Dr. hab. Bartosz Kołodziejek
Wydział Matematyki i Nauk Informacyjnych
Politechnika Warszawska

Warszawa, 27.07.2022 r.

Recenzja pracy doktorskiej mgr. Pawła Skolińskiego pt. „Estymacja parametrów niejednorodnych procesów gamma"

Przedstawiona rozprawa doktorska pana Pawła Skolińskiego została wykonana na Wydziale Matematyki Politechniki Wrocławskiej, a promotorem była dr hab. Alicja Jokiel-Rokita. Rozprawa liczy 112 stron; składa się ze strony tytułowej, dedykacji, streszczenia, spisu treści, sześciu rozdziałow, spisów rysunków, tabel i literatury.

1 Zawartość rozprawy

Praca doktorska mgr. Pawła Skolińskiego zawiera matematyczną i numeryczną analizę wybranych modeli opisujących momenty występowania zdarzeń powracających. Głównym problemem rozważanym w pracy jest estymacja parametrów tych modeli oraz twierdzenia graniczne dla tych estymatorów.

Rozdzial pierwszy jest poświęcony wprowadzeniu podstawowych pojęć dotyczących procesów punktowych, czyli głównego obiektu wykorzystywanego do modelowania momentów występowania zdarzeń powracających. Zdefiniowano tutaj niejednorodny proces gamma (IGP) oraz jego szczególne przypadki, w szczególności proces MGP (czyli proces IGP z log-liniową funkcją trendu) oraz proces MPLP (czyli proces IGP z potęgową funkcją trendu).

W rozdziale drugim przedstawiono twierdzenia dotyczące estymacji metodą największej wiarogodności parametrów w obu szczególnych przypadkach procesu IGP. W podrozdziale 2.1 przedstawiono wyniki dotyczące stymacji parametrów procesu MPLP. W podrozdziale tym doktorant referuje wyniki z prac [20,4,39], ale nie tylko, ponieważ również uzupełnia pewne rozumowania - np. Lemat 2.1.1. Prezentowane twierdzenia graniczne są niestandardowe: z Twierdzenia 2.1.2 wynika, że żeby otrzymać nietrywialną granicę estymatorów największej wiarogodności, jeden z nich należy dodatkowo podzielić przez $\log n$, a odpowiadająca granica ma singularny rozkład normalny (macierz kowariancji jest osobliwa).

Wyniki dla modelu MPLP pokazują jakiego typu asymptotyki można się spodziewać również w przypadku procesu MGP o rosnącej funkcji trendu (parametr $\beta > 0$). Model ten był analizowany wcześniej tylko w pracy Jokiel-Rokity i Magiery [19]. W [19] przedstawiono jakościowe i ilościowe analizy dotyczące istnienia estymatora największej wiarogodności procesu MGP, które stanowią punkt wyjścia do analiz asymptotyki tych estymatorów. W podrozdziale 2.2.2. zaprezentowano wyniki z pracy [21], czyli wspólnej publikacji mgr. Skolińskiego i jego promotorki. Dowody twierdzeń zostały umieszczone w odległym rozdziale 6. Z jednej strony rozumiem z czego wynika taki układ, z drugiej, taki zabieg zaburza płynność czytania osoby zainteresowanej wkładem matematycznym. Twierdzenia dotyczą asymptotycznych własności estymatorów największej wiarogodności parametrów modelu. Wyniki te są analogiczne do wcześniej opisanego procesu MPLP, ale są od nich niezależne.

W rozdziale trzecim zaproponowano kilka alternatywnych metod estymacji procesu IGP: metoda momentów, metoda najmniejszych kwadratów, a przede wszystkim estymatorów bazujących na tzw. funkcji estymujących. Funkcja estymująca zależy m.in. od wyboru własności rozkładów momentów zdarzeń. Wykorzystując związki procesów IGP z rozkładami Beta i Gamma, zdefiniowano estymatory typu BD i typu GD parametrów modeli MPLP i MGP.

Dla procesu MPLP podejście typu BD było już rozważane wcześniej w [20]. Należy tu podkreślić, że wybór postaci funkcji estymującej jest dość arbitralny i opiera się heurystycznych przesłankach.

Rozdział czwarty zawiera wyniki symulacji numerycznych, których celem było porównanie zachowania różnych estymatorów parametrów procesów MPLP i MGP wprowadzonych wcześniej. Wyniki zaprezentowano w formie tabel i wykresów. Sformułowano również rekomendacje dotyczące stosowalności tych estymatorów w różnych scenariuszach. Estymatory typu BD i GD są konkurencyjne do estymatorów największej wiarogodności.

Rozdział 5 poświęcony jest estymacji parametrów modeli w oparciu o rzeczywiste zbiory danych: awarii generatora lotniczego oraz awarii silnika napędowego. Z estymatorów tych można łatwo odczytać m.in. czy rozpatrywany system po naprawie jest w stanie lepszym czy gorszym niż tuż przed wystąpieniem awarii.

Rozdział 6 zawiera dowody twierdzeń sformułowanych w podrozdziale 2.2.2. Jest to główna część matematyczna pracy. Dowody pochodzą z pracy [21].

2 Ocena merytoryczna

Główny wkład matematyczny doktoranta stanowią twierdzenia z podrozdziału 2.2.2. Twierdzenia te pochodzą z artykułu mgr. Skolińskiego opublikowanego w *Journal of Statistical Theory and Practice* [21] wspólnie z promotorką, która jest ekspertem w dziedzinie. Nie znalazłem w internecie śladów innych prac doktoranta. Według mnie, jest to bardzo słaby dorobek naukowy.

Twierdzenia z rozdziału 2.2.2 nie są trywialne, a ich dowody często opierają się na wielu sprytnych szacowaniach dowodzonych w różnych lematach. Dotyczą one twierdzeń granicznych dość skomplikowanych funkcji od ciągu i.i.d. zmiennych losowych o rozkładzie Gamma. Doktorant musiał zidentyfikować główne czynniki odpowiadające za asymptotykę oraz udowodnić, że reszta nie kontrybuuję w sposób istotny w asymptotykę. Co do zasady, rachunki są przedstawiane dokładnie, jednak wzory bywają tak rozbudowane (jak np. wzór na c_{22} na stronie 97), że należy ważyć korzyści związane z przedstawieniem wszystkich przekształceń. Wydaje mi się, że doktorant dobrze sobie poradził w tym aspekcie. Pomimo że poszczególne elementy dowodu zwykle nie są wymagające, to połączenie ich w całość z pewnością wymagało bardzo dużego wkładu od doktoranta. W moim odczuciu doktorant wykazał się tutaj pewną dojrzałością matematyczną i dobrze poradził sobie w prezentacji dość złożonych dowodów.

Niemniej mam liczne zastrzeżenia do dowodów. O ile część redaktorską pracy uważam za wykonaną bardzo starannie, to część merytoryczna nie jest już tak dopracowana i tam gdzie pojawiają się wzory, często pojawiają się też błędy. Wszystkie zidentyfikowane przeze mnie błędy w dowodach są moim zdaniem naprawialne i nie wymuszają żadnych zmian w sformułowaniach twierdzeń. Niektóre z tych błędów w mojej ocenie są poważne. Listę moich zastrzerzeń różnego charakteru przedstawiam w następnym podrozdziale.

2.1 Uwagi krytyczne

Istotniejsze błędy oznaczam symbolem wykrzyknika: !.

2.1.1 Rozdział 1

Uwagi do dowodu Twierdzenia 1.2.3.:

1. Twierdzone jest, że „rozkład zmiennej losowej T_{i+1} zależy wyłącznie od momentu i-tego zdarzenia", co jest nieprawdą. Oczywiście chodzi o rozkład warunkowy i własność Markowa.

2. W dowodzie zostało ponadto wykorzystane oznaczenie $N(t+)$, które nie zostało wcześniej wprowadzone. Jak mniemam $N(t+) = N(t) + 1$.

3. ! Tożsamość (1.3) jest nieprawdziwa i nie wynika z poprzedniej linijki, a także nie implikuje kolejnej. Wydaje mi się, że zamiast (1.3) powinno być

$$g(t) = \lim_{\Delta t \to 0+} \frac{\mathbb{P}(W_{N(t+)} \in (Y(t), Y(t+\Delta t)] \mid T_{N(t-)})}{\mathbb{P}(W_{N(t+)} > Y(t) \mid T_{N(t-)})\Delta t}.$$

   Dalsze wnioskowanie powinno być uargumentowane mocną własnością Markowa. Dowód obecnie jest niepełny, chociaż dotyczy klasycznego zagadnienia.

   Co więcej, funkcja g jest granicą ciągu zmiennych losowych. Wypadałoby wskazać w jakim sensie jest rozumiana ta granica.

Uwagi do Twierdzenia 1.2.5:

1. W sformułowaniu powinno być $Be(\kappa i, \kappa(n-i))$, a nie $Be(i, n-i)$.

2. Na stronie 95 wykorzystywana jest mocniejsza własność precesu IGP, a mianowicie że zmienne $\Lambda(T_{i+1})/\Lambda(T_i)$, $i = 1, \ldots, n-1$, nie tylko mają rozkłady Beta, ale również są niezależne.

2.1.2 Rozdział 2.

Uwagi do Lematu 2.1.1.:

1. Strona 30, linijka 3: przy $A(t)$ stoi niepotrzebne „n".

2. Ostatnia równość w rozpisce $\kappa_n(\beta, t)$ powinna być nierównością.

2.1.3 Rozdział 6.1. Dowód twierdzenia 2.2.2

Uwagi do Lematu 6.1.1.:

1. Błąd we wzorze (6.6). Powinno być

$$Y_n = \frac{1}{n\kappa_0}\left(\sum_{i=1}^{n} X_i + \frac{\rho_0\kappa_0}{\beta_0}\right).$$

   Błąd ten propaguje się do wzoru z rozpiską $n^{1/2}(Y_n - 1)$ na środku strony 90.

2. Dowód punktu (ii) można istotnie uprościć: Punkt (ii) wynika z punktu (i) oraz dodatkowej obserwacji, że $\sqrt{n}(\bar{Y}_n - 1) \xrightarrow{p} 0$. Ta obserwacja może być prosto pokazana przez analizę wariancji tej zmiennej.

3. Błąd we wzorze (6.7). Powinno być

$$\sum_{i=1}^{n}\sum_{j=1}^{i} a_i b_j = \sum_{i=1}^{n}\sum_{j=i}^{n} b_i a_j.$$

4. Błąd w dwóch ostatnich wzorach na stronie 90. W przedostatnim wzorze w środkowym wyrażeniu powinna być $\sum_{j=1}^{i}$, a nie $\sum_{j=1}^{n}$. Prawa strona tego wzoru jest poprawna. W ostatnium wzorze drugi składnik po pierwszej i drugiej równosci powinien być $\frac{\rho_0}{\beta_0\kappa_0}$, a nie $\frac{\rho_0}{n\beta_0\kappa_0}$. Po drugiej równości powinno być $(n-i+1)^{-1}$, a nie $(n-i-1)^{-1}$.

5. Wzory (6.8) i (6.9) są błędne. Powinno być

$$K_n = n^{-1/2}\kappa_0^{-1}\sum_{i=1}^{n} X_i e_{in} + n^{-1/2}\frac{\rho}{\beta_0\kappa_0}\sum_{i=2}^{n} i^{-1}$$

oraz

$$e_{in} = \left(\sum_{j=i}^{n} j^{-1}\right) - 1 = \sum_{j=i}^{n}\left(j^{-1} - (n-i+1)^{-1}\right)$$

Błędy te propagują się do wzoru na K_{n2}.

6. Oznaczenie E_{ϑ_0} jest stosowane niekonsekwentnie - str 91 i 93.

7. Na stronie 92 wypadałoby wyjaśnić dlaczego $\frac{1}{n}\sum_{i=1}^{n}(1/Y_i - 1) = o_p(1)$. Własności ta jest potrzebna by $B_1 = o_P(1)$.

8. ! W 6 linijce strony 93 następująca nierówność została zastosowana:

$$\left(\frac{1}{Y_i} - 1\right)^2 \le \left(\frac{1}{\tilde{Y}_i} - 1\right)^2.$$

Nierówność ta nie jest w ogólności prawdziwa, co powoduje, że dalsze szacowania wyrażenia B_2 są nieprawidłowe. W dalszych szacowaniach również znajdują się błąd: wartość oczekiwana $\mathbb{E}[(1/\tilde{Y}_i - 1)^2]$ jest skończona tylko jeśli $i\,\kappa_0 > 2$.

9. Strona 93: Nie jest prawdą, że $\mathbb{E}[(Y_i - 1)^2] = \mathrm{Var}(Y_i)$, przez co ciąg równości na końcu dowodu Lematu 6.1.1. jest nieprawidłowy.

Uwagi do Lematu 6.1.2.:

1. Błąd we wzorze na U_{2n} na stronie 94: brakuje po prawej stronie składnika $\frac{1}{\sqrt{n}}\frac{\rho_0(\kappa_0-1)}{\beta_0}$.

2. Wypadałoby uzasadnić, że jeśli $X \sim \mathrm{Gamma}(\kappa, 1)$, to $\mathrm{Cov}(X, \log X) = 1$.

3. ! W dowodzie jest twierdzone, że zmienna losowa U_{1n} jest niezależna od wektora (U_{2n}, U_{3n}), co jest nieprawdą. Z Lematu 6.1.1 (iii) i (iv) wynika, że U_{1n} i U_{3n} są asymptotycznie nieskorelowane (a nie niezależne!) o ile reszta zbiega w L_2. Należy pokazać, że U_{1n} i U_{2n} również są asymptotycznie nieskorelowane.

Uwagi do Lematu 6.1.3.:

1. Błąd we wzore na B_{L1} na stronie 95. Suma powinna być od $i = 1$ lub inną wartość należy odejmować.

2. ! W pierwszej linijce ostatniego wzoru dowodu powinna być zmienna V_i, a nie Y_i. Co więcej, wariancja V_i^{-1} jest skończona tylko jeśli $\kappa_0(i-1) > 2$. Tym samym argument w ostatnim wzorze jest nieprawidłowy/niepełny.

Uwagi do Lematu 6.2.2.:

1. Funkcja ψ'' jest funkcją ujemną, więc we wzorze z linijki -8 na stronie 100 brakuje minusa. Co więcej funkcja $-\psi''$ jest malejąca, więc oszacowanie górne otrzymujemy w punkcie κ_0, a nie $\kappa_0 + hn^\delta$.

2. Na tronie 101 brakuje indeksu n przy $\mathbf{V}^*$ w czterech miejscach.

2.1.4 Inne

1. Twierdzenie 1.1.2. oraz Wniosek 1.2.1. zostały sformułowane bez dowodu ani referencji.

2. Ostatnia linijka na stronie 21: $+\Delta t$ powinna być w argumencie funkcji Λ.

3. W wielu miejscach doktorant miesza pojęcie funkcji f z wartością funkcji w punkcie $f(t)$. Prowadzi to do dziwnych oznaczeń procesów, np. $IGP(\lambda(t), \kappa)$, gdy nie wiadomo czym jest t. Błąd ten nie jest popełniany konsekwentnie, doktorant czasem pisze np. $TRP(\lambda(\cdot), \kappa)$.

4. Strona 47, drugi parametr rozkładu zmiennej U_i to $(n-i)\kappa$, a nie $(n-1)\kappa$. Ponadto, rozkład ten był wcześniej oznaczany przez Be, a nie B.

5. Bibliografia prowadzona niekonsekwentnie. Niektóre nazwy czasopism są w skróconej formie, a inne w pełnej. Ponadto znajduje się tam szereg literówek: referencje [6, 11, 17, 47].

3 Podsumowanie

Mimo stosunkowo licznych uwag krytycznych przedstawionych powyżej, pozytywnie oceniam przedstawioną mi do recenzji rozprawę doktorską mgr. Pawła Skolińskiego. Uważam, że główny cel pracy został osiągnięty, a doktorant włożył dużo wysiłku w przygotowanie pracy. Rozprawa doktorska Pana Pawła Skolińskiego spełnia moim zdaniem ustawowe i zwyczajowe wymagania i wnoszę o dopuszczenie jej do dalszych etapów przedowu doktorskiego.

Bartosz Kołodziejek