody (metoda momentów i metoda najmniejszych kwadratów) nie prowadzą do rozsądnych rozwiązań. W rozdziale 3 nie ma nowych twierdzeń, jedynie opisy możliwych metod oraz kilka faktów z dowodami, Fakt 3.2.1 wzięty z pracy [20], Fakt 3.2.2, Fakt 3.3.1, które są prostymi obserwacjami, Fakt 3.4.1 oraz Twierdzenie 3.4.1 wzięte z [48]. Całość rozdziału ma charakter szerokiego komentarza co do stosowania innych metod estymacji niż metoda największej wiarygodności.

Rozdział 4 zawiera wyniki symulacji przeprowadzonych w celu porównania dokładności estymatorów badanych modeli przy użyciu oszacowań obciążenia (kryterium $\widehat{bias}$) i pierwiastka ze średniego błędu kwadratowego (kryterium $R\widehat{MS}E$), jak również w celu porównanie dokładności predykcji kolejnej chwili $n+1$ pierwszego punktu procesu punktowego,

$$\hat{t}_{n+1} = \hat{\Lambda}^{-1}(\hat{\kappa} + \hat{\Lambda}(t_n)),$$

ponownie w oparciu o wyliczenie oszacowań obciążenia i pierwiastka średniego błędu kwadratowego oraz w celu analizy istnienia estymatorów.

Wyniki symulacji przedstawione są obszernie przy pomocy zestawień tabelarycznych oraz wykresów. W podrozdziale 4.1 znajdują się opisy symulacji dla modelu $MPLP$, natomiast dla modelu MGP w 4.2. Całość rozdziału jest obszerna, a opisy są detaliczne powodując, że całość tego rozdziału ma charakter raportu technicznego. Trudno jest wyłowić z tego materiału jednoznaczne konkluzje. Cytując opis z rozprawy dowiadujemy się na przykład, że dla modelu $MPLP$: *W zdecydowanej większości przypadków estymator NW zapewnia dokładniejsze oszacowania parametru α pod względem $R\widehat{MS}E$ niż estymatory alternatywne*, czy: *Najdokładniejsze wartości predyktorów, względem przyjętych kryteriów, uzyskaliśmy przy zastosowaniu metody NW*, ale już dla modelu MGP :*Po przeanalizowaniu wyników symulacji komputerowych, nie możemy wskazać metody estymacji parametrów procesu MGP, która jest najlepsza, przy przyjętych kryteriach.* Należy podkreślić, że w wielu przypadkach estymatory największej wiarygodności nie istnieją.

Rozdział 5 zawiera ilustarcję faktu, iż doktorant potrafi zastosować wyniki teoretyczne do danych rzeczywistych wielokrotnie badanych już w literaturze. Autor rozprawy zaproponował własną metodę graficzną sprawdzania, czy dany model jest dobrze dopasowany do danych.

Rozprawa jest dobrze zredagowana z małą ilością usterek technicznych. Drobnym mankamentem dla czytelnika rozprawy może być niedokładność w trakcie odwoływania się do literatury, czy faktów, na przykład odwoływanie się do książki bez podania dokładniejszych wskazówek takich jak numer rozdziału, zakres stron, numer twierdzenia itp., czy powoływanie się ogólnie na przykład na twierdzenie Słuckiego, czy metodę delta bez podania wygodnych źródeł.

Sumując, jest to solidna praca zawierająca interesujące i nowe rezultaty, których uzyskanie wymagało dobrej znajomości teorii parametrycznej estymacji procesów punktowych oraz biegłości w symulacji badanych procesów. Wiedza z tych obszarów została wykorzystana w sposób twórczy, a autor wykazał się sprawnością techniczną.

Uważam, że rozprawa magistra Pawła Skolińskiego spełnia wszystkie wymagania stawiane rozprawom doktorskim w Ustawie z dnia 20 lipca 2018 r. *Prawo o szkolnictwie wyższym i nauce* (Art. 187. 1.). Wnoszę o dopuszczenie magistra Pawła Skolińskiego do dalszych etapów postępowania w sprawie nadania stopnia doktora.

Wydział Matematyki i Informatyki

Prof. dr hab. Ryszard Szekli

10 CZE. 2022