Biblioteka WIADOMOŚCI STATYSTYCZNYCH

STATYSTYKA MAŁYCH OBSZARÓW – SYMULACYJNE BADANIE PRÓBNIKA GIBBSA W HIERARCHICZNYM MODELU BAYESOWSKIM

Tom 62

WARSZAWA 2010

GŁÓWNY URZĄD STATYSTYCZNY

Wojciech Niemiro
Jacek Wesołowski

STATYSTYKA MAŁYCH OBSZARÓW – SYMULACYJNE BADANIE PRÓBNIKA GIBBSA W HIERARCHICZNYM MODELU BAYESOWSKIM

Tom 62

WARSZAWA 2010

Rada Programowa

dr Halina Dmochowska (przewodnicząca), prof. dr hab. Czesław Domański, mgr Małgorzata Fronk, prof. dr hab. Jan Kordos, dr Tomasz Pawlak, mgr Stanisława Szwałek, dr Teresa Śmiłowska, prof. dr hab. Kazimierz Zając

Prace merytoryczne

Departament Metodologii, Standardów i Rejestrów:
dr hab. Wojciech Niemiro, prof. dr hab. Jacek Wesołowski

Opracowanie redakcyjne

Anatol Kula, Alina Świderska

Projekt okładki

Krzysztof Dobrowolski

CENTRALNA BIBLIOTEKA STATYSTYCZNA WARSZAWA 3

183751

ZWS Druk i oprawa: Zakład Wydawnictw Statystycznych
al. Niepodległości 208, 00-925 Warszawa

Przedmowa

Przekazujemy Państwu sześćdziesiąty drugi tom serii: „Biblioteka Wiadomości Statystycznych", nawiązując do nurtu metodologicznego tej serii. Opracowanie dra hab. Wojciecha Niemiro i prof. dra hab. Jacka Wesołowskiego dotyczy statystyki małych obszarów.

Statystyka ta niesie ze sobą ogromny potencjał metodologiczny i aplikacyjny. Zainteresowanie tą dziedziną wzmogło się wyraźnie w XXI w. W życiu społeczno-gospodarczym Polski coraz większą rolę odgrywają lokalne jednostki terytorialne, które poszukują aktualnych, dokładnych i szczegółowych informacji. Należy podkreślić, że wraz ze statystyką małych obszarów w metodologii badań społeczno-ekonomicznych, prowadzonych na próbkach losowych, szersze zastosowanie znajduje podejście modelowe, wywodzące się wprost ze statystyki matematycznej.

Autorzy publikacji zajęli się badaniem jednego z podstawowych modeli bayesowskich w statystyce małych obszarów, w którym modelowaniu podlegają badane jednostki, a nie estymatory. Opracowanie koncentruje się na bogatej analizie symulacyjnej jakości aproksymacji wartości estymatorów w modelu bayesowskim za pomocą metod MCMC (*Monte Carlo Markov Chain*). Zawarte w pracy wyniki stabilności estymatorów, uzyskane w hierarchicznym modelu bayesowskim, wskazują na możliwość zastosowania tego typu podejścia w praktyce statystycznej.

Ze względu na badane cechy wyniki opisywanych symulacji są szczególnie obiecujące w sytuacji dużego zróżnicowania małych obszarów. Przekonuje to do dalszych poszukiwań metodologicznych, jak również do wyrażenia ostrożnego optymizmu co do możliwości zastosowania opracowanej metodologii w praktyce badań społeczno-ekonomicznych. Z tego powodu ten tom „Biblioteki Wiadomości Statystycznych" warto polecić zarówno pracownikom statystyki publicznej, jak i badaczom zainteresowanym statystyką małych obszarów.

Prezes
Głównego Urzędu Statystycznego

/-/ Józef Oleński

Spis treści

Wstęp 5

1. Model i metoda obliczeniowa 7

1.1. Hierarchiczny model bayesowski na poziomie jednostek 7

1.2. Próbnik Gibbsa 8

2. Doświadczenia wstępne.
Badanie rozkładów *a posteriori*
dla długich przebiegów próbnika 13

2.1. Zróżnicowane małe obszary,
nieinformacyjny rozkład *a priori* 14

2.2. Niezróżnicowane małe obszary,
nieinformacyjny rozkład *a priori* 25

2.3. Niezróżnicowane małe obszary,
informacyjny rozkład *a priori* 38

3. Błąd średniokwadratowy i obciążenie 53

4. Badanie wpływu punktu startowego na działanie próbnika 61

5. Badanie wpływu rozkładów *a priori* na działanie próbnika
i na estymatory bayesowskie 69

6. Konkluzje 83

Literatura 87

Wstęp

Opracowanie poświęcone jest symulacyjnym badaniom próbnika Gibbsa dla hierarchicznego modelu bayesowskiego małych obszarów. Model ten, w całej ogólności, opisany jest w podrozdziale 1.1. Jest to model nakładany na jednostki populacji, a nie na estymatory w małych obszarach. Te dwa podejścia są opisane jako podstawowe dla statystyki małych obszarów w monografii Rao (2003). Dotychczas mało wiadomo o dokładności estymacji bądź predykcji parametrów na poziomie małych obszarów przy wykorzystaniu właśnie tego podstawowego modelu, a w szczególności, przy wykorzystaniu próbnika Gibbsa. W tej sytuacji problem się dodatkowo komplikuje, bo zaburzenia wprowadza sam próbnik, tzn. nie realizuje on dokładnych wartości estymatorów, czy predyktorów bayesowskich, a jedynie je przybliża.

Nasza praca koncentruje się właśnie na tym drugim aspekcie: przedmiotem zainteresowania jest, bardziej niż jakość estymatorów bayesowskich w modelu hierarchicznym, analiza zbieżności próbnika Gibbsa do tych właśnie estymatorów, czyli odpowiednich średnich *a posteriori*. Należy jednak podkreślić, że wyniki symulacyjne pozwalają na sformułowanie pewnych wniosków również dotyczących jakości samych estymatorów *a posteriori*. Próbnik jest opisany w podrozdziale 1.2. Implementowany był za pomocą pakietu R.

W rozdziale 2 przedstawiono wyniki trzech eksperymentów z bardzo długimi przebiegami o 20500 krokach. Zbadano symulacyjnie zachowanie zmiennych występujących w modelu, zarówno zmiennych o charakterze pomocniczym, jak i zmiennych, których estymacja jest celem ostatecznym zastosowania próbnika. Okazuje się, że duży wpływ na

jakość otrzymywanych wyników ma stosunek zmienności błędów losowych do zmienności między małymi obszarami.

Jeżeli małe obszary są znacząco zróżnicowane (podrodział 2.1), to próbnik szybko się stabilizuje i bardzo dobrze estymuje odpowiednie parametry *a posteriori*. W szczególności wtedy, gdy rozkłady *a priori* są nieinformacyjne, tzn. jedyna informacja poza modelem, którą wykorzystuje próbnik, to dane z obserwacji. Jeżeli to zróżnicowanie jest niewielkie w stosunku do wariancji błędu losowego (podrozdział 2.2), okres stabilizacji próbnika znacząco się wydłuża i otrzymywane estymatory parametrów *a posteriori* są mało efektywne. W dodatku są one dalekie od parametrów modelowych. Nieco można poprawić sytuację zadając informacyjne rozkłady *a priori* , tzn. takie, które zawierają przybliżoną informację o rozkładach modelowych parametrów. Jednak i w tym przypadku (podrozdział 2.3), mimo że próbnik dobrze przybliża parametry *a posteriori*, są one odległe od interesujących nas parametrów małych obszarów.

Doświadczenie zdobyte w badaniach długich trajektorii wpłynęło na to, iż do dalszych analiz wybrano model ze znaczącym zróżnicowaniem małych obszarów. Dla takiego modelu przeprowadzono serię 200 symulacji, przy czym w każdej liczba kroków próbnika wynosiła 10500. Ponownie do analiz wykorzystywano trajektorie bez początkowych 500 kroków. W takiej sytuacji, przede wszystkim zbadano empirycznie błąd średniokwadratowy dla estymatorów parametrów *a posteriori* oraz ich obciążenie. Tego dotyczy rozdział 3. W rozdziale 4 zajęto się wpływem warunków początkowych na wyniki otrzymywane przez próbnik. Natomiast w rozdziale 5 wrócono do zagadnienia wpływu rozkładu *a priori* na szybkość zbieżności próbnika i na jakość otrzymywanych rezultatów. Należy podkreślić, że w odróżnieniu od wyników uzyskanych w tym kierunku w podrozdziale 2.3, tym razem poziom wariancji błędów nie jest wysoki w porównaniu ze zmiennością między małymi obszarami. W krótkim rozdziale 6 zebrano wnioski wynikające z przeprowadzonej analizy symulacyjnej.

Rozdział 1

Model i metoda obliczeniowa

1.1. Hierarchiczny model bayesowski na poziomie jednostek

W rozpatrywanym modelu przez y_j oznaczono wartość obserwowanej cechy $\mathcal{Y}$ dla j-tej jednostki populacji o liczebności J ($j = 1, \ldots, J$). Zakłada się hierarchiczny model bayesowski opisany dalej. W modelu tym x_j oznacza wektor znanych zmiennych pomocniczych dla j-tej jednostki, β oznacza nieznany wektor efektów stałych, wspólny dla wszystkich jednostek populacji, θ_i oznacza efekt losowy wspólny dla wszystkich jednostek z małego obszaru o numerze i ($i = 1, \ldots, I$, gdzie I jest liczbą małych obszarów), natomiast s^2 oznacza wspólną nieznaną wariancję błędu obserwacji.

Celem estymacyjnym w statystyce małych obszrów jest prognoza wektora (θ_i) oraz, znacznie łatwiejsza, estymacja parametru β. Rozważany model zdefiniowany jest przez następujący hierarchiczny układ prawdopodobieństw warunkowych:

$$y = (y_j)|\beta, \theta, s^2 \sim \bigotimes_{j=1}^{J} \mathcal{N}\left(x_j^T\beta + \theta_{ma(j)}, s^2\right),$$

$$\beta \sim \mathcal{N}(b, \Sigma),$$

$$\theta = (\theta_i)|v^2, \mu \sim \bigotimes_{i=1}^{I} \mathcal{N}(\mu, v^2),$$

$$v^{-2} \sim \mathcal{G}(p, \lambda),$$

$$\mu \sim \mathcal{N}(m, \tau^2),$$

$$s^{-2} \sim \mathcal{G}(q, \kappa),$$

gdzie $\mathcal{N}$ oznacza rozkład normalny, $\mathcal{G}$ rozkład gamma, symbol $ma(j)$ oznacza numer małego obszaru, do którego należy jednostka j, a symbol $\bigotimes$ produkt rozkładów (niezależność odpowiednich zmiennych losowych). We wzorach przyjęto, że populacja składająca się z J jednostek podzielona jest na I małych obszarów.

Modele tego rodzaju są standardowo rozważane w statystyce małych obszarów, zob. np. Rao (2003), Ghosh (2001) albo wcześniej w Ghosh i Rao (1994). Brak jest natomiast opracowań dotyczących jakości otrzymywanych oszacowań. Z tego względu dogłębna analiza ich własności jest bardzo istotna dla pełnego wprowadzenia tego podejścia w realnych badaniach próbkowych.

Stosowane w naszym opracowaniu podejście bayesowskie polega na tym, że estymatory uzyskiwane mają być jako odpowiednie średnie *a posteriori*. Warto mocno podkreślić, że takie podejście wyklucza wpływ schematu losowania na otrzymywane predyktory i estymatory. Rozkłady *a posteriori* są analitycznie nieosiągalne, dlatego stosuje się do ich otrzymania podejście MCMC (*Markov Chain Monte Carlo*) wykorzystując próbnik Gibbsa.

1.2. Próbnik Gibbsa

Próbnik Gibbsa jest zdefiniowany przez podanie wszystkich brzegowych rozkładów warunkowych, które w przypadku rozważanego mod-

elu hierarchicznego dają się *explicite* wypisać wzorami analitycznymi i w konsekwencji realizować na drodze symulacji komputerowej. Rozkłady te mają postać:

$$v^{-2}|y,\beta,\theta,s^2,\mu \sim \mathcal{G}\left(\frac{I}{2}+p,\ \lambda+\frac{1}{2}\sum_i(\theta_i-\mu)^2\right),$$

$$s^{-2}|y,\beta,\theta,v^2,\mu \sim \mathcal{G}\left(\frac{J}{2}+q,\ \kappa+\frac{1}{2}\sum_j(y_j-x_j^T\beta-\theta_{ma(j)})^2\right),$$

$$(\theta_i)|y,\beta,s^2,v^2,\mu \sim$$

$$\bigotimes_i \mathcal{N}\left(\frac{s^2}{s^2+v^2J_i}\mu+\frac{v^2}{s^2+v^2J_i}\sum_{j:ma(j)=i}(y_j-x_j^T\beta),\ \frac{s^2v^2}{s^2+v^2J_i}\right),$$

$$\beta|y,\theta,s^2,v^2,\mu \sim$$

$$\mathcal{N}\left(\left(\sum_j x_jx_j^T+s^2\Sigma^{-1}\right)^{-1}\left(\sum_j(y_j-\theta_{ma(j)})x_j+s^2\Sigma^{-1}b\right),\right.$$

$$\left. s^2\left(\sum_j x_jx_j^T+s^2\Sigma^{-1}\right)^{-1}\right),$$

$$\mu|y,\beta,s^2,v^2,\theta \sim \mathcal{N}\left(\frac{v^2}{v^2+\tau^2I}m+\frac{\tau^2}{v^2+\tau^2I}\sum_i\theta_i,\ \frac{v^2\tau^2}{v^2+\tau^2I}\right).$$

We wzorach tych liczby (J_i) oznaczają liczebnoś kolejnych małych obszarów.

Oczywiście, w wyniku stosowania próbnika Gibbsa można otrzymy-

wać jedynie przybliżenia rozkładów *a posteriori*. Istotne jest zatem zbadanie własności tego właśnie próbnika. Dla rozważanego modelu, ze względu na potencjalne zastosowania, podejmowane są w ostatnich latach liczne próby szacowania szybkości zbieżności próbnika Gibbsa. Najczęściej stosowanym kryterium jest odległość pełnego wahania pomiędzy rozkładami generowanych przez próbnik zmiennych losowych a docelowymi rozkładami *a posteriori*. W tym celu wykorzystuje się podejście oparte na tzw. funkcji dryfu, zob. np. Rosenthal (1995), Jones i Hobert (2004), Roberts i Rosenthal (2004), Łatuszyński (2008) oraz Łatuszyński i Niemiro (2009).

W pewnym sensie, model rozważany w tym opracowaniu jest modelem testowym dla różnych teoretycznych metod analizy algorytmów. Należy podkreślić, że otrzymywane oszacowania są bardzo niedokładne. Przy tym brak dogłębnych analiz symulacyjnych. Tę lukę ma wypełnić nasza praca. W zasadzie dotyczy ona pełnego modelu i odpowiedniego próbnika Gibbsa. Jedyne uproszczenie polega na zaniedbaniu paramteru β, tzn. przyjęciu, że jest on równy 0 (albo równoważnie, że wszystkie x_j-ty są zerowe).

Oprócz trajektorii generowanych bezpośrednio przez próbnik, obliczano także trajektorie „warunkowej metody Monte Carlo" typu Rao-Blackwella. Są one zdefiniowane w sposób następujący: Niech funkcja f przyporządkowuje wielowymiarowej zmiennej $x = (s, v, \mu, \theta)$ jedną ze współrzędnych. Trajektorię generowaną bezpośrednio przez próbnik stanowi ciąg $f_k := f(X_k)$. Przez trajektorię „warunkowej metody Monte Carlo" typu Rao-Blackwella rozumiemy ciąg

$$f_k^{\mathrm{R-B}} := \mathbb{E}[f(X_{k+1})|X_k].$$

Równolegle do standardowego estymatora „średnia wzdłuż trajektorii", czyli

$$\bar{f}_{b,n} := \frac{1}{n} \sum_{k=b+1}^{b+n} f(X_k),$$

obliczano „warunkowy estymator Monte Carlo" lub inaczej estymator typu Rao-Blackwella,

$$\bar{f}_{b,n}^{\mathrm{R-B}} := \frac{1}{n} \sum_{k=b+1}^{b+n} \mathbb{E}[f(X_{k+1})|X_k].$$

Jeśli, np. przedmiotem zainteresowania jest współrzędna μ, czyli

$$f(x) = f(s, v, \mu, \theta) = \mu,$$

to stosowane będą oznaczenia

$$\mu_k := f_k\,, \quad \mu_k^{\mathrm{R-B}} := f_k^{\mathrm{R-B}}\,, \quad \bar{\mu}_{b,n} := \bar{f}_{b,n} \quad \text{oraz} \quad \bar{\mu}_{b,n}^{\mathrm{R-B}} := \bar{f}_{b,n}^{\mathrm{R-B}}.$$

Podobne symbole używane będą dla trajektorii i „średnich wzdłuż trajektorii" dla zmiennych s i v. Dla zmiennych (θ_i) trzeba używać oznaczeń z dodatkowym wskaźnikiem i numerującym obszary, a więc $\theta_{i,k}$, $\theta_{i,k}^{\mathrm{R-B}}$, $\bar{\theta}_{i,b,n}$ i $\bar{\theta}_{i,b,n}^{\mathrm{R-B}}$.

Rozdział 2

Doświadczenia wstępne. Badanie rozkładów *a posteriori* dla długich przebiegów próbnika

Metodologia badań symulacyjnych była następująca: Dla wybranego zestawu parametrów modelu, czyli zmiennych μ, s, v oraz zadanego zestawu liczebności małych obszarów (J_i) generowano sztuczne dane, czyli układ zmiennych (θ_i) oraz (y_j) zgodnie z modelem opisanym w podrozdziale 1.1.

Doświadczenia dotyczyły modelu bez zmiennych pomocniczych x_j. Innymi słowy zakładano, że $\beta = 0$. Dla tych danych wielokrotnie powtarzano przebieg algorytmu (próbnika Gibbsa) opisanego w podrozdziale 1.2. Badano wpływ zadanych parametrów rozkładu *a priori*, czyli $p, \lambda, q, \kappa, m, \tau$ oraz wyboru punktu startowego na zachowanie algorytmu, a w szczególności na cechy rozkładu *a posteriori* oraz na szybkość zbieżności algorytmu.

Należy podkreślić, że dla ustalonych danych, estymatory generowane przez algorytm nie są zbieżne do „modelowych" wielkości μ, s, v (użytych do uprzedniego wygenerowania danych i wobec tego znanych), tylko do wartości oczekiwanych *a posteriori* $\mathbb{E}(\mu|y)$, $\mathbb{E}(s|y)$, $\mathbb{E}(v|y)$. Podobna uwaga dotyczy układu zmiennych losowych (θ_i). Rozkład *a posteriori* w rozważanym modelu nie jest możliwy do obliczenia *explicite*.

W istocie jedynym znanym sposobem obliczania wielkości charakteryzujących ten rozkład jest użycie metod MCMC. Wobec tego, nawet

dla sztucznych danych, rozkład docelowy i jego charakterystyki nie są znane. Aby temu zaradzić, każdą serię doświadczeń rozpoczyna się od długiego przebiegu próbnika, wykonanego w celu możliwie dokładnego wyestymowania rozkładu docelowego. Estymatory uzyskane na podstawie tego wstępnego etapu są następnie traktowane jako dobre przybliżenie wielkości „docelowych" (granicznych) i używane jako punkt odniesienia do porównań krótszych przebiegów i szacowania błędów.

2.1. Zróżnicowane małe obszary, nieinformacyjny rozkład *a priori*

W pierwszej serii doświadczeń zestaw parametrów jest następujący:

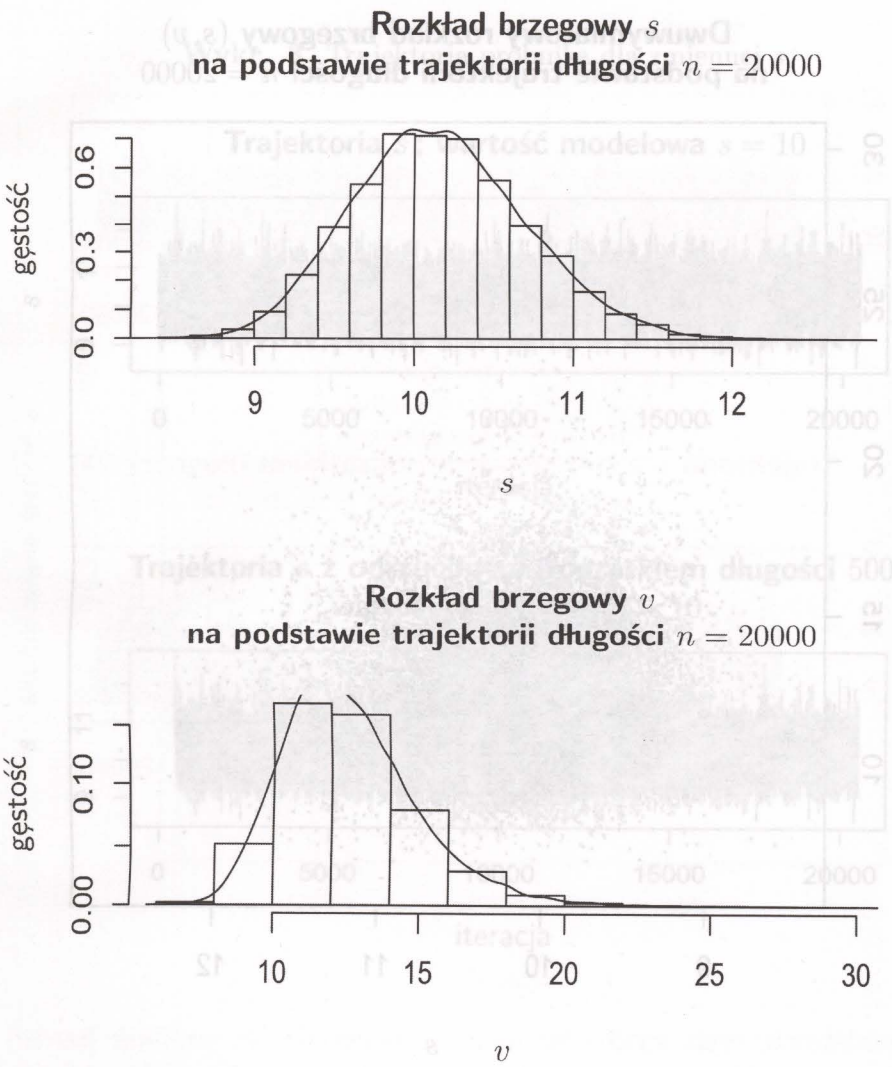
$$\mu = 0, \quad s = 10, \quad v = 10, \quad J_i = 10 \ (i = 1, \ldots, I = 20).$$

Rozważa się niewłaściwy, nieinformacyjny rozkład *a priori* z parametrami:

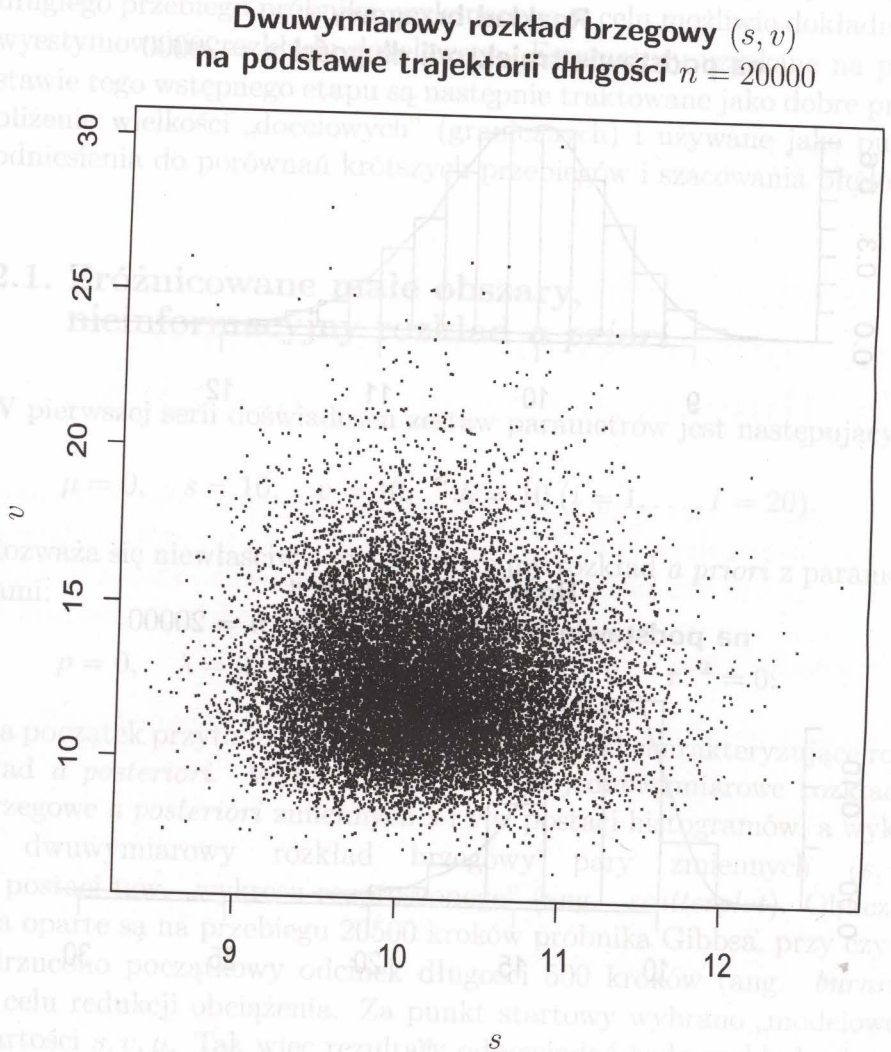
$$p = 0, \quad \lambda = 0, \quad q = 0, \quad \kappa = 0, \quad m = 0, \quad \tau^{-2} = 0.$$

Na początek przytoczone są przykładowe wyniki charakteryzujące rozkład *a posteriori*. Wykr. 1 przedstawia jednowymiarowe rozkłady brzegowe *a posteriori* zmiennych s i v w postaci histogramów, a wykr. 2 dwuwymiarowy rozkład brzegowy pary zmiennych (s, v) w postaci tzw. „wykresu rozproszonego" (ang. *scatterplot*). Obliczenia oparte są na przebiegu 20500 kroków próbnika Gibbsa, przy czym odrzucono początkowy odcinek długości 500 kroków (ang. *burnin*) w celu redukcji obciążenia. Za punkt startowy wybrano „modelowe" wartości s, v, μ. Tak więc rezultaty odpowiadać będą rozkładowi empirycznemu dla próbki 20000 (zależnych) zmiennych losowych. Jak zostanie uzasadnione nieco później, można przyjąć, że próbnik jest w tym przykładzie szybko zbieżny, i że wielkość relacjonowanego doświadczenia jest dostateczna do wiarygodnego poznania rozkładu *a posteriori*.

Wykr.1. Jednowymiarowe rozkłady brzegowe zmiennych s i v



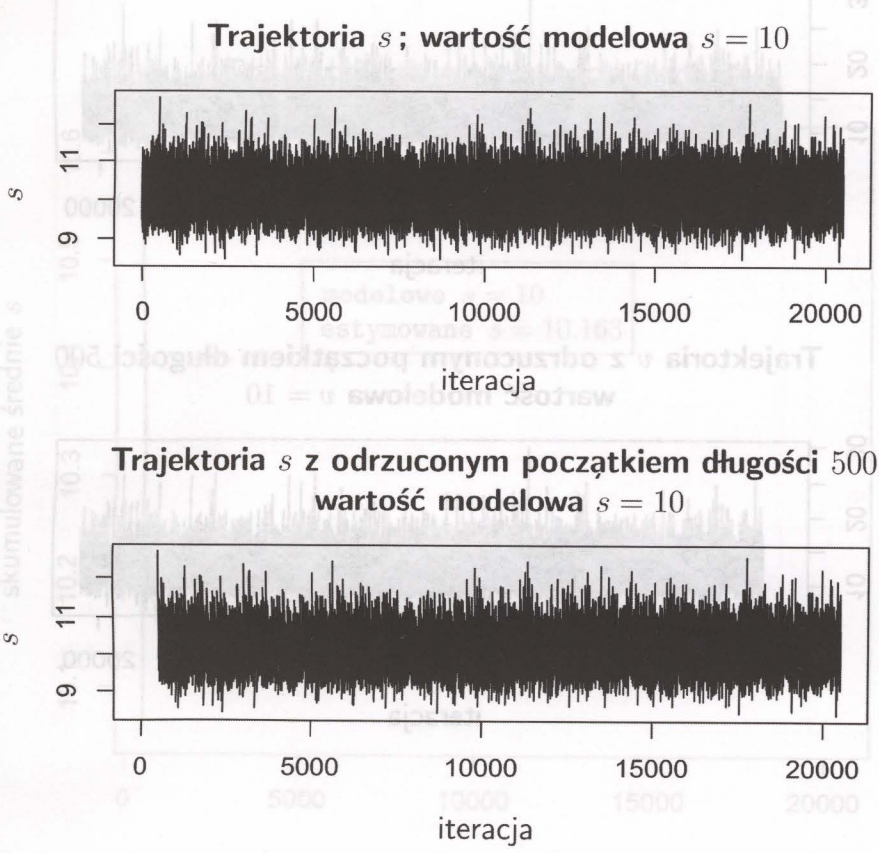
Wykr. 2. Dwuwymiarowy rozkład brzegowy pary zmiennych (s, v)



Rozkłady zmiennych μ i θ wyglądają dość podobnie, więc nie ma potrzeby przytaczania odpowiednich wykresów.

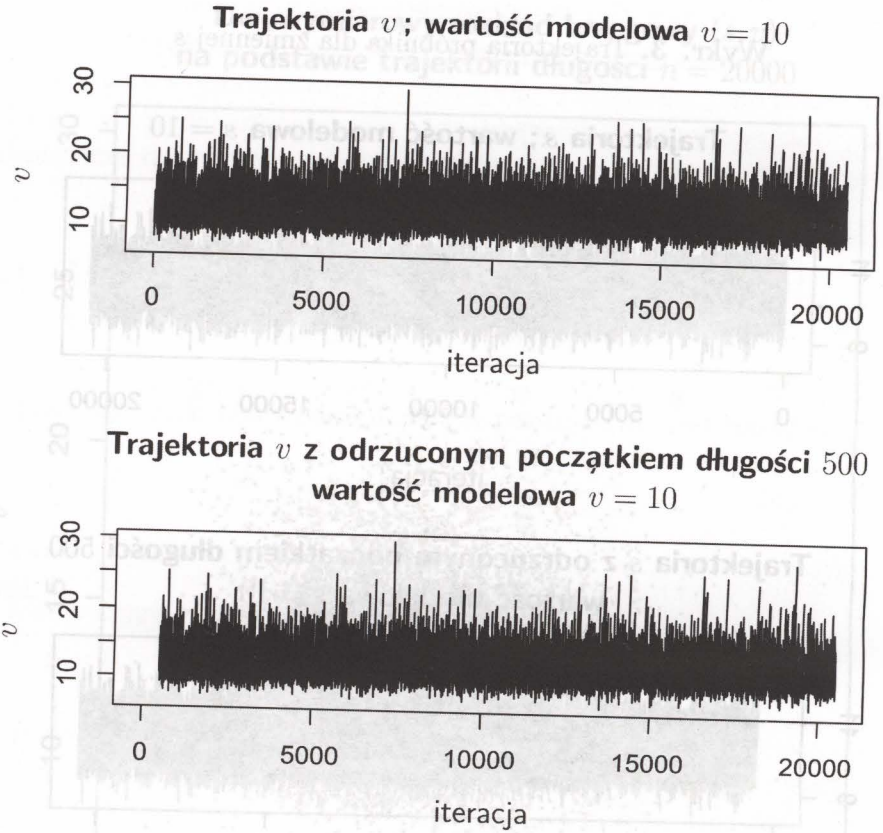
Następne wykresy (wykr. 3 i 4) obrazują przebieg trajektorii próbnika i przedstawiają ewolucję zmiennych s_k i v_k.

Wykr. 3. Trajektoria próbnika dla zmiennej s



CENTRALNA BIBLIOTEKA STATYSTYCZNA
WARSZAWA

Wykr. 4. Trajektoria próbnika dla zmiennej v



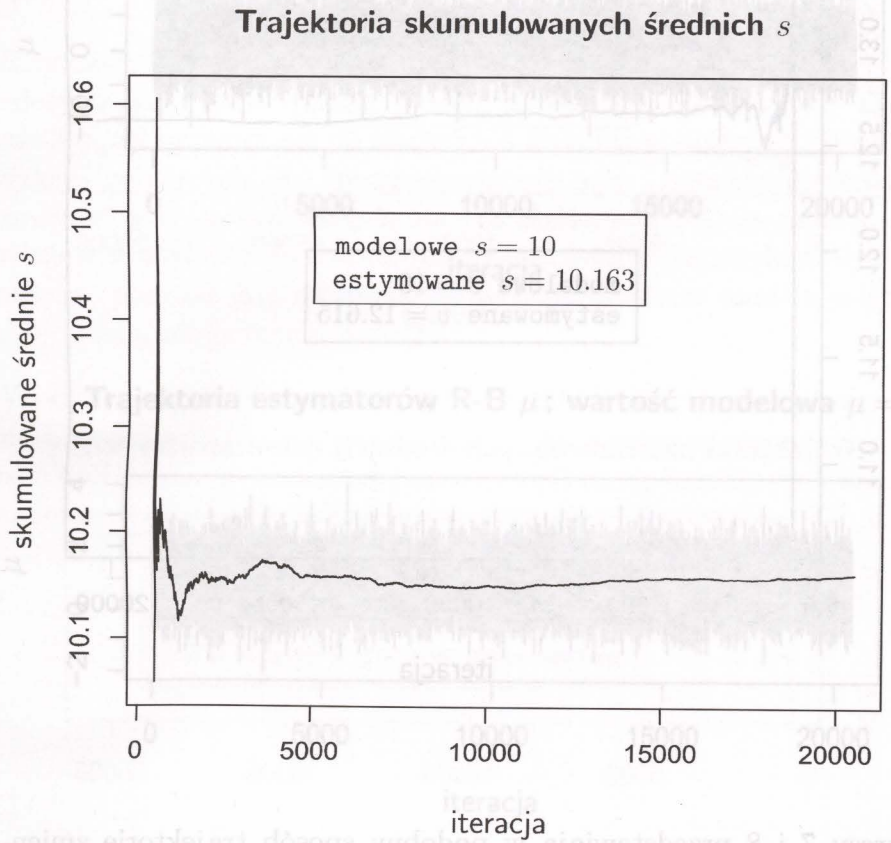
Na podstawie tych wykresów można stwierdzić, że próbnik bardzo szybko „osiąga" stan stacjonarny. W istocie jest tak szybko zbieżny do rozkładu stacjonarnego, że odrzucenie początkowych $b = 500$ obserwacji wydaje się przesadną ostrożnością.

Na wykresach 5 i 6 przedstawiono ewolucję „średnich wzdłuż trajektorii", czyli estymatorów $\bar{s}_{b,n}$ i $\bar{v}_{b,n}$ w zależności od długości n uśrednianej trajektorii. Zwróćmy uwagę, że średnie obliczano na podstawie trajektorii:

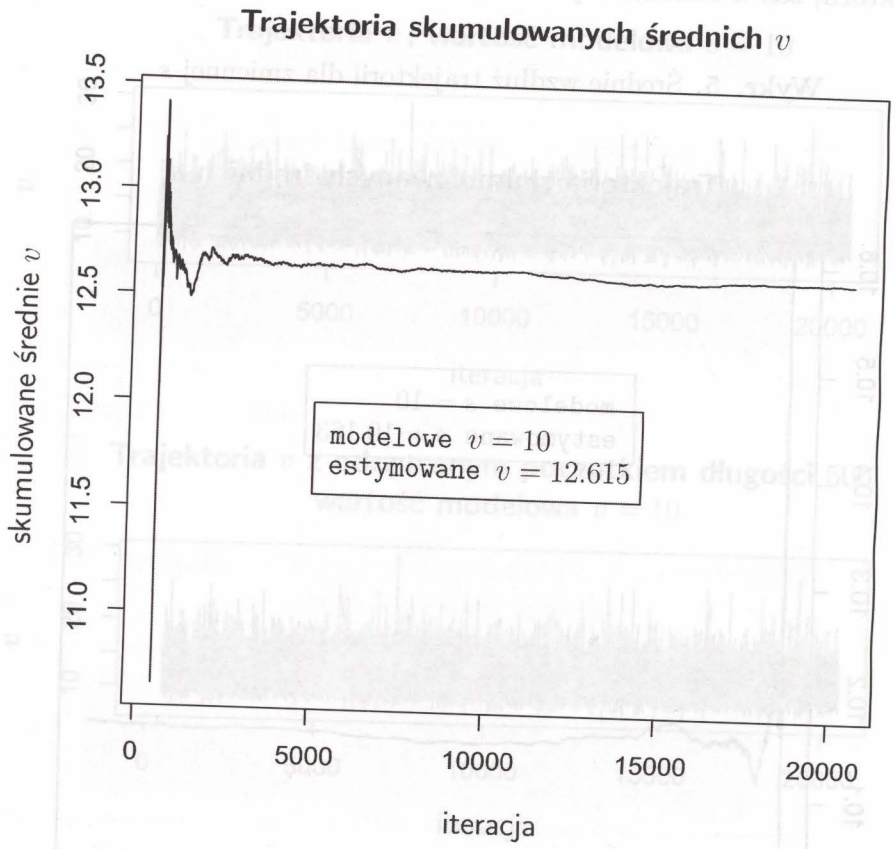
$$X_{b+1}, \ldots, X_{b+n},$$

gdzie $b = 500$ oznacza długość odrzuconego początkowego odcinka trajektorii, zaś n zmienia się od 1 do 20000.

Wykr. 5. Średnie wzdłuż trajektorii dla zmiennej s

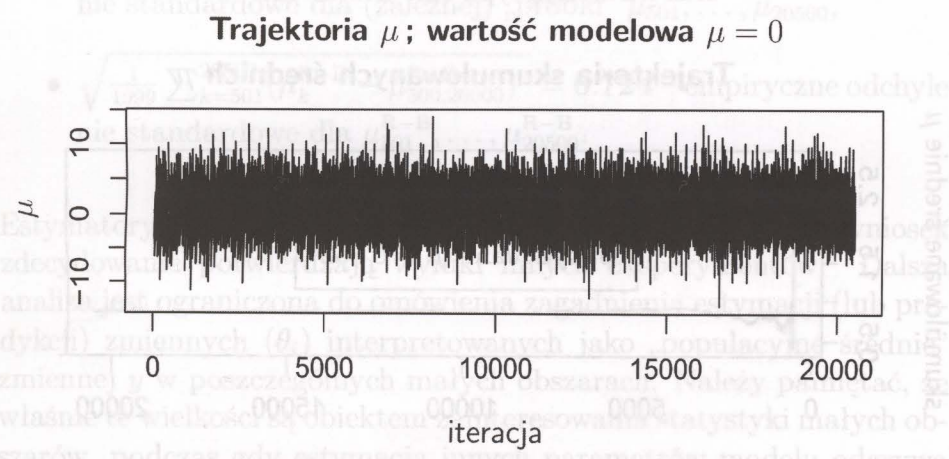


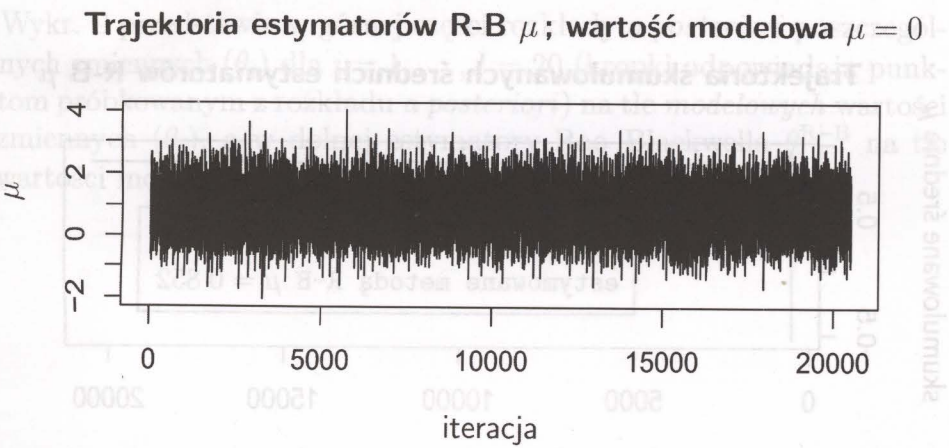
Wykr. 6. Średnie wzdłuż trajektorii dla zmiennej v



Wykresy 7 i 8 przedstawiają w podobny sposób trajektorię zmiennej μ. Zrezygnowano tu z prezentowania początkowego odcinka 500 iteracji. Równolegle do trajektorii wartości μ_k generowanych przez próbnik obliczano także trajektorię $\mu_k^{\mathrm{R-B}}$ „warunkowej metody Monte Carlo" typu Rao-Blackwella.

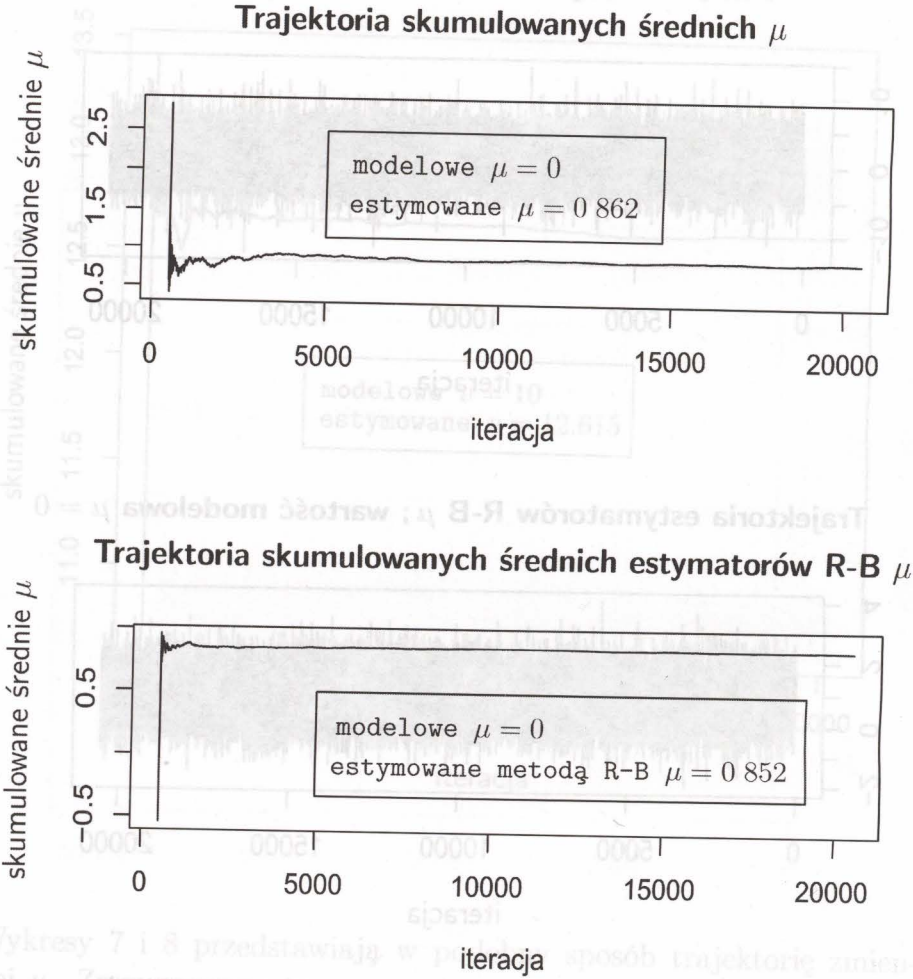
Wykr. 7. Trajektoria próbnika dla zmiennej μ





Zwraca uwagę mniejszy rozrzut warunkowych wartości oczekiwanych $\mu_k^{\mathrm{R-B}}$ (w dolnej części wykr. 7). W rezultacie obliczone estymatory typu Rao-Blackwella maja wyraźnie mniejszą wariancję i mniejszy błąd średniokwadratowy. Poniżej przytoczone są wartości estymatorów $\bar{\mu}_{b,n}$ i $\bar{\mu}_{b,n}^{\mathrm{R-B}}$ obliczone na podstawie trajektorii długości n w zależności od n, dla $n = 1, \ldots, 20000$. Jak zwykle, w przedstawianych tu doświadczeniach, $b = 500$.

Wykr. 8. Średnie wzdłuż trajektorii dla zmiennej μ oraz ich wersja typu Rao-Blackwella



Prawidłowości widoczne na omawianych wykresach potwierdzają obliczenia estymatorów i ich błędów:

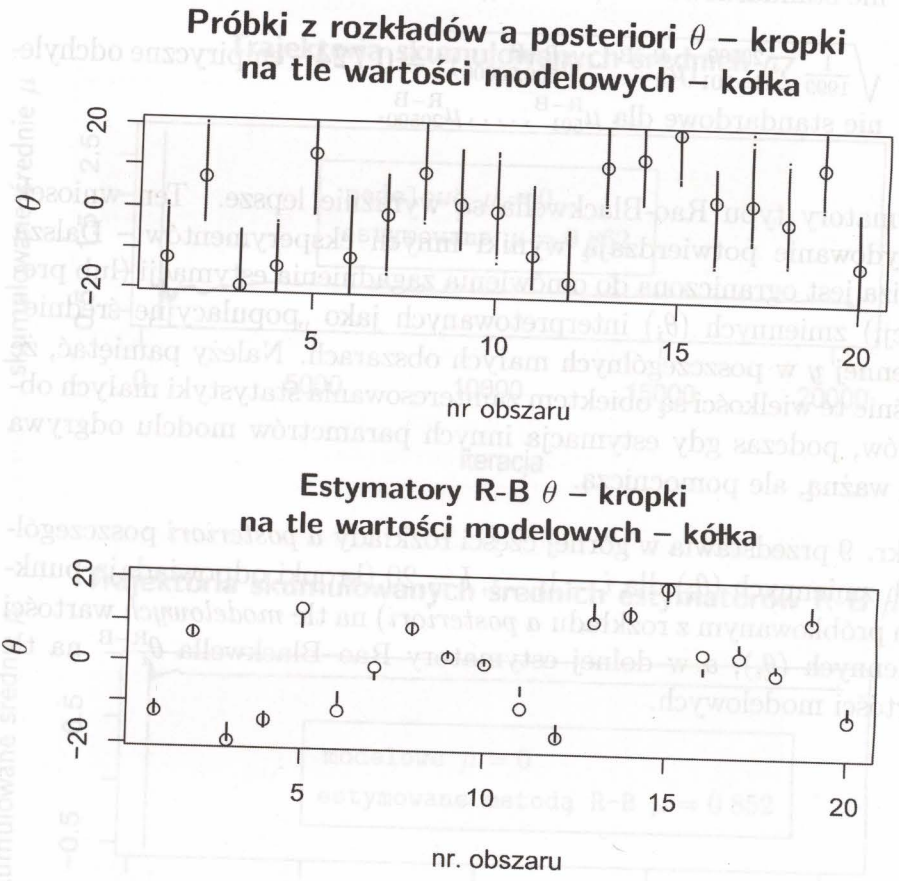
- $\bar{\mu}_{500,20000} = 0.862$ – estymator „średnia wzdłuż trajektorii",
- $\bar{\mu}^{\mathrm{R-B}}_{500,20000} = 0.852$ – estymator Rao-Blackwella,

- $\sqrt{\frac{1}{1999}\sum_{k=501}^{20500}(\mu_k - \bar{\mu}_{500,20000})^2} = \mathbf{2.966}$ – empiryczne odchylenie standardowe dla (zależnej) „próbki" $\mu_{501}, \ldots, \mu_{20500}$,
- $\sqrt{\frac{1}{1999}\sum_{k=501}^{20500}(\mu_k^{\mathrm{R-B}} - \bar{\mu}_{500,20000}^{\mathrm{R-B}})^2} = \mathbf{0.724}$ – empiryczne odchylenie standardowe dla $\mu_{501}^{\mathrm{R-B}}, \ldots, \mu_{20500}^{\mathrm{R-B}}$.

Estymatory typu Rao-Blackwella są wyraźnie lepsze. Ten wniosek zdecydowanie potwierdzają wyniki innych eksperymentów. Dalsza analiza jest ograniczona do omówienia zagadnienia estymacji (lub predykcji) zmiennych (θ_i) interpretowanych jako „populacyjne średnie" zmiennej y w poszczególnych małych obszarach. Należy pamiętać, że właśnie te wielkości są obiektem zainteresowania statystyki małych obszarów, podczas gdy estymacja innych parametrów modelu odgrywa rolę ważną, ale pomocniczą.

Wykr. 9 przedstawia w górnej części rozkłady *a posteriori* poszczególnych zmiennych (θ_i) dla $i = 1, \ldots, I = 20$ (kropki odpowiadają punktom próbkowanym z rozkładu *a posteriori*) na tle *modelowych* wartości zmiennych (θ_i), a w dolnej estymatory Rao–Blackwella $\theta_i^{\mathrm{R-B}}$ na tle wartości modelowych.

Wykr. 9. Rozkłady a posteriori i modelowe wartości zmiennych θ_i



Z kolei w tablicy 1 przedstawione są estymatory średnich *a posteriori* dla wielkości θ_i oraz odchylenia standardowe tych estymatorów liczonych wzdłuż trajektorii, również dla wersji Rao–Blackwella.

Tabl. 1. Mały obszar: i, wartości modelowe: θ_i, ich estymatory: $\hat{\theta}_i$ i odchylenia standardowe rozkładów *a posteriori*: σ_i, estymatory R-B: $\hat{\theta}_i^{(R-B)}$ i ich odchylenia standardowe: $\sigma_i^{(R-B)}$

i	θ_i	$\hat{\theta}_i$	σ_i	$\hat{\theta}_i^{(R-B)}$	$\sigma_i^{(R-B)}$
1	-12.337	-12.112	3.132	-12.117	0.367
2	7.250	7.934	3.127	7.934	0.254
3	-19.174	-18.013	3.136	-17.967	0.496
4	-14.073	-14.701	3.133	-14.676	0.422
5	13.369	11.000	3.114	11.026	0.310
6	-11.583	-8.822	3.130	-8.807	0.300
7	-0.745	-3.087	3.102	-3.081	0.209
8	9.276	9.439	3.130	9.454	0.281
9	2.089	3.053	3.115	3.112	0.194
10	0.544	1.152	3.114	1.157	0.186
11	-10.011	-6.085	3.130	-6.101	0.252
12	-16.685	-16.299	3.125	-16.265	0.458
13	11.766	15.355	3.120	15.325	0.400
14	13.753	14.752	3.158	14.721	0.387
15	20.090	20.808	3.132	20.817	0.522
16	4.152	-0.031	3.145	-0.043	0.187
17	3.518	6.215	3.077	6.207	0.227
18	-0.740	0.775	3.129	0.781	0.186
19	12.426	15.543	3.153	15.561	0.405
20	-11.306	-9.839	3.130	-9.834	0.320

Uwaga. Wielkości w tabl. 1 są obliczone w identyczny sposób jak analogiczne wielkości dla zmiennej μ, podane wcześniej.

2.2. Niezróżnicowane małe obszary, nieinformacyjny rozkład *a priori*

W drugiej serii doświadczeń zestaw parametrów był następujący:

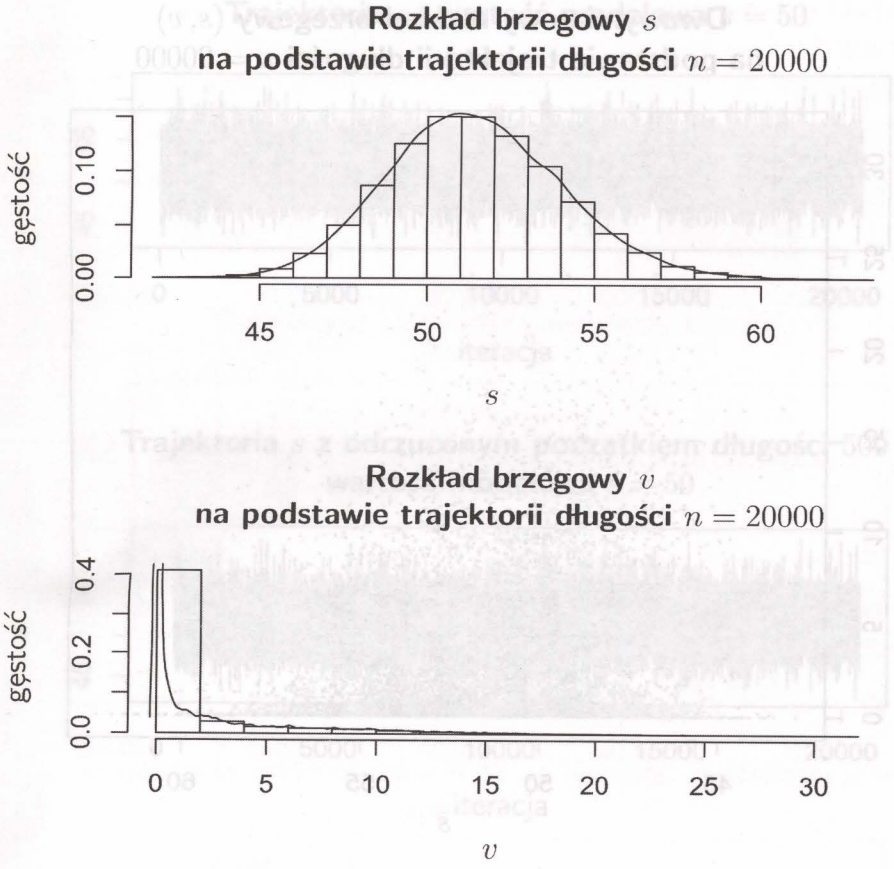
$$\mu = 0, \quad s = 50, \quad v = 10, \quad J_i = 10 \ (i = 1, \ldots, I = 20).$$

Błąd losowy miał zatem znacznie większą zmienność ($s = 50$), niż zmienność między małymi obszarami ($v = 10$). Rozważano ponownie niewłaściwy, nieinformacyjny rozkład *a priori* z następującymi parametrami:

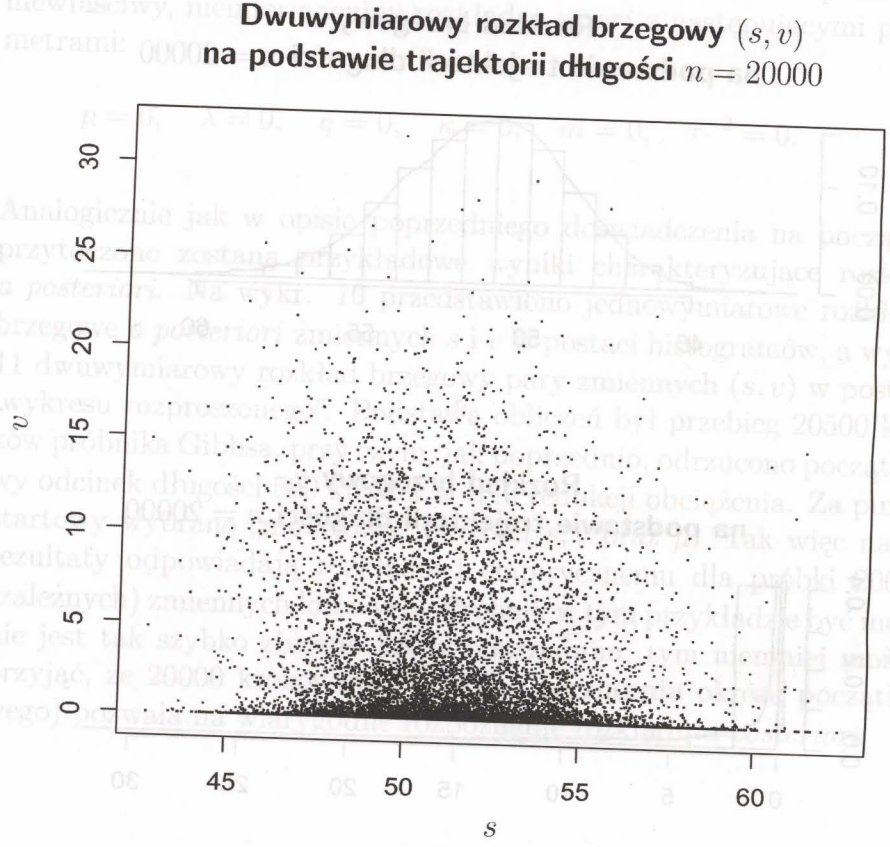
$$p = 0, \quad \lambda = 0, \quad q = 0, \quad \kappa = 0, \quad m = 0, \quad \tau^{-2} = 0.$$

Analogicznie jak w opisie poprzedniego doświadczenia na początku przytoczone zostaną przykładowe wyniki charakteryzujące rozkład *a posteriori*. Na wykr. 10 przedstawiono jednowymiarowe rozkłady brzegowe *a posteriori* zmiennych s i v w postaci histogramów, a wykr. 11 dwuwymiarowy rozkład brzegowy pary zmiennych (s, v) w postaci „wykresu rozproszonego". Podstawą obliczeń był przebieg 20500 kroków próbnika Gibbsa, przy czym, jak poprzednio, odrzucono początkowy odcinek długości 500 kroków w celu redukcji obciążenia. Za punkt startowy wybrano „modelowe" wartości s, v oraz μ. Tak więc nasze rezultaty odpowiadają rozkładowi empirycznemu dla próbki 20000 (zależnych) zmiennych losowych. Próbnik w tym przykładzie być może nie jest tak szybko zbieżny jak w poprzednim, tym niemniej można przyjąć, że 20000 kroków próbnika (po odrzuceniu okresu początkowego) pozwala na wiarygodne rozpoznanie rozkładu *a posteriori*.

Wykr. 10. Jednowymiarowe rozkłady brzegowe zmiennych s i v



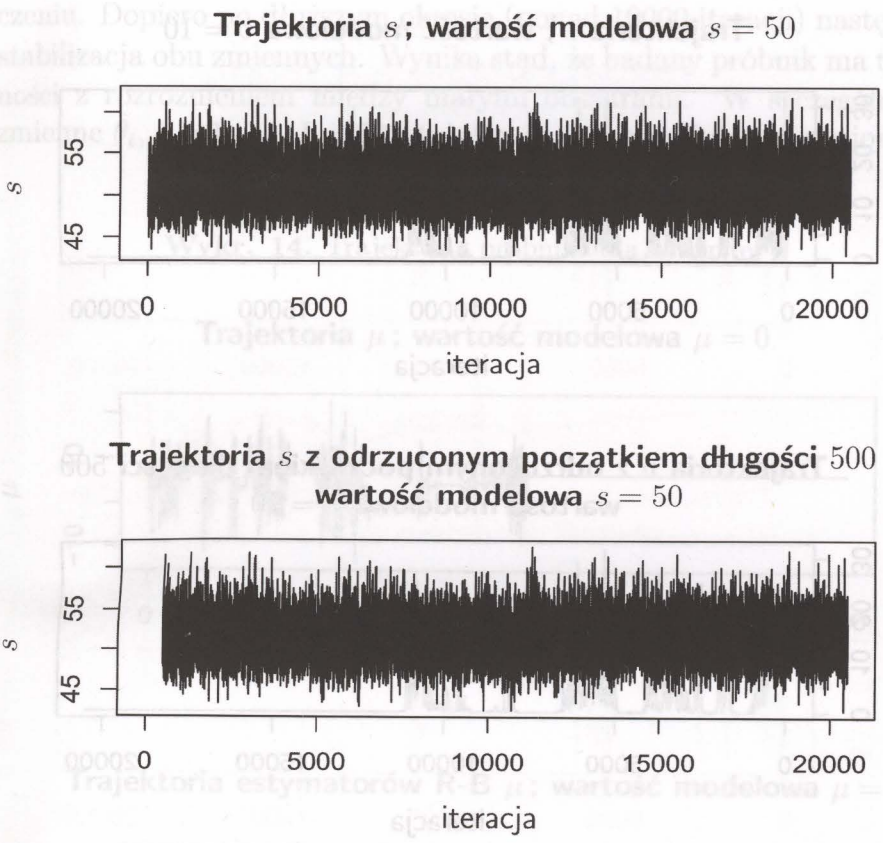
Wykr. 11. Dwuwymiarowy rozkład brzegowy pary zmiennych (s, v)



Obserwujemy znaczące różnice między rozkładami jednowymiarowymi v w tym przykładzie i w przykładzie rozpatrywanym poprzednio. Podobnie rzecz się ma z rozkładami dwuwymiarowymi (s, v). Próbnik ma trudności ze zróżnicowaniem małych obszarów, ściągając wartości v do zera.

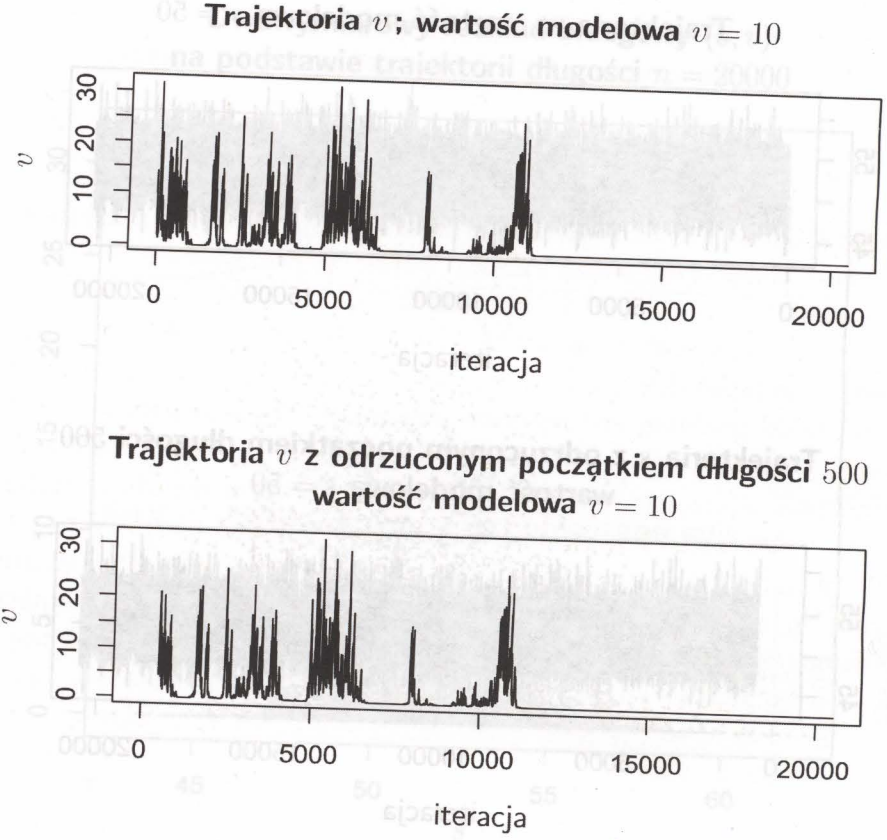
Wykr. 12 i 13 obrazują przebieg trajektorii próbnika i przedstawiają ewolucję zmiennych s_k i v_k.

Wykr. 12. Trajektoria próbnika dla zmiennej s



Trajektoria dla zmiennej s w zasadzie nie różni się od trajektorii w pierwszym doświadczeniu.

Wykr. 13. Trajektoria próbnika dla zmiennej v

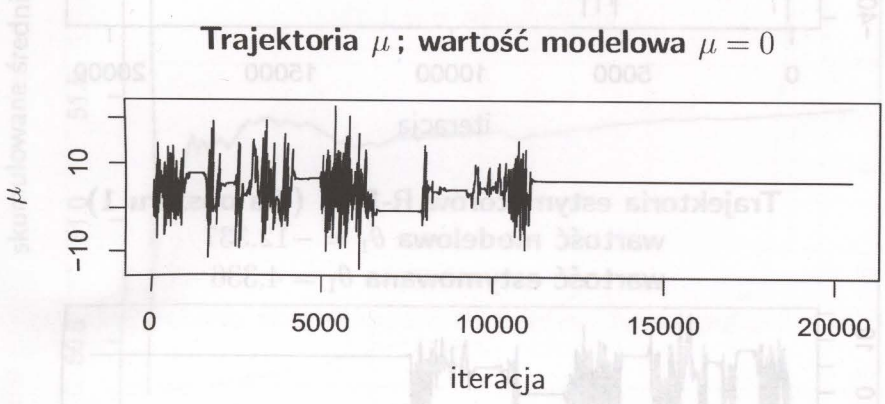


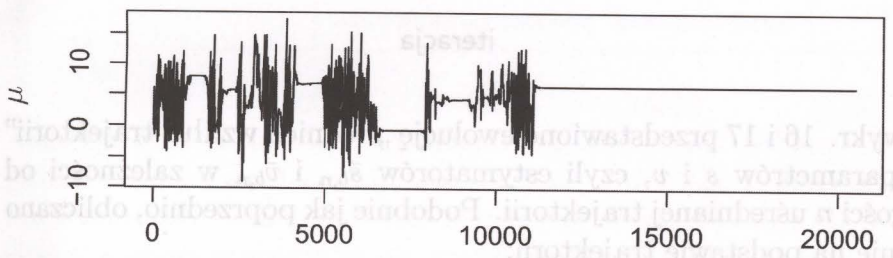
Natomiast rozkład *a posteriori* dla v jest znacząco inny. To już było obserwowane na histogramie. Trajektoria wiele czasu spędza w okolicy zera, ale zdarzają się też nagłe wyskoki do góry, co wskazuje, że rozkład *a posteriori* dla v oprócz znaczącej mody w zerze, może mieć też drugą modę. Te nieregularności rozkładu zmiennej v mają też wpływ na długi czas osiągania stacjonarności przez próbnik. Wpływa to również na zachowanie zmiennych μ i (θ_i).

Kolejne wykresy (wykr. 14 i 15) przedstawiają w podobny sposób trajektorię zmiennej μ bez początkowego odcinka 500 iteracji. Równolegle do trajektorii generowanych przez próbnik wartości μ_k obliczano także trajektorię

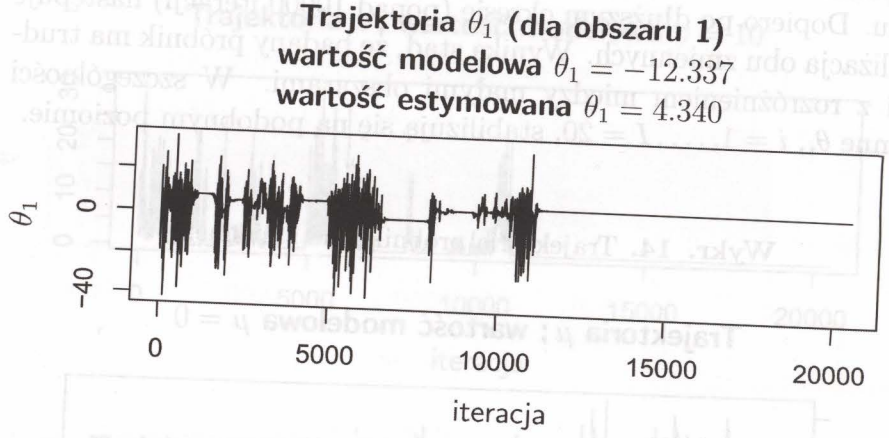
$\mu_k^{\mathrm{R-B}}$ „warunkowej metody Monte Carlo" typu Rao–Blackwella. W obu przypadkach zachowanie jest całkiem inne niż w pierwszym doświadczeniu. Dopiero po dłuższym okresie (ponad 10000 iteracji) następuje stabilizacja obu zmiennych. Wynika stąd, że badany próbnik ma trudności z rozróżnieniem między małymi obszarami. W szczególności zmienne θ_i, $i = 1, \ldots, I = 20$, stabilizują się na podobnym poziomie.

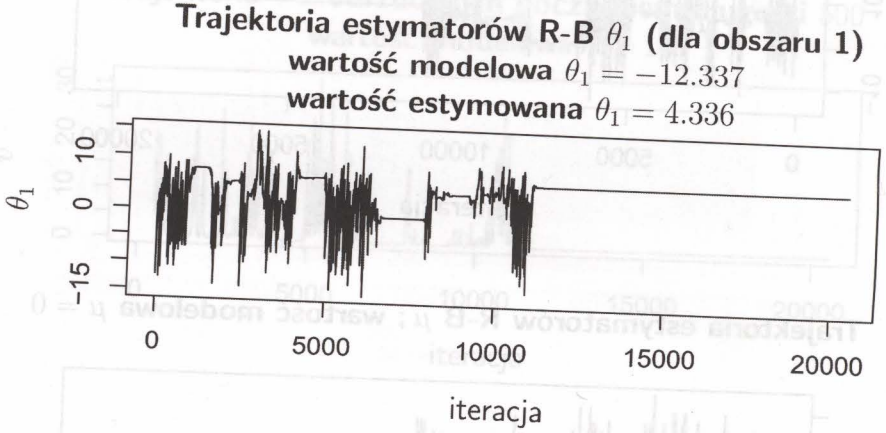
Wykr. 14. Trajektoria próbnika dla zmiennej μ





Wykr. 15. Trajektoria próbnika dla zmiennej θ_1



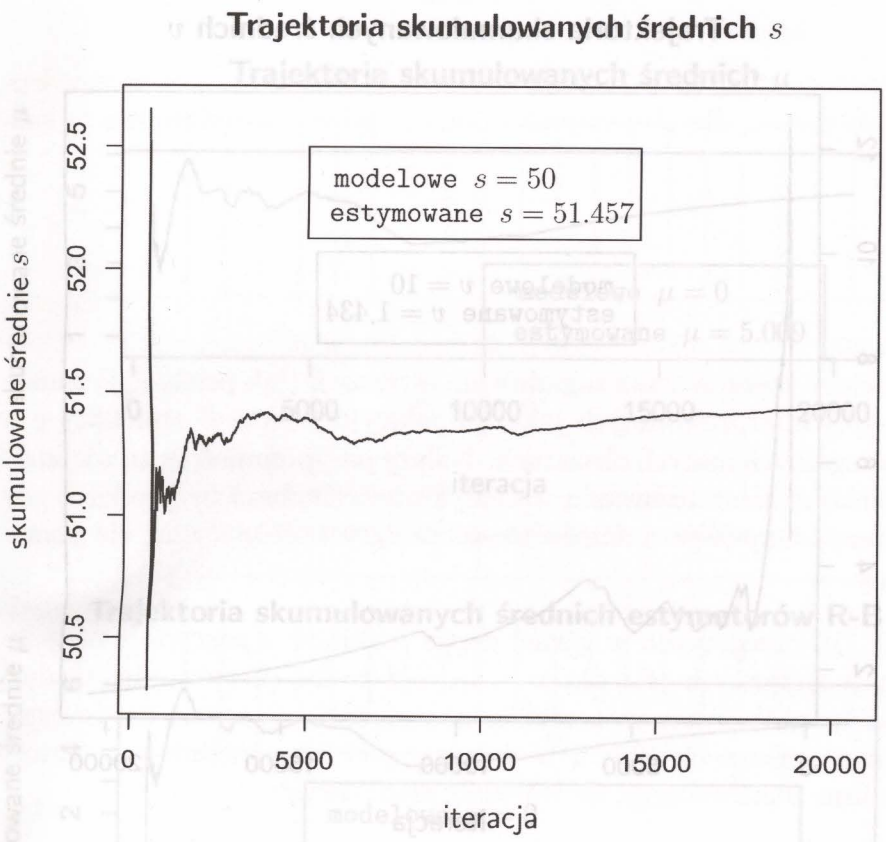


Na wykr. 16 i 17 przedstawiono ewolucję „średnich wzdłuż trajektorii" dla parametrów s i v, czyli estymatorów $\bar{s}_{b,n}$ i $\bar{v}_{b,n}$ w zależności od długości n uśrednianej trajektorii. Podobnie jak poprzednio, obliczano średnie na podstawie trajektorii:

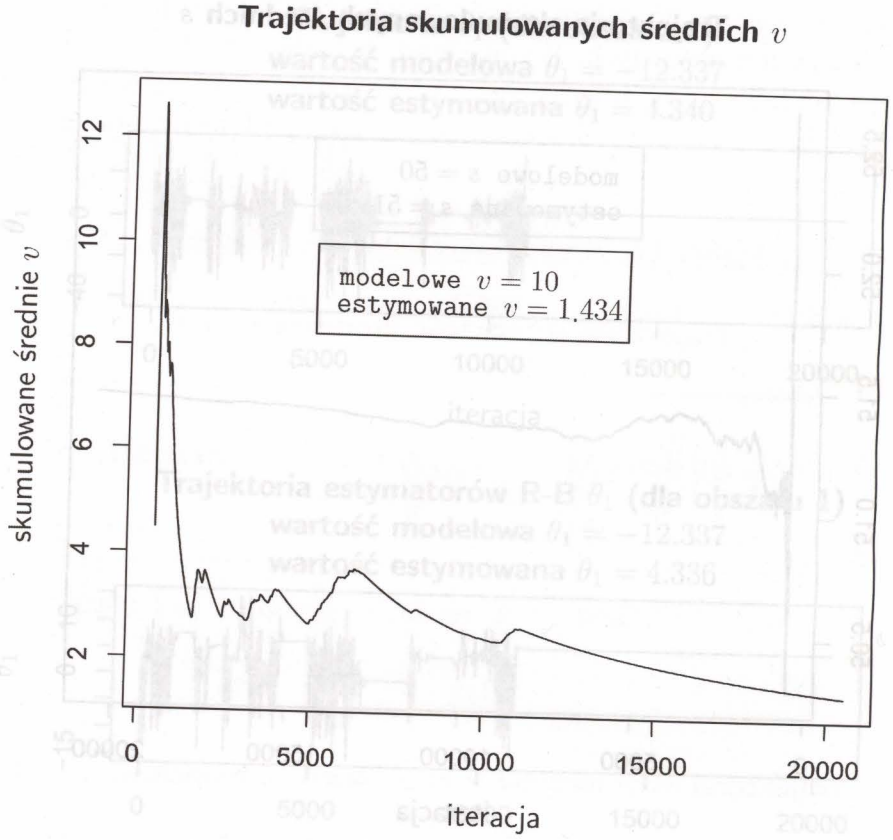
$$X_{b+1}, \ldots, X_{b+n},$$

gdzie $b = 500$ oznacza długość odrzuconego początkowego odcinka trajektorii, zaś n zmienia się od 1 do 20000.

Wykr. 16. Średnie wzdłuż trajektorii dla zmiennej s

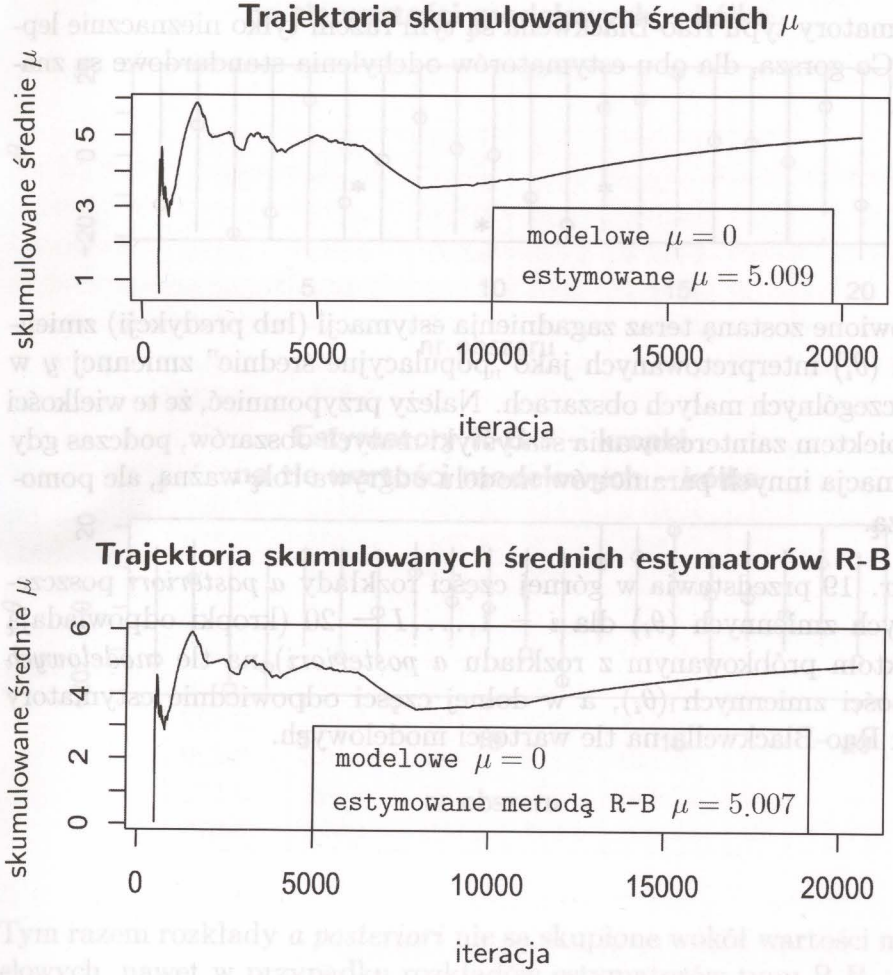


Wykr. 17. Średnie wzdłuż trajektorii dla zmiennej v



Średnie wzdłuż trajektorii dla zmiennej s całkiem dobrze aproksymują średnią *a posteriori*, choć ich stabilizacja trwa znacznie dłużej (ok. 8000 iteracji) niż w pierwszym doświadczeniu. Natomiast dla zmiennej v średnie wzdłuż trajektorii źle przybliżają średnią *a posteriori*. Wiąże się to z dużymi i stosunkowo rzadkimi wyskokami w górę zmiennej v. Nawet dla trajektorii mającej 20500 iteracji nie nastąpiła widoczna stabilizacja średniej v wzdłuż trajektorii. Nie pozostaje to bez wpływu na średnie wzdłuż trajektorii dla zmiennej μ co widać na wykr. 18.

Wykr. 18. Średnie wzdłuż trajektorii dla zmiennej μ oraz ich wersja typu Rao-Blackwella



Liczby potwierdzają widoczną na tych rwykresach prawidłowość:

- $\bar{\mu}_{500,20000} = 5.009$ – estymator „średnia wzdłuż trajektorii',
- $\bar{\mu}^{\mathrm{R-B}}_{500,20000} = 5.007$ – estymator Rao-Blackwella,
- $\frac{1}{1999}\sum_{k=501}^{20500}(\mu_k - \bar{\mu}_{500,20000})^2 = \mathbf{2.85}$ – empiryczne odchylenie standardowe dla (zależnej) „próbki" $\mu_{501}, \ldots, \mu_{20500}$,

- $\frac{1}{1999}\sum_{k=501}^{20500}(\mu_k^{\mathrm{R-B}} - \bar{\mu}_{500,20000}^{\mathrm{R-B}})^2 = \mathbf{2.734}$ – empiryczne odchylenie standardowe dla $\mu_{501}^{\mathrm{R-B}}, \ldots, \mu_{20500}^{\mathrm{R-B}}$.

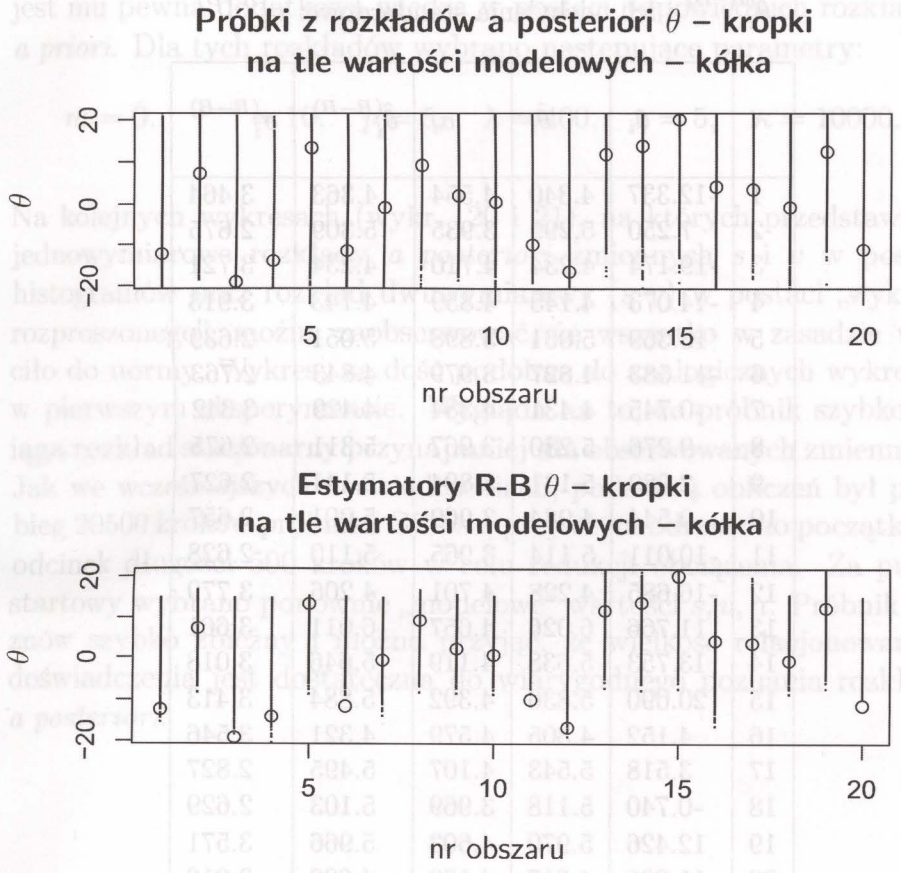
Estymatory typu Rao-Blackwella są tym razem tylko nieznacznie lepsze. Co gorsza, dla obu estymatorów odchylenia standardowe są znaczne.

* *
*

Omówione zostaną teraz zagadnienia estymacji (lub predykcji) zmiennych (θ_i) interpretowanych jako „populacyjne średnie" zmiennej y w poszczególnych małych obszarach. Należy przypomnieć, że te wielkości są obiektem zainteresowania statystyki małych obszarów, podczas gdy estymacja innych parametrów modelu odgrywa rolę ważną, ale pomocniczą.

Wykr. 19 przedstawia w górnej części rozkłady *a posteriori* poszczególnych zmiennych (θ_i) dla $i = 1, \ldots, I = 20$ (kropki odpowiadają punktom próbkowanym z rozkładu *a posteriori*) na tle *modelowych* wartości zmiennych (θ_i), a w dolnej części odpowiednie estymatory typu Rao-Blackwella na tle wartości modelowych.

Wykr. 19. Rozkłady a posteriori i modelowe wartości zmiennych θ_i



Tym razem rozkłady *a posteriori* nie są skupione wokół wartości modelowych, nawet w przypadku rozkładów estymatorów typu R-B.

Potwierdza ten efekt również poniższa tabl. 2, w której przedstawiono wartości estymatorów liczonych wzdłuż trajektorii dla średnich *a posteriori* zmiennych (θ_i). Zwraca uwagę fakt, iż wartości te odbiegają od wartości modelowych, a ich odchylenia standardowe są znaczne. Wartości te obliczono w identyczny sposób jak analogiczne wielkości dla zmiennej μ, podane wcześniej.

Tabl. 2. Mały obszar: i, wartości modelowe: θ_i, ich estymatory: $\hat{\theta}_i$ i odchylenia standardowe rozkładów *a posteriori*: σ_i, estymatory R-B: $\hat{\theta}_i^{(R-B)}$ i ich odchylenia standardowe: $\sigma_i^{(R-B)}$

i	θ_i	$\hat{\theta}_i$	σ_i	$\hat{\theta}_i^{(R-B)}$	$\sigma_i^{(R-B)}$
1	-12.337	4.340	4.554	4.363	3.464
2	7.250	5.299	3.935	5.309	2.675
3	-19.174	4.234	4.710	4.234	3.721
4	-14.073	4.145	4.859	4.143	3.913
5	13.369	5.061	3.898	5.051	2.639
6	-11.583	4.827	3.979	4.843	2.763
7	-0.745	4.431	4.384	4.429	3.342
8	9.276	5.280	3.967	5.311	2.675
9	2.089	5.131	3.894	5.141	2.627
10	0.544	4.984	3.969	5.001	2.657
11	-10.011	5.114	3.965	5.110	2.628
12	-16.685	4.228	4.701	4.206	3.779
13	11.766	6.026	4.657	6.011	3.661
14	13.753	5.638	4.119	5.646	3.018
15	20.090	5.836	4.392	5.884	3.413
16	4.152	4.305	4.579	4.321	3.546
17	3.518	5.543	4.107	5.495	2.827
18	-0.740	5.118	3.969	5.103	2.629
19	12.426	5.970	4.603	5.966	3.571
20	-11.306	4.617	4.188	4.628	3.018

2.3. Niezróżnicowane małe obszary, informacyjny rozkład *a priori*

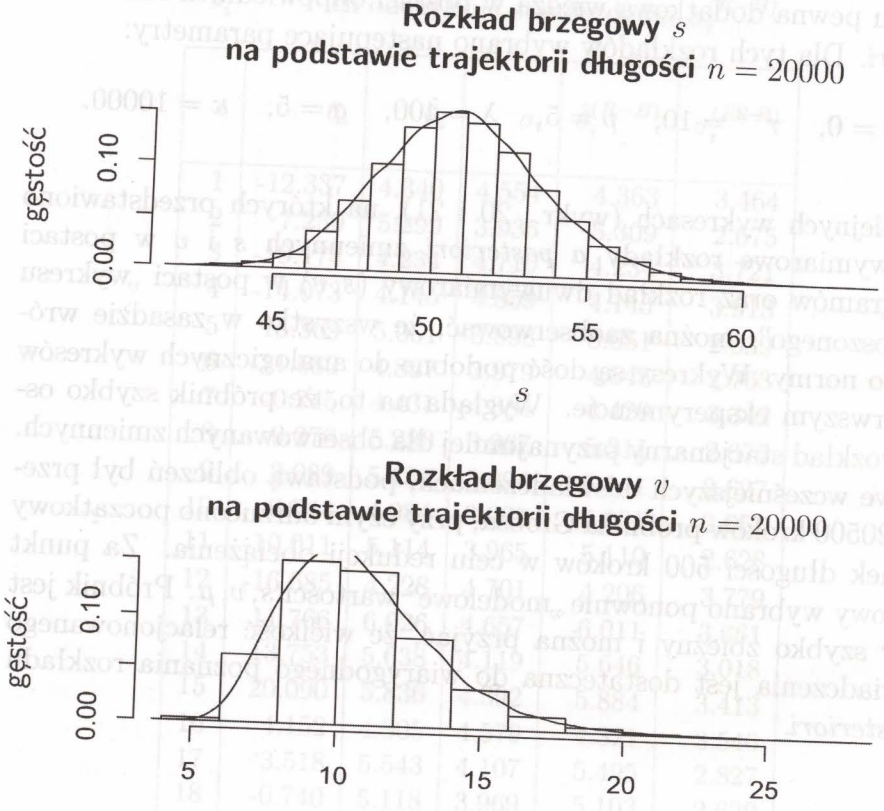
Trzeci eksperyment dotyczy, tak jak i drugi, sytuacji, gdy błąd losowy ma znacznie wiekszą zmienność niż zmienność między małymi obszarami. Ponownie jako parametry modelowe przyjęto $\mu = 0$, $s = 50$ i $v = 10$. Jak można było zaobserwować podczas analizy drugiego eksperymentu, same dane przy nieinformacyjnych rozkładach *a priori*, w tej sytuacji nie zapewniają poprawnego działania próbnika. Pomysł poprawy działania próbnika polegał na wzięciu takich parametrów roz-

kładów *a priori*, które są bliższe wartościom modelowym. Innymi słowy, próbnik nie tylko korzysta z obserwacji (y-ów) ale dostarczana jest mu pewna dodatkowa wiedza w postaci odpowiednich rozkładów *a priori*. Dla tych rozkładów wybrano następujące parametry:

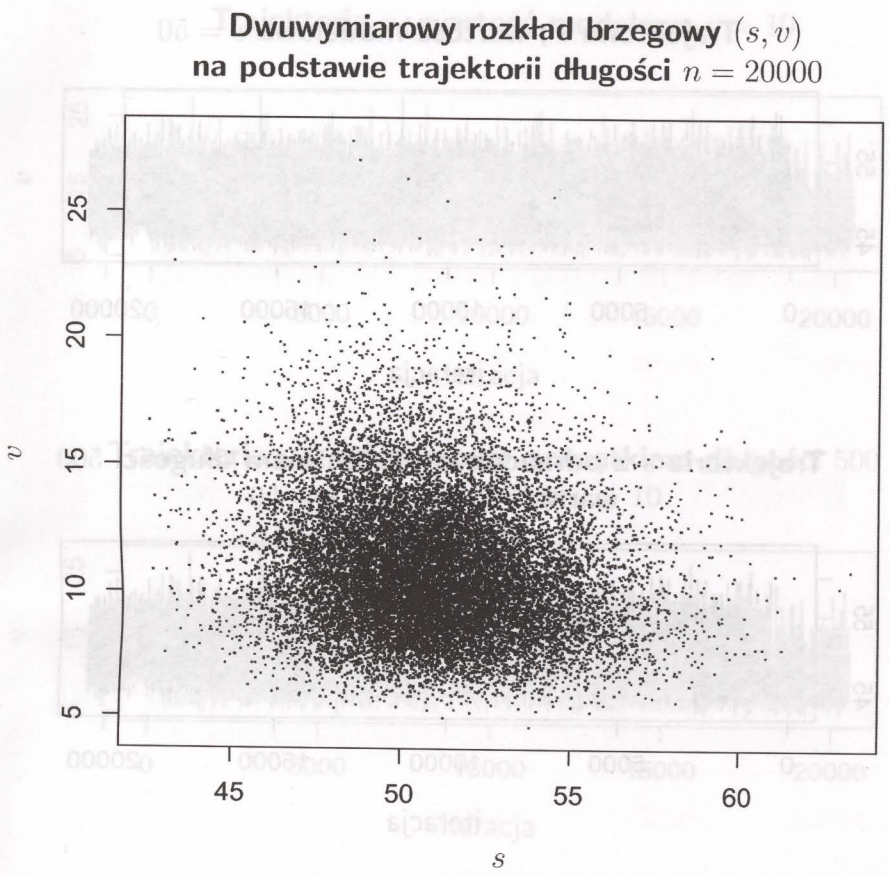
$$m = 0, \quad \tau^{-2} = 10, \quad p = 5, \quad \lambda = 400, \quad q = 5, \quad \kappa = 10000.$$

Na kolejnych wykresach (wykr. 20 i 21), na których przedstawiono jednowymiarowe rozkłady *a posteriori* zmiennych s i v w postaci histogramów oraz rozkład dwuwymiarowy (s, v) w postaci „wykresu rozproszonego", można zaobserwować, że wszystko w zasadzie wróciło do normy. Wykresy są dość podobne do analogicznych wykresów w pierwszym eksperymencie. Wygląda na to, że próbnik szybko osiąga rozkład stacjonarny przynajmniej dla obserwowanych zmiennych. Jak we wcześniejszych doświadczeniach, podstawą obliczeń był przebieg 20500 kroków próbnika Gibbsa, przy czym odrzucono początkowy odcinek długości 500 kroków w celu redukcji obciążenia. Za punkt startowy wybrano ponownie „modelowe" wartości s, v, μ. Próbnik jest znów szybko zbieżny i można przyjąć, że wielkość relacjonowanego doświadczenia jest dostateczna do wiarygodnego poznania rozkładu *a posteriori*.

Wykr. 20. Jednowymiarowe rozkłady brzegowe zmiennych s i v

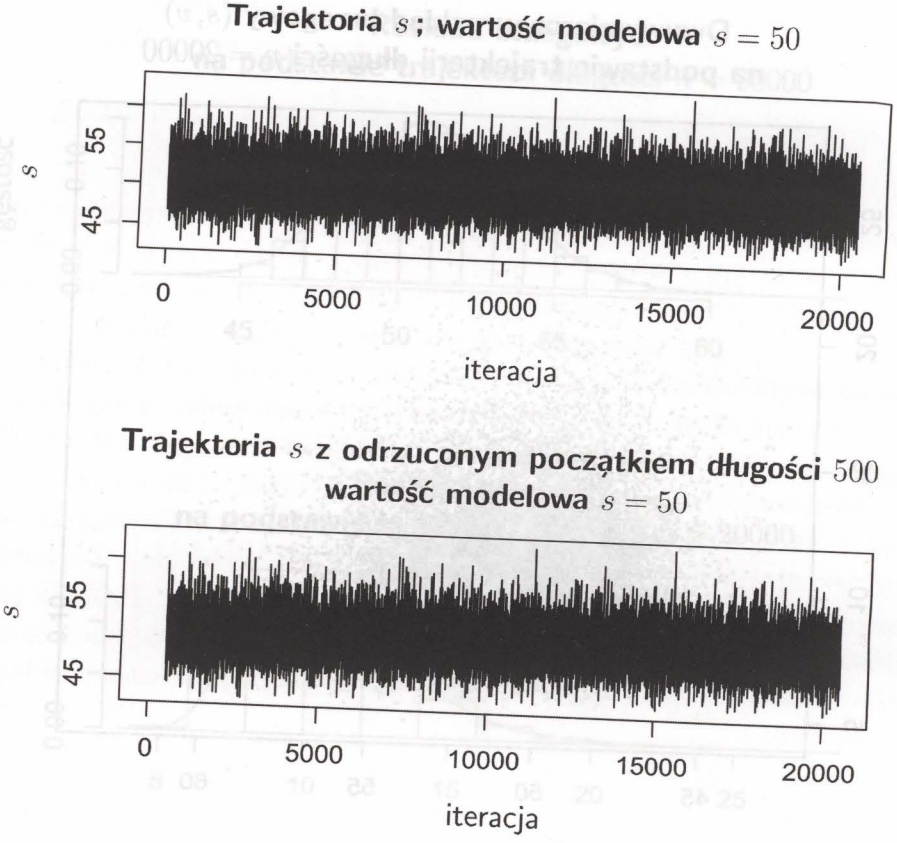


Wykr. 21. Dwuwymiarowy rozkład brzegowy pary zmiennych (s, v)

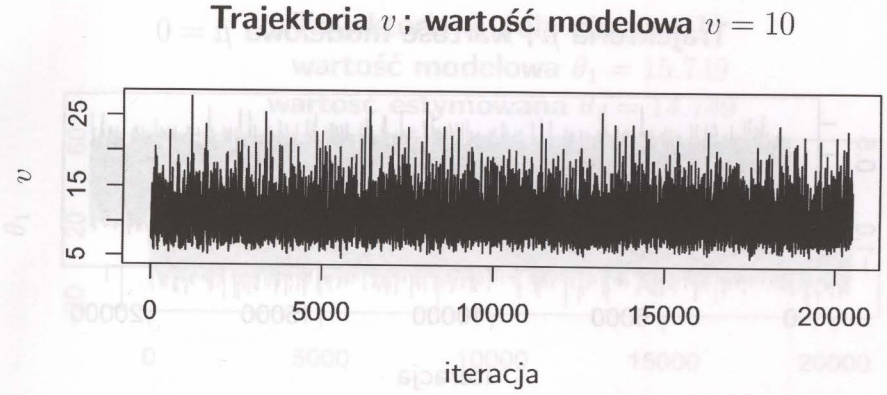


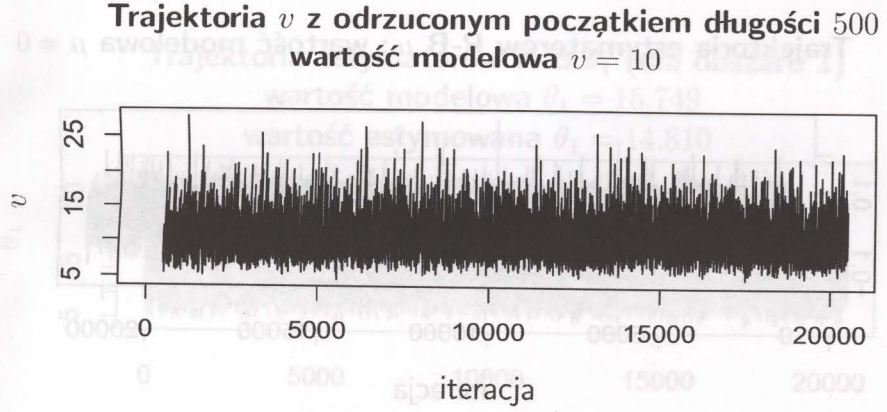
Wykr. 22 i 23 obrazują przebieg trajektorii próbnika i przedstawiają ewolucję zmiennych s_k i v_k.

Wykr. 22. Trajektoria próbnika dla zmiennej s



Wykr. 23. Trajektoria próbnika dla zmiennej v

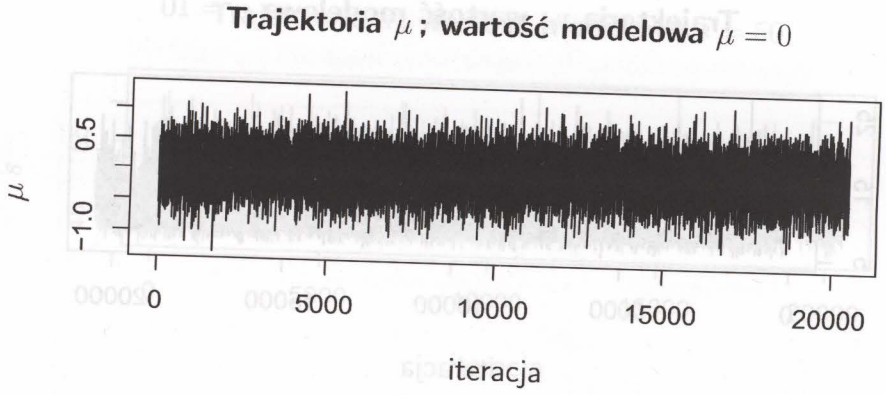


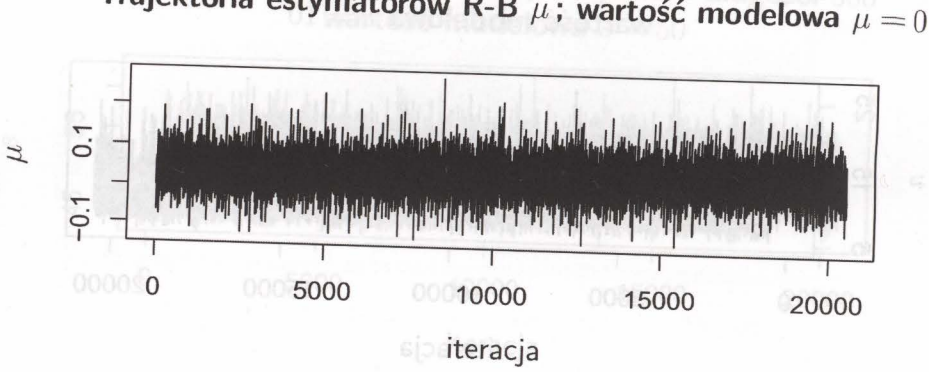


Ponownie przebiegi trajektorii wskazują, że dla obu zmiennych próbnik stosunkowo szybko osiąga rozkład stacjonarny.

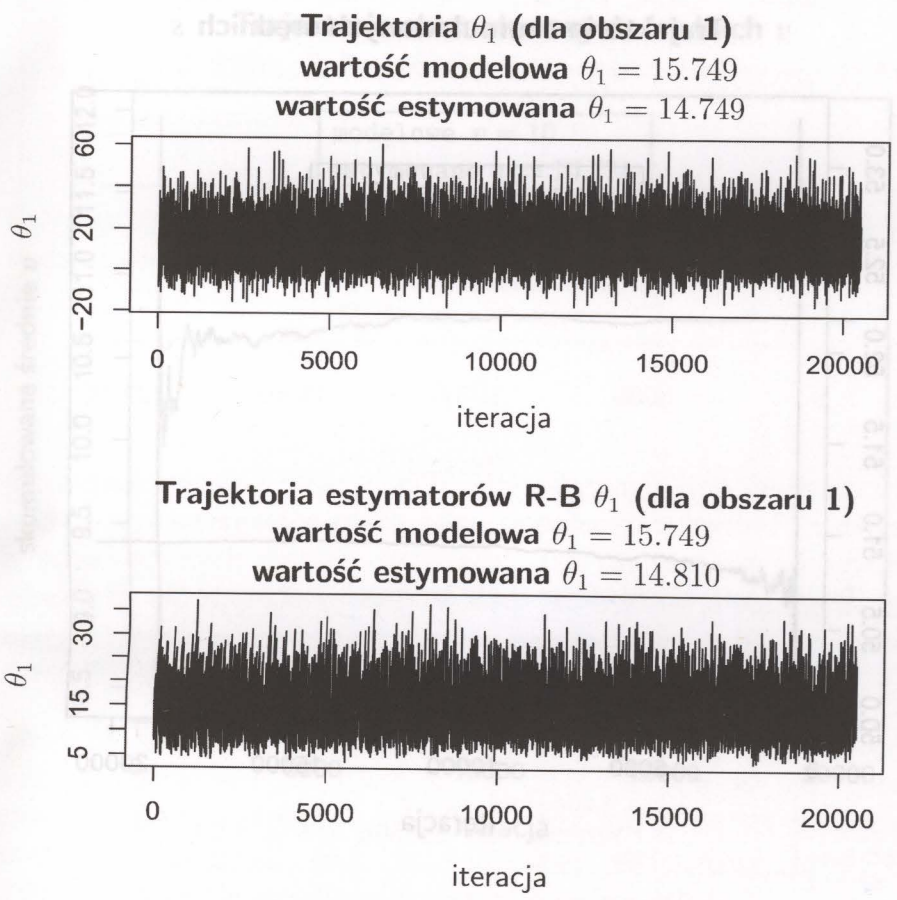
Trajektorie dla zmiennych μ i (θ_i) zobrazowane na wykr. 24 i 25 (przy czym na wykr. 25 wybrano mały obszar $i = 1$) wyglądają podobnie.

Wykr. 24. Trajektoria próbnika dla zmiennej μ





Wykr. 25. Trajektoria próbnika dla zmiennej θ_1

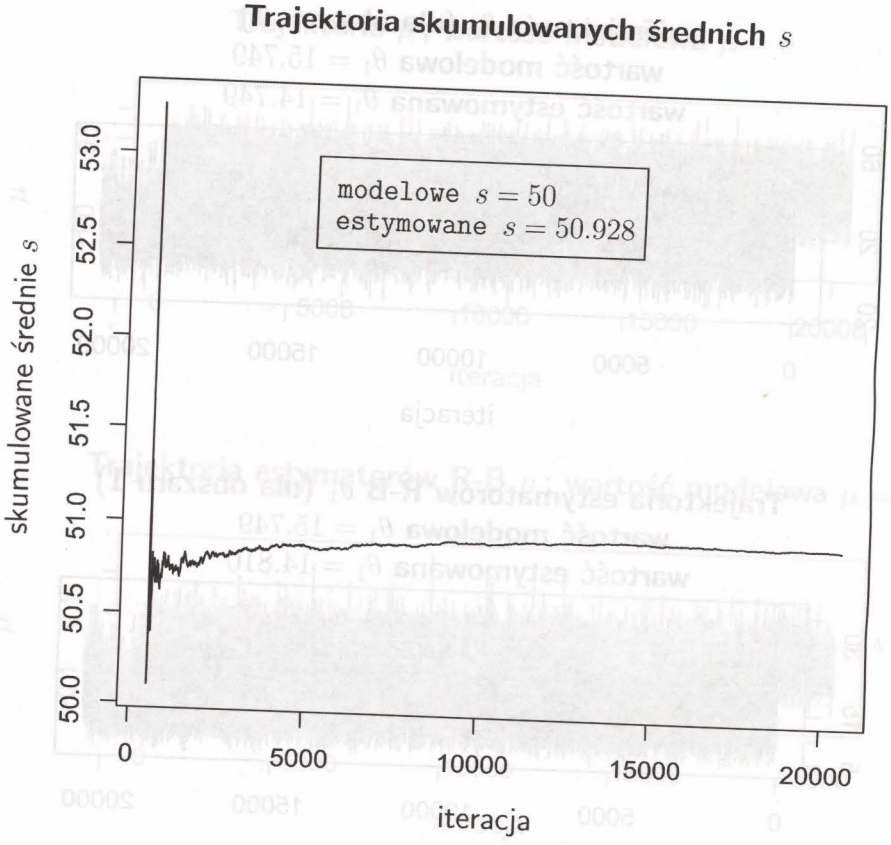


Na kolejnych wykresach (wykr. 26 i 27) przedstawiono ewolucję „średnich wzdłuż trajektorii", czyli estymatorów $\bar{s}_{b,n}$ i $\bar{v}_{b,n}$ w zależności od długości n uśrednianej trajektorii. Średnie obliczano, jak poprzednio, na podstawie trajektorii:

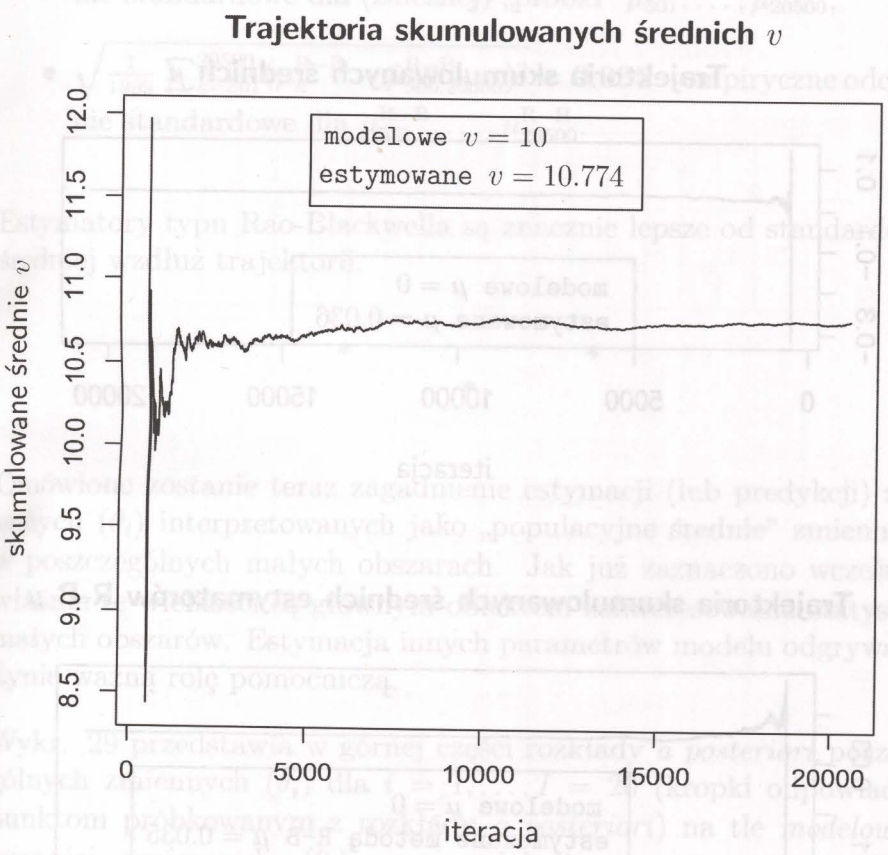
$$X_{b+1}, \ldots, X_{b+n},$$

gdzie $b = 500$ oznacza długość odrzuconego początkowego odcinka trajektorii, zaś n zmienia się od 1 do 20000.

Wykr. 26. Średnie wzdłuż trajektorii dla zmiennej s



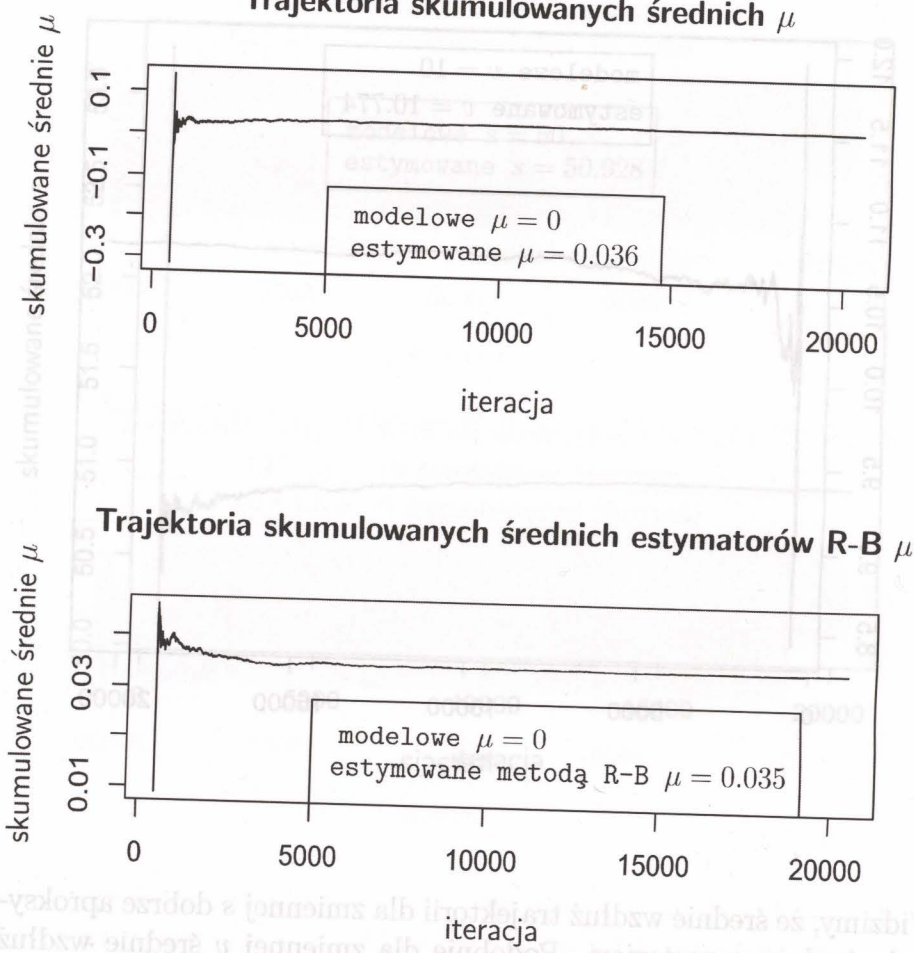
Wykr. 27. Średnie wzdłuż trajektorii dla zmiennej v



Widzimy, że średnie wzdłuż trajektorii dla zmiennej s dobrze aproksymują średnią *a posteriori*. Podobnie dla zmiennej v średnie wzdłuż trajektorii nieźle przybliżają średnią *a posteriori*. Problem, który pojawił się w drugim eksperymencie został zażegnany (jak się okaże dalej jedynie częściowo) poprzez wzięcie informacyjnego rozkładu *a priori*.

Średnie wzdłuż trajektorii dla zmiennej μ też stabilizują się stosunkowo szybko i dobrze przybliżają odpowiednie średnie *a posteriori*, co znakomicie obrazuje wykr. 28.

Wykr. 28. Średnie wzdłuż trajektorii dla zmiennej μ oraz ich wersja typu Rao-Blackwella



Liczbowo, obserwowana na poprzednich wykresach prawidłowość wygląda następująco:

- $\bar{\mu}_{500,20000} = 0.036$ – estymator „średnia wzdłuż trajektorii",
- $\bar{\mu}^{\mathrm{R-B}}_{500,20000} = 0.035$ – estymator Rao-Blackwella,

- $\sqrt{\frac{1}{1999}\sum_{k=501}^{20500}(\mu_k - \bar{\mu}_{500,20000})^2} = \mathbf{0.315}$ – empiryczne odchylenie standardowe dla (zależnej) „próbki" $\mu_{501}, \ldots, \mu_{20500}$,
- $\sqrt{\frac{1}{1999}\sum_{k=501}^{20500}(\mu_k^{\mathrm{R-B}} - \bar{\mu}_{500,20000}^{\mathrm{R-B}})^2} = \mathbf{0.035}$ – empiryczne odchylenie standardowe dla $\mu_{501}^{\mathrm{R-B}}, \ldots, \mu_{20500}^{\mathrm{R-B}}$.

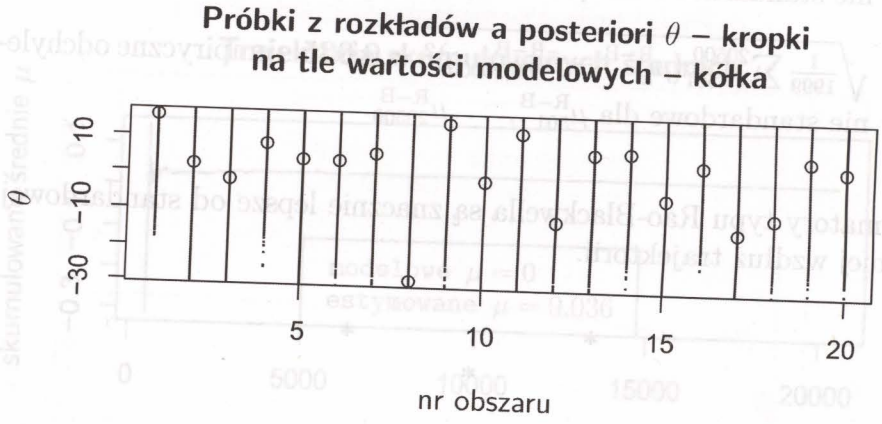
Estymatory typu Rao-Blackwella są znacznie lepsze od standardowej średniej wzdłuż trajektorii.

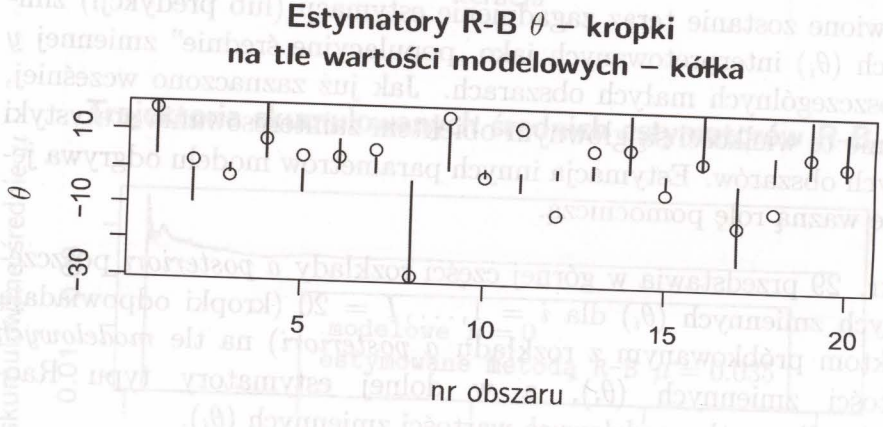
* * *

Omówione zostanie teraz zagadnienie estymacji (lub predykcji) zmiennych (θ_i) interpretowanych jako „populacyjne średnie" zmiennej y w poszczególnych małych obszarach. Jak już zaznaczono wcześniej, właśnie te wielkości są głównym obiektem zainteresowania statystyki małych obszarów. Estymacja innych parametrów modelu odgrywa jedynie ważną rolę pomocniczą.

Wykr. 29 przedstawia w górnej części rozkłady *a posteriori* poszczególnych zmiennych (θ_i) dla $i = 1, \ldots, I = 20$ (kropki odpowiadają punktom próbkowanym z rozkładu *a posteriori*) na tle *modelowych* wartości zmiennych (θ_i), a w dolnej estymatory typu Rao-Blackwella na tle *modelowych* wartości zmiennych (θ_i).

Wykr. 29. Rozkłady a posteriori i modelowe wartości zmiennych θ_i





Widzimy, w szczególności w przypadku rozkładów estymatorów typu Rao-Blackwella, że wartości modelowe mają większe zróżnicowanie niż rozkłady *a posteriori* w poszczególnych małych obszarach. Innymi słowy, zastosowanie informacyjnego rozkładu *a priori*, choć poprawiło estymację zmiennych pomocniczych, to nie wystarczyło do uzyskania dobrych estymatorów *a posteriori* wielkości θ_i w poszczególnych małych obszarach. Należy przypuszczać, że przeważył efekt małego odchylenia standardowego v w stosunku do s. Efekt ten obrazuje też tabl. 3. Obserwuje się zbyt małe zróżnicowanie estymatorów w małych obszarach. Tym niemniej, tak jak w przypadku pierwszego eksperymentu, zauważyć można znacząco mniejsze

odchylenie standardowe dla estymatorów typu Rao-Blackwella, w stosunku do zwykłych estymatorów uzyskiwanych z próbnika Gibbsa.

Tabl. 3. Mały obszar: i, wartości modelowe: θ_i, ich estymatory: $\hat{\theta}_i$ i odchylenia standardowe rozkładów *a posteriori*: σ_i, estymatory R-B: $\hat{\theta}_i^{(R-B)}$ i ich odchylenia standardowe: $\sigma_i^{(R-B)}$

i	θ_i	$\hat{\theta}_i$	σ_i	$\hat{\theta}_i^{(R-B)}$	$\sigma_i^{(R-B)}$
1	15.749	14.749	10.199	14.810	4.957
2	2.203	-3.813	8.930	-3.829	1.316
3	-1.825	0.1585	8.903	0.130	0.222
4	8.193	13.129	9.956	13.044	4.365
5	3.999	-2.373	8.897	-2.372	0.839
6	3.858	3.423	8.956	3.409	1.150
7	6.022	1.576	8.931	1.510	0.539
8	-28.868	-12.686	9.784	-12.748	4.294
9	14.859	5.253	9.051	5.297	1.776
10	-0.972	0.058	8.948	0.256	0.231
11	12.561	-1.903	8.935	-1.808	0.658
12	-11.291	0.600	8.854	0.596	0.288
13	7.179	1.270	8.918	1.214	0.451
14	7.737	6.568	9.205	6.643	2.224
15	-4.615	-1.174	8.994	-1.092	0.439
16	4.481	11.2623	9.731	11.245	3.763
18	-9.073	2.598	8.916	2.654	0.903
19	6.598	6.551	9.162	6.624	2.218
20	4.154	5.380	9.084	5.437	1.822

Rozdział 3

Błąd średniokwadratowy i obciążenie

Przeprowadzono 200 powtórzeń przebiegu próbnika dla tego samego zestawu parametrów „modelowych" co w pierwszej części rozdziału 2, to znaczy:

$$\mu = 0, \quad s = 10, \quad v = 10, \quad n_i = 10 \ (i = 1, \ldots, I = 20).$$

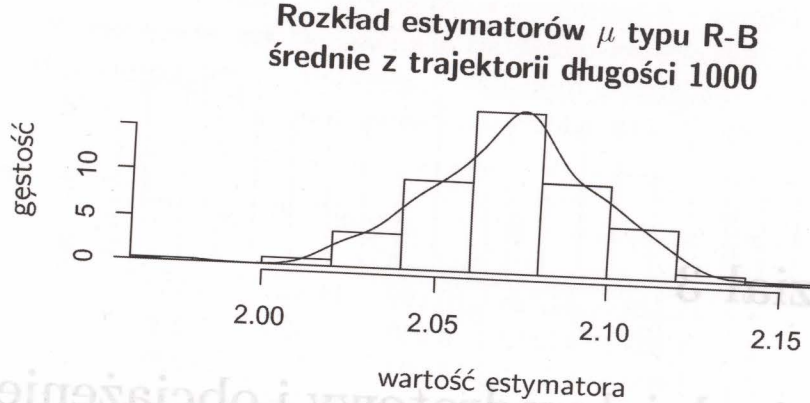
Podobnie jak poprzednio, zastosowano niewłaściwy, nieinformacyjny rozkład *a priori* z następującymi parametrami:

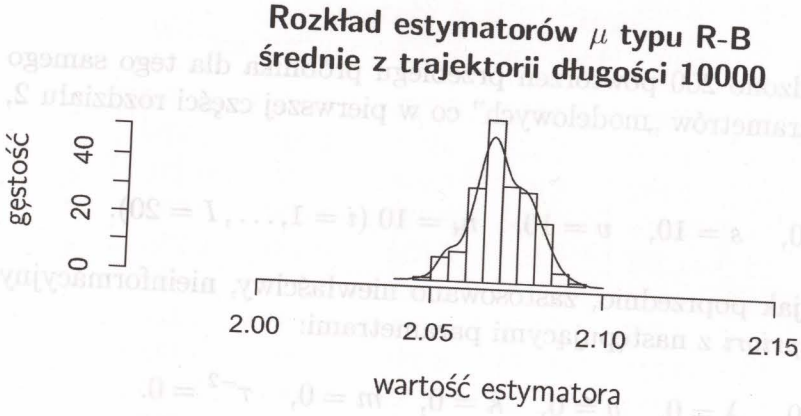
$$p = 0, \quad \lambda = 0, \quad q = 0, \quad \kappa = 0, \quad m = 0, \quad \tau^{-2} = 0.$$

W każdym z 200 powtórzeń, badano przebieg 10500 kroków próbnika Gibbsa, przy czym odrzucano początkowy odcinek długości $b = 500$ kroków w celu redukcji obciążenia. Za punkt startowy wybierano każdorazowo „modelowe" wartości s, v, μ. Uwaga skupiona była na badaniu estymatorów parametrów μ i (θ_i) typu Rao-Blackwella.

Wykr. 30 przedstawia histogram estymatorów parametru μ dla dwóch długości trajektorii 1000 i 10000 kroków. Wyraźnie widoczne jest, że dla długiej trajektorii rozkład średniej z trajektorii znacznie bardziej koncentruje się wokół prawdziwej wartości średniej *a posteriori* niż dla trajektorii krótszej. Ten efekt będzie jeszcze lepiej zobrazowany w dalszej części rozdziału za pomocą sekwencji wykresów pudełkowych.

Wykr. 30. Rozkłady estymatorów parametru μ dla dwóch długości trajektorii



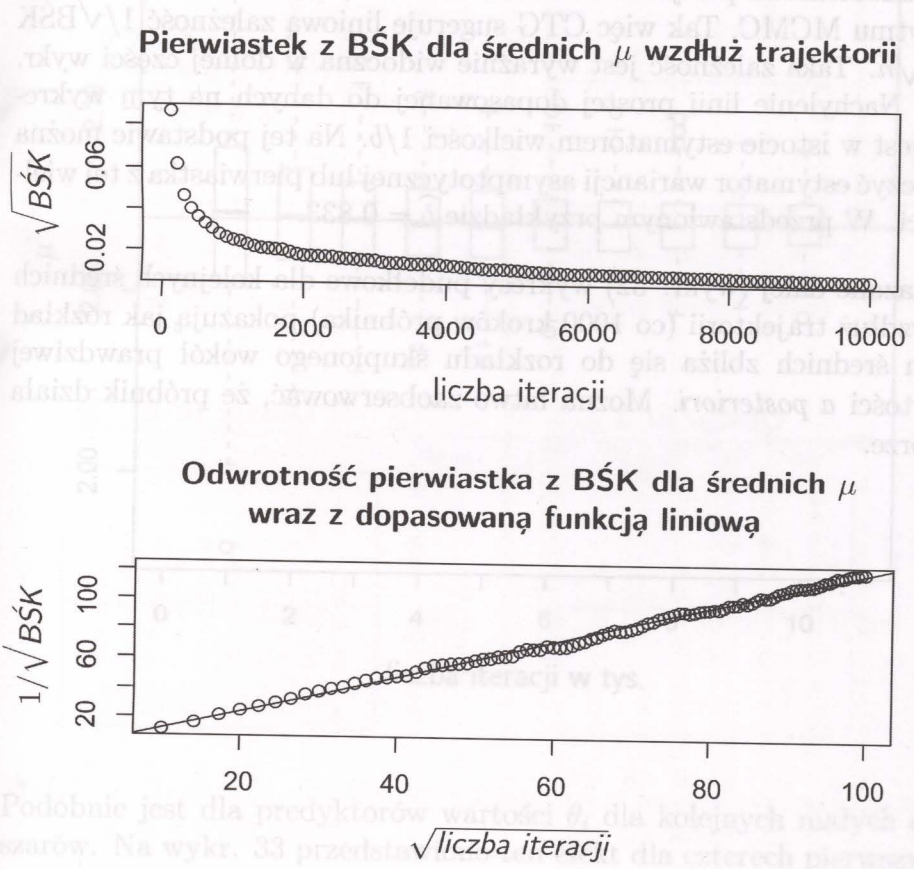


Następnie przedstawiona zostanie zależność błędów średniokwadratowych estymatorów od długości trajektorii. Należy podkreślić, że głównym przedmiotem zainteresowania jest tu dokładność obliczeń charakterystyk *a posteriori* metodą MCMC przy pomocy próbnika Gibbsa. Innymi słowy, badana jest jakość aproksymacji wartości średniej *a posteriori* $\mu^{\text{post}} := \mathbb{E}(\mu|y)$ przez obliczone w kolejnych krokach próbnika wartości $\bar{\mu}_{b,n}^{\text{R-B}}$ (a *nie* jakość aproksymacji „modelowej wielkości" μ przez estymator bayesowski μ^{post}). Wobec tego interesującą charakterystyką jest błąd średniokwadratowy zdefiniowany następująco:

$$\text{BŚK} := \mathbb{E}\left(\bar{\mu}_{b,n}^{\text{R-B}} - \mu^{\text{post}}\right)^2.$$

Oczywiście, wartość oczekiwana we wzorze na błąd średniokwadratowy jest obliczana metodą Monte Carlo na podstawie 200 niezależnych powtórzeń doświadczenia.

Wykr. 31. BŚK estymatorów parametru μ w zależności od długości trajektorii



Jak widać z pierwszego z pierwszego wykresu (górna część wykr. 31), błąd średniokwadratowy zmierza dość szybko do zera wraz z liczbą iteracji próbnika (każde kółko na wykresie odpowiada błędowi średniokwadratowemu obliczonemu na podstawie 200 estymatorów średniej *a posteriori* μ). Wyniki symulacji wykazują zaskakująco dużą zgodność z asymptotycznymi twierdzeniami teorii łańcuchów Markowa.

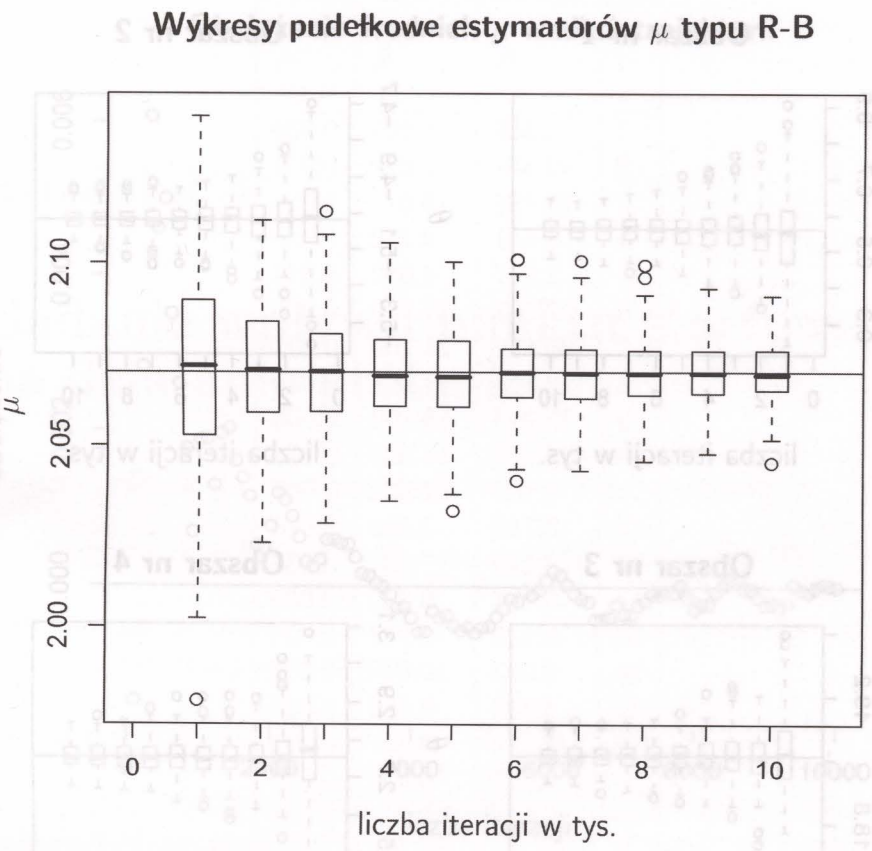
Centralne Twierdzenie Graniczne mówi, że

$$\sqrt{n}\left(\bar{\mu}_{b,n}^{\mathrm{R-B}} - \mu^{\mathrm{post}}\right) \to \mathcal{N}(0, b^2),$$

gdzie b^2 jest tzw. wariancją asymptotyczną i zależy w skomplikowany sposób nie tylko od rozkładu docelowego (*a posteriori*) ale i od prawdopodobieństwa przejścia łańcucha Markowa, czyli od konstrukcji algorytmu MCMC. Tak więc CTG sugeruje liniową zależność $1/\sqrt{\mathrm{BŚK}}$ od $\sqrt{n}$. Taka zależność jest wyraźnie widoczna w dolnej części wykr. 31. Nachylenie linii prostej dopasowanej do danych na tym wykresie jest w istocie estymatorem wielkości $1/b$. Na tej podstawie można obliczyć estymator wariancji asymptotycznej lub pierwiastka z tej wielkości. W przedstawionym przykładzie $\widehat{b} = 0.833$.

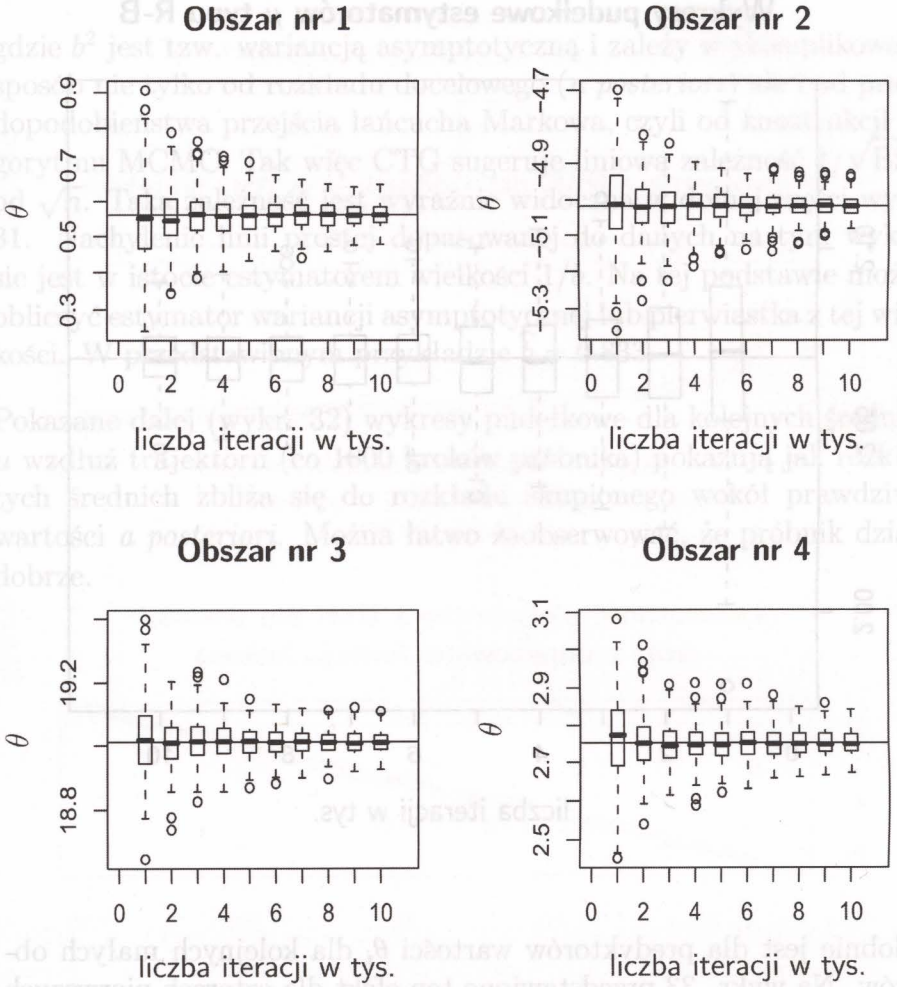
Pokazane dalej (wykr. 32) wykresy pudełkowe dla kolejnych średnich μ wzdłuż trajektorii (co 1000 kroków próbnika) pokazują jak rozkład tych średnich zbliża się do rozkładu skupionego wokół prawdziwej wartości *a posteriori*. Można łatwo zaobserwować, że próbnik działa dobrze.

Wykr. 32. Wykresy pudełkowe dla średniej parametru μ



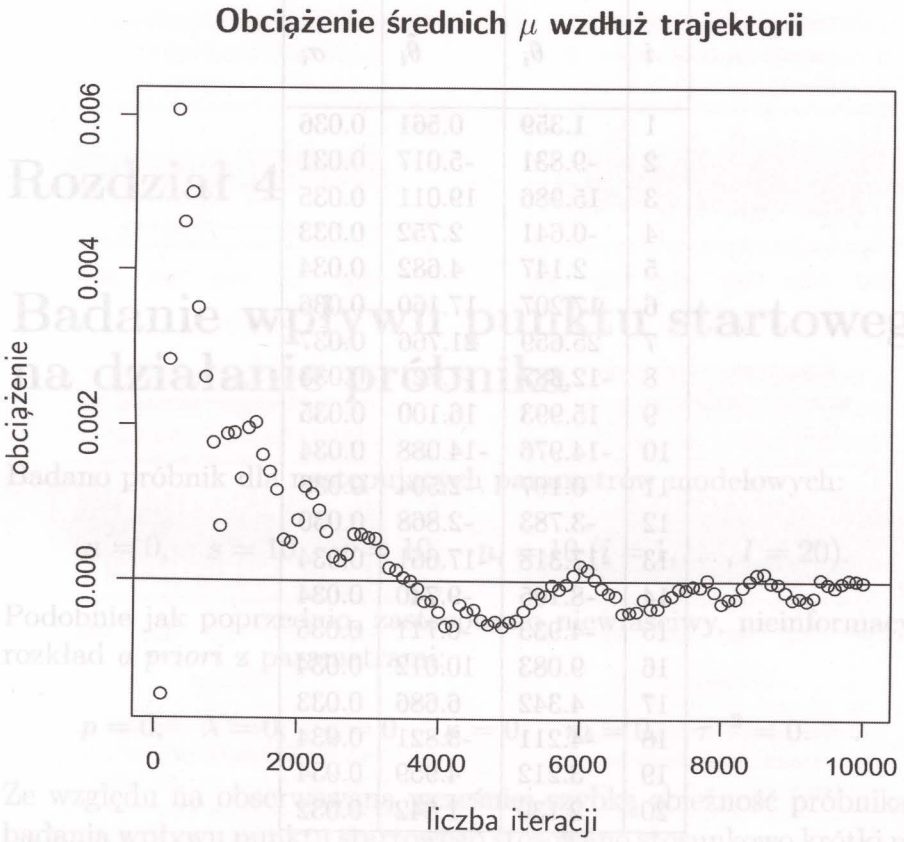
Podobnie jest dla predyktorów wartości θ_i dla kolejnych małych obszarów. Na wykr. 33 przedstawiono ten efekt dla czterech pierwszych małych obszarów.

Wykr. 33. Wykresy pudełkowe dla średnich parametru θ_i, $i = 1, 2, 3, 4$



Z kolei, na wykr. 34 zaobserwować można, że również obciążenie średniej wzdłuż trajektorii jako predyktora średniej *a posteriori* parametru μ jest niewielkie i szybko staje się wręcz znikome.

Wykr. 34. Obciążenie dla średnich wzdłuż trajektorii parametru μ



Wyniki liczbowe dla estymatora parametru μ wyglądają nastepująco:

- $\bar{\mu} = \frac{1}{200}\sum_{k=1}^{200} \hat{\mu}^{(k)}_{500,10000} = 2.07$ - średnia z 200 symulacji ze średnich wzdłuż trajektorii,
- $\sqrt{\frac{1}{199}\sum_{k=1}^{200} (\hat{\mu}^{(k)}_{500,10000} - \bar{\mu})^2} = 0.00832886$ - odchylenie standardowe średnich z 200 symulacji.

Analogiczne wartości dla parametrów małych obszarów (θ_i) podane są w tabl. 4.

Tabl. 4. Mały obszar: i, wartości modelowe: θ_i, ich estymatory: $\hat{\theta}_i$ i odchylenia standardowe rozkładów *a posteriori*: σ_i

i	θ_i	$\hat{\theta}_i$	σ_i
1	1.359	0.561	0.036
2	-9.831	-5.017	0.031
3	15.986	19.011	0.035
4	-0.641	2.752	0.033
5	2.147	4.682	0.034
6	17.207	17.160	0.036
7	25.659	21.766	0.037
8	-12.387	-7.721	0.033
9	15.993	16.100	0.035
10	-14.976	-14.088	0.034
11	0.197	-2.504	0.033
12	-3.783	-2.868	0.030
13	-17.318	-17.661	0.034
14	-8.105	-9.740	0.034
15	-4.933	-0.711	0.035
16	9.083	10.012	0.034
17	4.342	6.686	0.033
18	-4.211	-3.821	0.034
19	3.212	4.959	0.034
20	2.130	1.842	0.032

Rozdział 4

Badanie wpływu punktu startowego na działanie próbnika

Badano próbnik dla następujących parametrów modelowych:

$$\mu = 0, \quad s = 10, \quad v = 10, \quad n_i = 10 \ (i = 1, \ldots, I = 20).$$

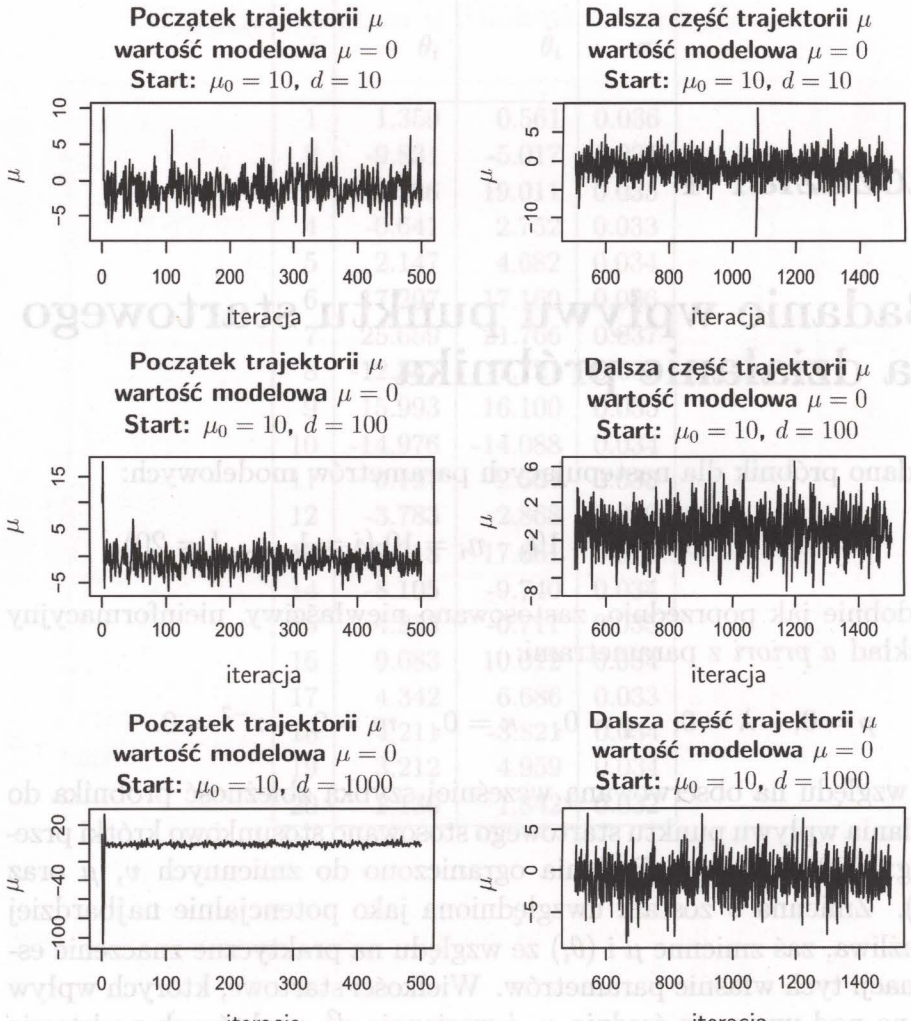
Podobnie jak poprzednio, zastosowano niewłaściwy, nieinformacyjny rozkład *a priori* z parametrami:

$$p = 0, \quad \lambda = 0, \quad q = 0, \quad \kappa = 0, \quad m = 0, \quad \tau^{-2} = 0.$$

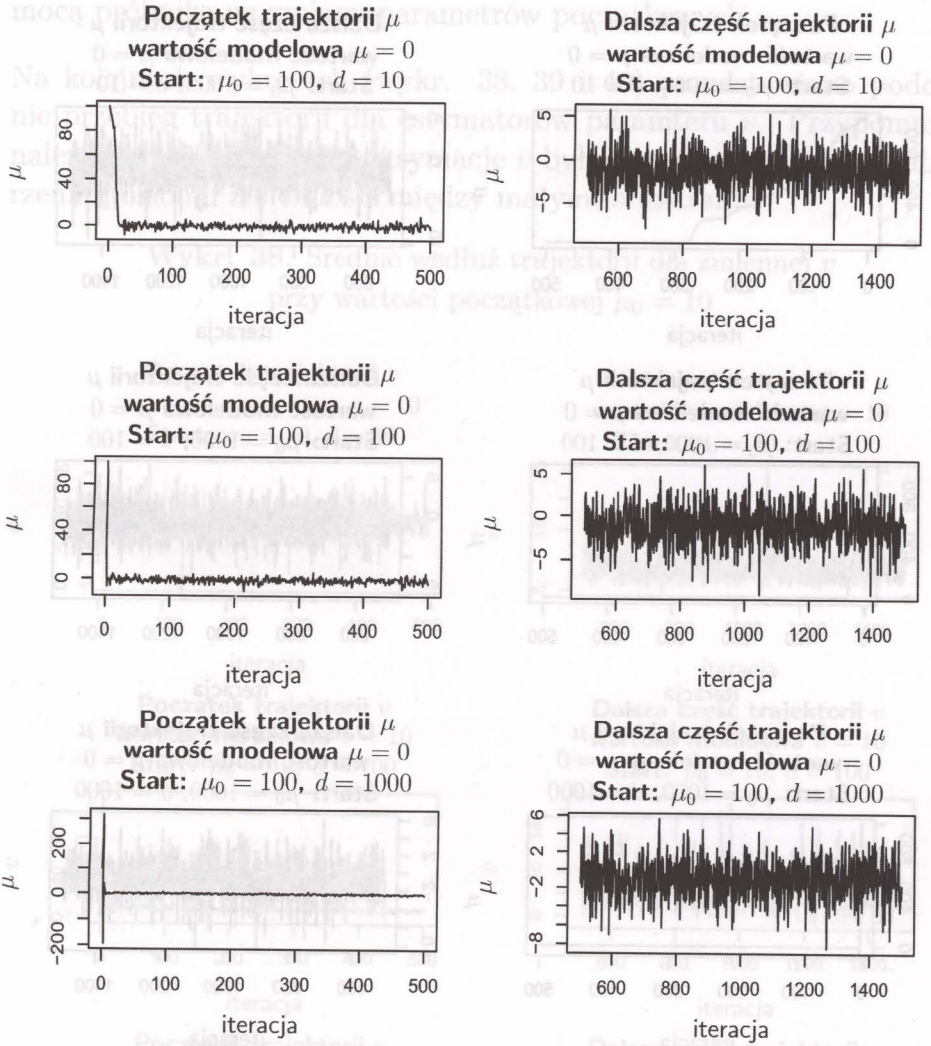
Ze względu na obserwowaną wcześniej szybką zbieżność próbnika do badania wpływu punktu startowego stosowano stosunkowo krótki przebieg: 2000 kroków. Badania ograniczono do zmiennych v, μ oraz (θ_i). Zmienna v została uwzględniona jako potencjalnie najbardziej wrażliwa, zaś zmienne μ i (θ_i) ze względu na praktyczne znaczenie estymacji tych właśnie parametrów. Wielkości startowe, których wpływ brano pod uwagę to średnia μ_0 i wariancja d^2, na których podstawie generowany jest początkowy wektor losowy (θ_i) składający się z niezależnych zmiennych losowych o tym samym rozkładzie normalnym $\mathcal{N}(\mu_0, d^2)$. Wielkość μ_0 yo również wartość początkowa zmiennej μ.

Przyjęto następujące wartości parametrów startowych: $\mu_0 = 10$, 100, 1000 oraz $d = 10$, 100, 1000. Na wykr. 35, 36 i 37 przedstawione są trajektorie typu Rao-Blackwella zmiennej μ dla wszystkich możliwych kombinacji rozważanych wartości początkowych parametrów μ_0 i d.

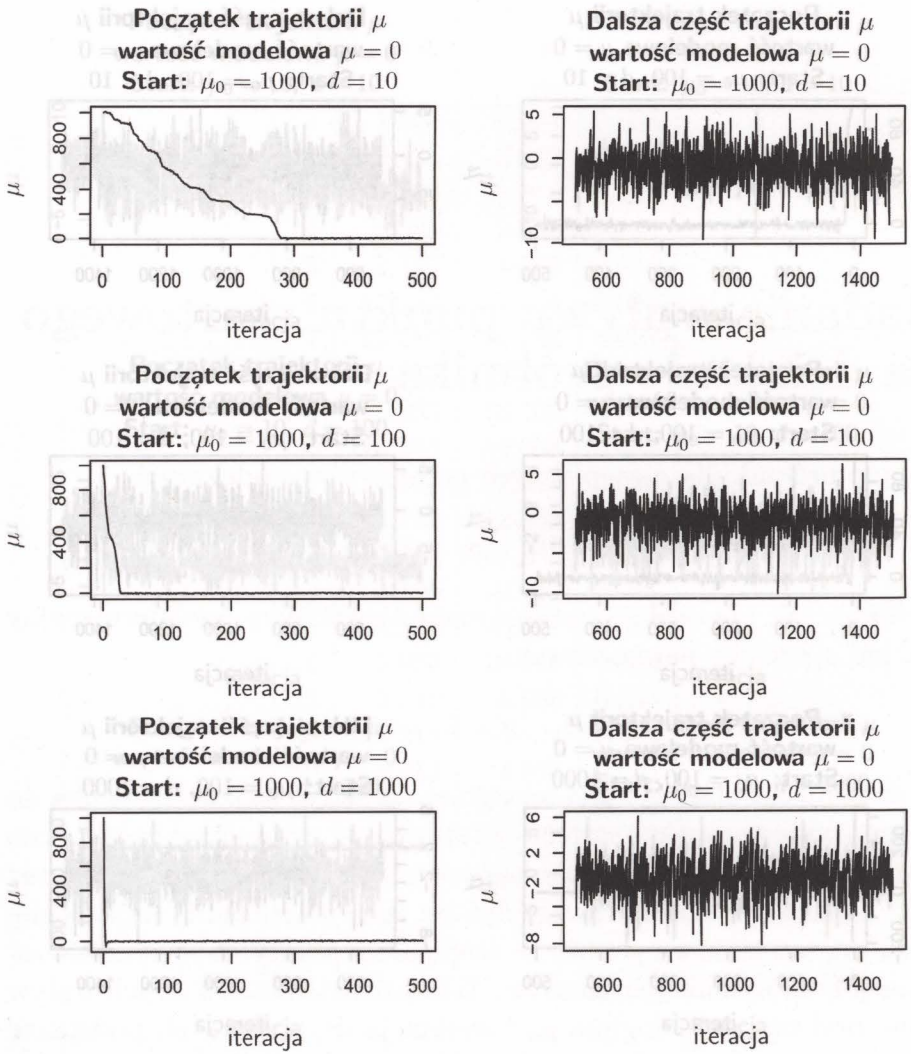
Wykr. 35. Średnie typu Rao-Blackwella wzdłuż trajektorii dla zmiennej μ przy wartości początkowej $\mu_0 = 10$



Wykr. 36. Średnie typu Rao-Blackwella wzdłuż trajektorii dla zmiennej μ przy wartości początkowej $\mu_0 = 100$



Wykr. 37. Średnie typu Rao-Blackwella wzdłuż trajektorii dla zmiennej μ przy wartości początkowej $\mu_0 = 1000$

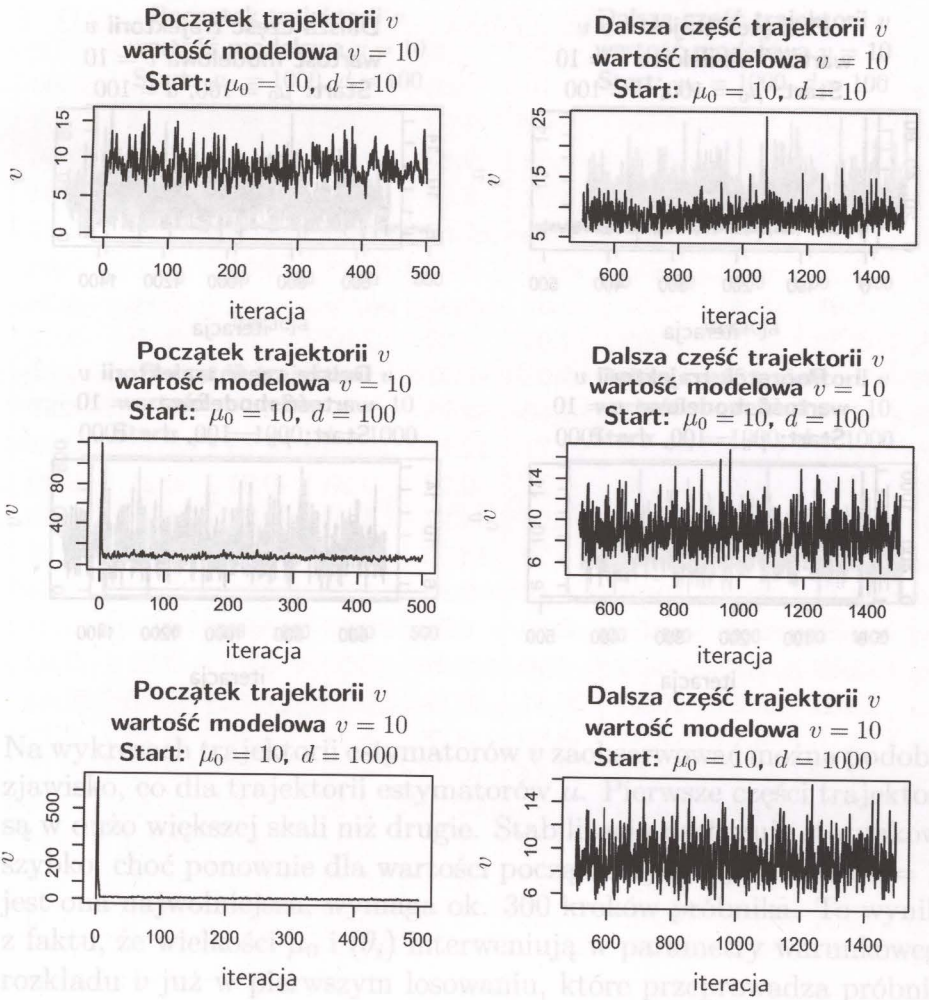


Należy zwrócić uwagę, że na wykresach tych początkowa część trajektorii każdorazowo jest przedstawiona w innej skali niż część dalsza trajektorii. Dzieje się tak ze względu na dużą różnicę między modelową wartościa $\mu = 0$, a wartością początkową $\mu = \mu_0$. Tym niemniej dość szybko osiągany jest stan stacjonarny, każdorazowo już w pierwszej części trajektorii, choć w przypadku $\mu_0 = 1000$ i $d = 10$ osiągnięcie tego stanu trwa dość długo (ok. 300 kroków próbnika). Zaobserwować

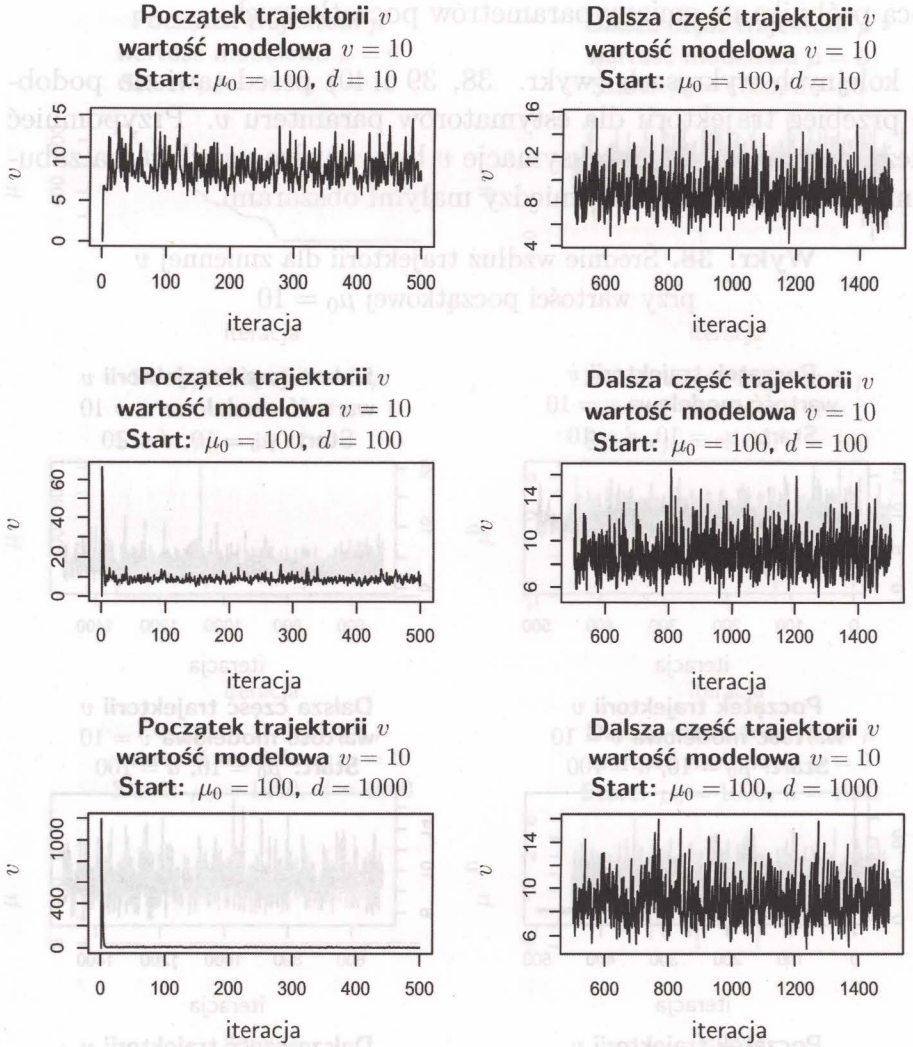
można, że dalsze części trajektorii estmatorów μ dla każdej z dziewięciu kombinacji początkowych wyglądają bardzo podobnie. Świadczy to o niewrażliwości estymatorów parametru μ otrzymywanych za pomocą próbnika na zmiany parametrów początkowych.

Na kolejnych wykresach (wykr. 38, 39 i 40) przedstawiono podobnie przebieg trajektorii dla estymatorów paramteru v. Przypomnieć należy, że próbnikowe aproksymacje v były bardzo wrażliwe na zaburzenia poziomu zmienności między małymi obszarami.

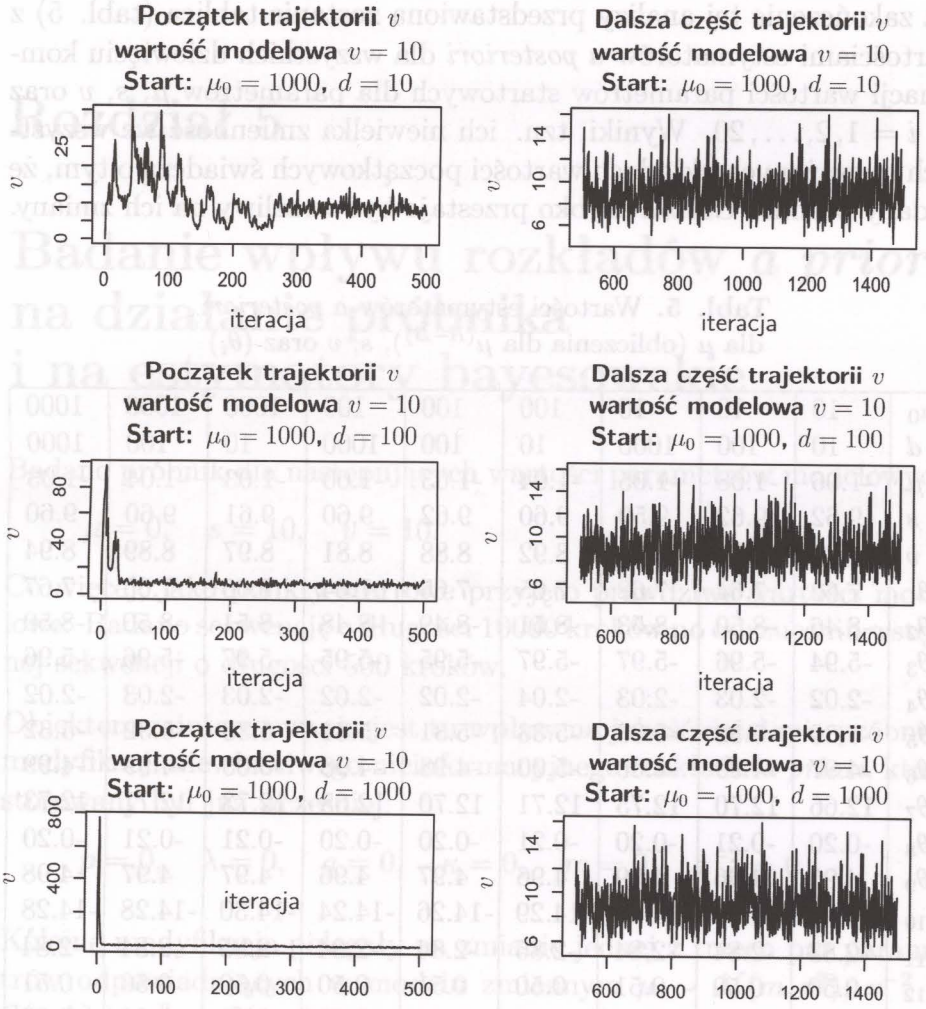
Wykr. 38. Średnie wzdłuż trajektorii dla zmiennej v przy wartości początkowej $\mu_0 = 10$



Wykr. 39. Średnie wzdłuż trajektorii dla zmiennej v przy wartości początkowej $\mu_0 = 100$



Wykr. 40. Średnie wzdłuż trajektorii dla zmiennej v przy wartości początkowej $\mu_0 = 1000$



Na wykresach trajektorii estymatorów v zaobserwować można podobne zjawisko, co dla trajektorii estymatorów μ. Pierwsze części trajektorii są w dużo większej skali niż drugie. Stabilizacja następuje stosunkowo szybko, choć ponownie dla wartości początkowych $\mu_0 = 1000$ i $d = 10$ jest ona najwolniejsza, wymaga ok. 300 kroków próbnika. To wynika z faktu, że wielkości μ_0 i (θ_i) interweniują w parametry warunkowego rozkładu v już w pierwszym losowaniu, które przeprowadza próbnik.

Ponownie stabilizacja każdorazowo następuje już w pierwszej części trajektorii.

Na zakończenie tej analizy przedstawiona zostanie tablica (tabl. 5) z wartościami estymatorów *a posteriori* dla wszystkich dziewięciu kombinacji wartości parametrów startowych dla parametrów μ, s, v oraz θ_i, $i = 1, 2, \ldots, 20$. Wyniki, tzn. ich niewielka zmienność we wszystkich rozważanych układach wartości początkowych świadczą o tym, że badany próbnik Gibbsa szybko przestaje być wrażliwy na ich zmiany.

Tabl. 5. Wartości estymatorów *a posteriori* dla μ (obliczenia dla $\mu^{(R-B)}$), s, v oraz (θ_i)

μ_0	10	10	10	100	100	100	1000	1000	1000
d	10	100	1000	10	100	1000	10	100	1000
μ	-1.06	-1.08	-1.05	-1.04	-1.03	-1.00	-1.03	-1.04	-1.03
s	9.62	9.62	9.59	9.60	9.62	9.60	9.61	9.60	9.60
v	8.80	8.87	8.99	8.92	8.88	8.81	8.97	8.89	8.94
θ_1	7.62	7.64	7.68	7.65	7.65	7.64	7.66	7.65	7.67
θ_2	-8.46	-8.50	-8.53	-8.51	-8.49	-8.48	-8.51	-8.50	-8.50
θ_3	-5.94	-5.96	-5.97	-5.97	-5.95	-5.95	-5.97	-5.96	-5.96
θ_4	-2.02	-2.03	-2.03	-2.04	-2.02	-2.02	-2.03	-2.03	-2.02
θ_5	-5.30	-5.32	-5.33	-5.33	-5.31	-5.31	-5.33	-5.32	-5.32
θ_6	-4.97	-4.99	-4.50	-5.00	-4.98	-4.98	-5.00	-4.99	-4.99
θ_7	12.66	12.70	12.75	12.71	12.70	12.68	12.73	12.71	12.73
θ_8	-0.20	-0.21	-0.20	-0.21	-0.20	-0.20	-0.21	-0.21	-0.20
θ_9	4.95	4.96	4.99	4.96	4.97	4.96	4.97	4.97	4.98
θ_{10}	-14.22	-14.27	-14.31	-14.29	-14.26	-14.24	-14.30	-14.28	-14.28
θ_{11}	-2.83	-2.84	-2.84	-2.85	-2.84	-2.84	-2.85	-2.84	-2.84
θ_{12}	0.50	0.50	0.51	0.50	0.51	0.50	0.50	0.50	0.51
θ_{13}	3.89	3.90	3.92	3.90	3.91	3.90	3.91	3.91	3.92
θ_{14}	20.32	20.38	20.46	20.41	20.38	20.36	20.43	20.40	20.43
θ_{15}	-8.75	-8.79	-8.81	-8.80	-8.78	-8.77	-8.80	-8.79	-8.79
θ_{16}	-6.70	-6.73	-6.74	-6.74	-6.72	-6.72	-6.74	-6.73	-6.73
θ_{17}	-4.47	-4.49	-4.50	-4.50	-4.49	-4.48	-4.50	-4.49	-4.49
θ_{18}	2.91	2.92	2.94	2.92	2.92	2.92	2.925	2.92	2.93
θ_{19}	-6.23	-6.26	-6.27	-6.27	-6.25	-6.24	-6.27	-6.26	-6.26
θ_{20}	-3.47	-3.48	-3.49	-3.49	-3.48	-3.48	-3.50	-3.48	-3.48

Rozdział 5

Badanie wpływu rozkładów *a priori* na działanie próbnika i na estymatory bayesowskie

Badano próbnik dla następujących wartości parametrów modelowych:

$$\mu = 0, \quad s = 10, \quad v = 10.$$

Co więcej, jako punkty startowe przyjęto prawdziwe wartości modelowe. Badano sekwencję o długości 10000 kroków po odrzuceniu wstępnej sekwencji o długości 500 kroków.

Obiektem zainteresowania jest tu wpływ na jakość działania próbnika modyfikacji niewłaściwego nieinformacyjnego rozkładu *a priori*, który stosowany był już wcześniej:

$$p = 0, \quad \lambda = 0, \quad q = 0, \quad \kappa = 0, \quad m = 0, \quad \tau^{-2} = 0.$$

Kolejne modyfikacje polegały na zmianie jednej z trzech par parametrów odpowiadających w modelu zmiennym $\mu \sim \mathcal{N}(m, \tau^2)$, $v^{-2} \sim \mathcal{G}(p, \lambda)$ i $s^{-2} \sim \mathcal{G}(q, \kappa)$, przy utrzymaniu dla pozostałych dwóch par niewłaściwego rozkładu jednostajnego.

Dla rozkładu μ przyjmowane były następujące kombinacje parametrów: $m = 0, 1, 5$ oraz $\tau^{-2} = 0, 1, 10$. Przyjęto pozostałe parametry q, κ, p, λ równe zero.

Tabl. 6 przedstawia wartości estymatorów średniej *a posteriori* μ: bezpośredniego i typu Rao-Blackwella, dla wszystkich kombinacji parametrów m i τ^{-2}.

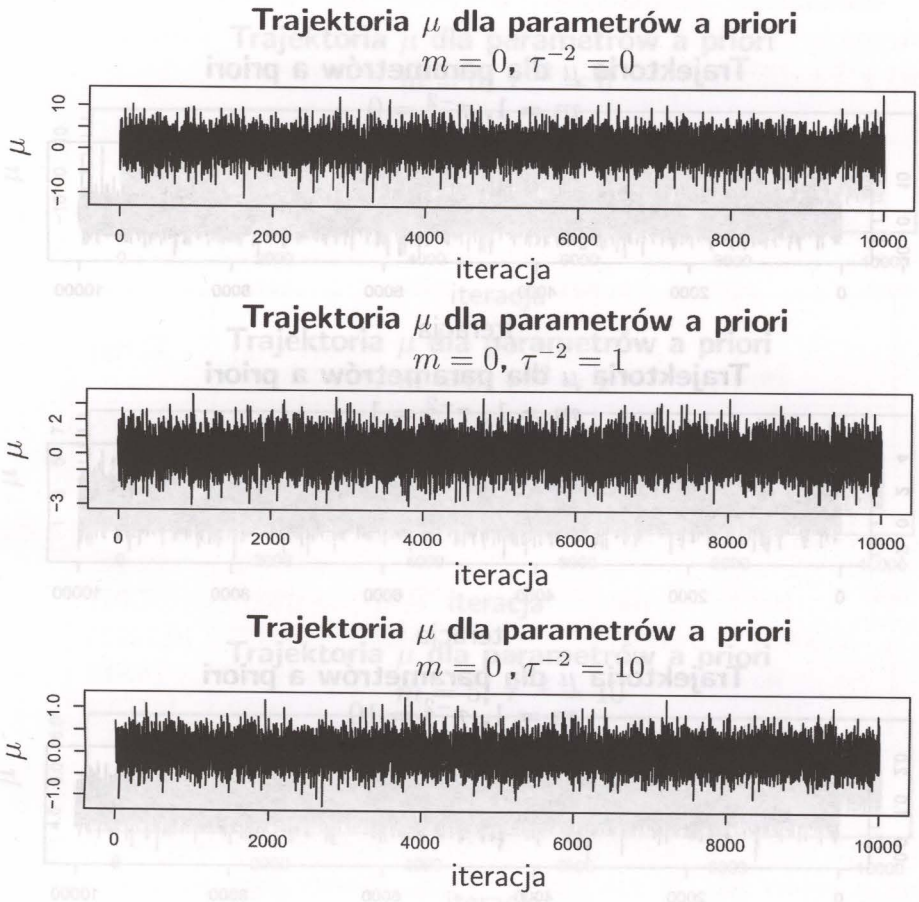
Tabl. 6. Wartości estymatorów średniej *a posteriori* $\mathbb{E}(\mu|Y)$ dla różnych kombinacji parametrów rozkładu *a priori* μ

m	τ^{-2}	μ	$\mu^{(R-B)}$
0	0	0.864	0.876
0	1	0.119	0.100
0	10	0.015	0.011
1	0	0.909	0.863
1	1	0.991	0.984
1	10	0.998	0.998
5	0	0.829	0.860
5	1	4.554	4.553
5	10	4.950	4.951

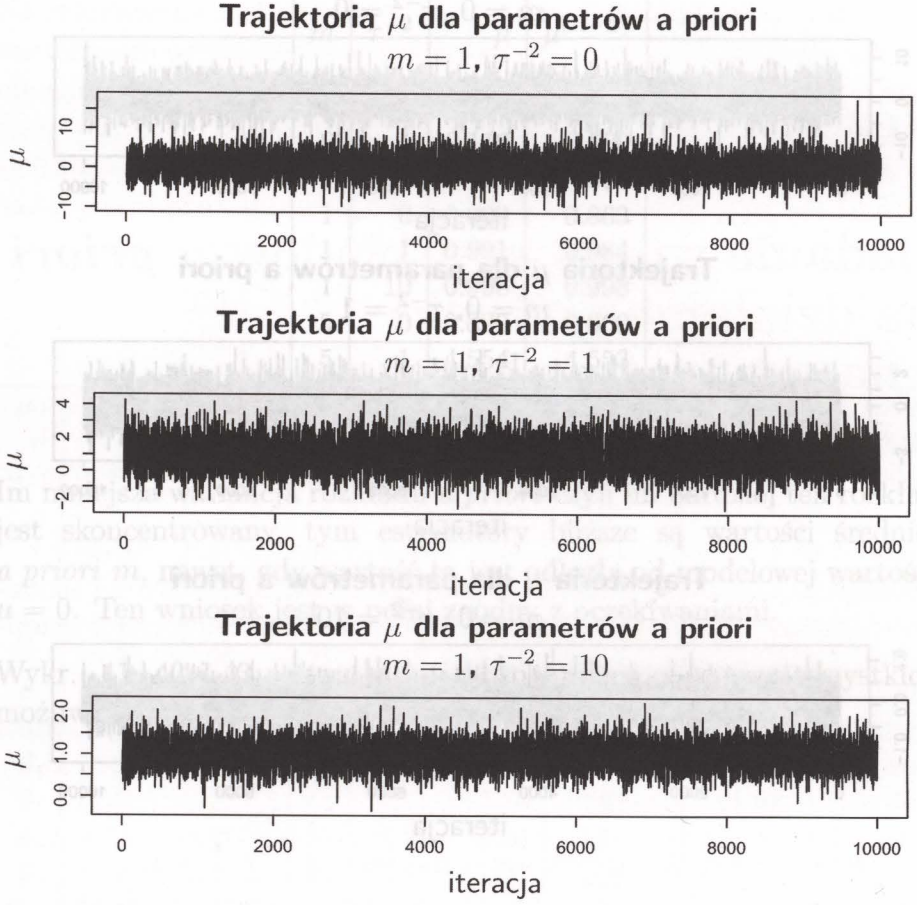
Im mniejsza wariancja rozkładu *a priori* czyli im bardziej ten rozkład jest skoncentrowany, tym estymatory bliższe są wartości średniej *a priori* m, nawet, gdy wartość ta jest odległa od modelowej wartości $\mu = 0$. Ten wniosek jest w pełni zgodny z oczekiwaniami.

Wykr. 41, 42 i 43 pokazują trajektorie zmiennej μ przy wszystkich możliwych kombinacjach parametrów m i τ^{-2}.

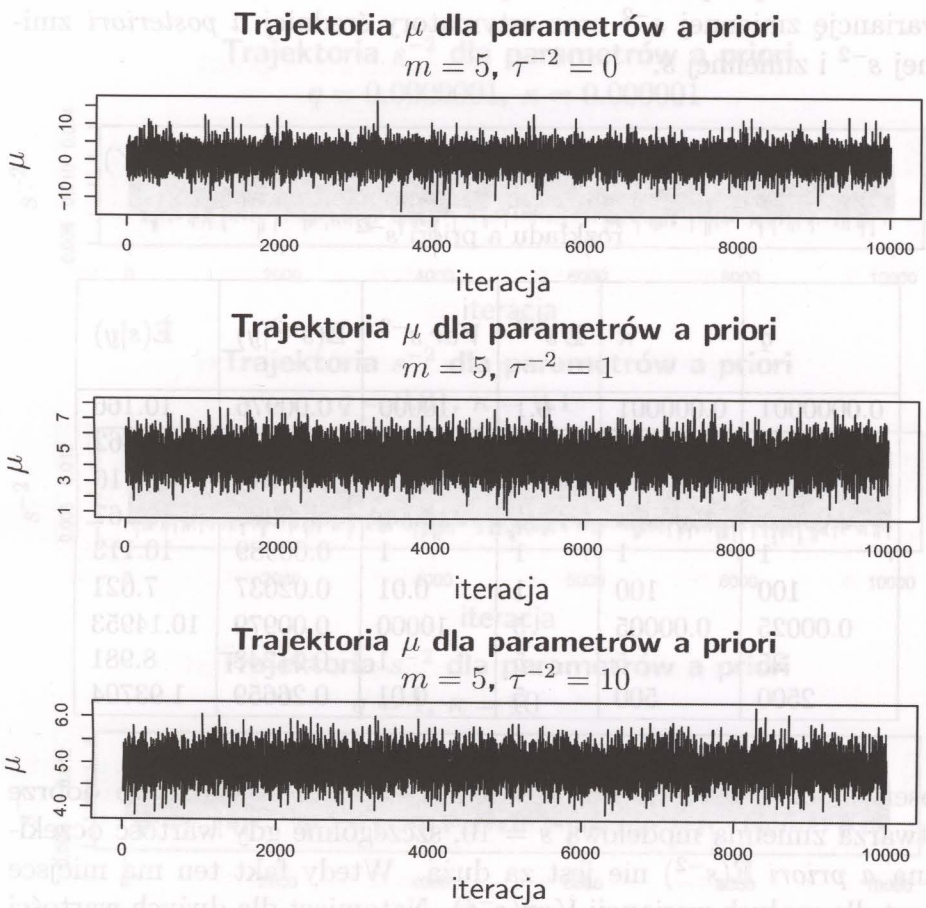
Wykr. 41. Trajektorie zmiennej μ dla $m = 0$ i $\tau^{-2} = 0, 1, 10$



Wykr. 42. Trajektorie zmiennej μ dla $m = 1$ i $\tau^{-2} = 0, 1, 10$



Wykr. 43. Trajektorie zmiennej μ dla $m = 5$ i $\tau^{-2} = 0, 1, 10$



Trajektorie te w każdej z rozważanych sytuacji mają bardzo podobny kształt; rozkłady są co najwyżej przesunięte. W każdym z obserwowanych przypadków wygląda, że próbnik szybko osiąga stan stacjonarny. Z przytoczonych wyników symulacji można wywnioskować, że niezależnie od parametrów *a priori* próbnik dobrze radzi sobie z osiąganiem średniej *a posteriori*. Inna rzecz, że dla nieadekwatnie dobranych rozkładów *a priori* (o zbyt małej wariancji i nietrafnie wybranej średniej) estymatory bayesowskie źle przybliżają wielkości modelowe.

W kolejnej tablicy (tabl. 7) przedstawione są wartości parametrów rozkładu s^{-2}, dla których prowadzone były symulacje. Pozostałe para-

metry m, τ^{-2}, p, λ równe były zero. W tabelce uwzględniono przeliczenie kombinacji parametrów q i κ na wartość oczekiwaną i wariancję zmiennej s^{-2} oraz estymatory średniej *a posteriori* zmiennej s^{-2} i zmiennej s.

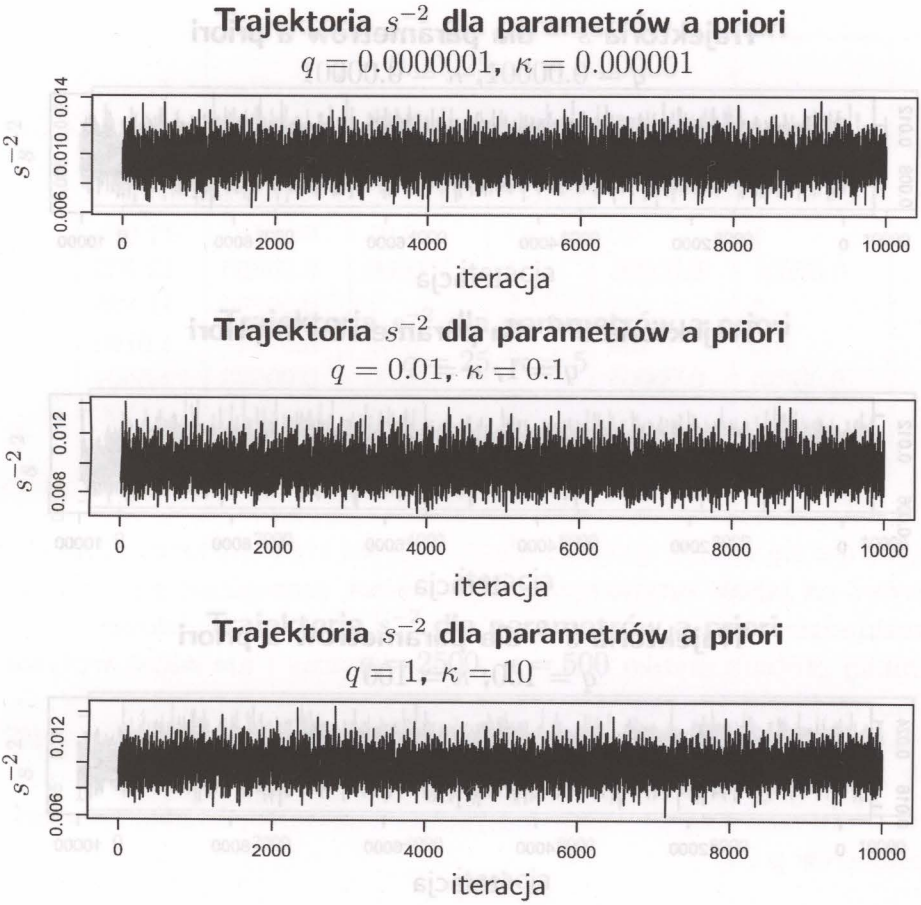
Tabl. 7. Wartości estymatorów średniej a posteriori $\mathbb{E}(s|Y)$ (oraz $\mathbb{E}(s^{-2}|Y)$) dla różnych kombinacji parametrów rozkładu a priori s^{-2}

q	κ	$\mathbb{E}\, s^{-2}$	$Var\, s^{-2}$	$\mathbb{E}(s^{-2}\|y)$	$\mathbb{E}(s\|y)$
0.0000001	0.000001	0.1	10000	0.00976	10.166
0.01	0.1	0.1	1	0.00977	10.162
1	10	0.1	0.01	0.00985	10.116
0.00001	0.00001	1	10000	0.00976	10.167
1	1	1	1	0.00989	10.113
100	100	1	0.01	0.02037	7.021
0.00025	0.00005	5	10000	0.00979	10.14953
25	5	5	1	0.01248	8.981
2500	500	5	0.01	0.26659	1.93704

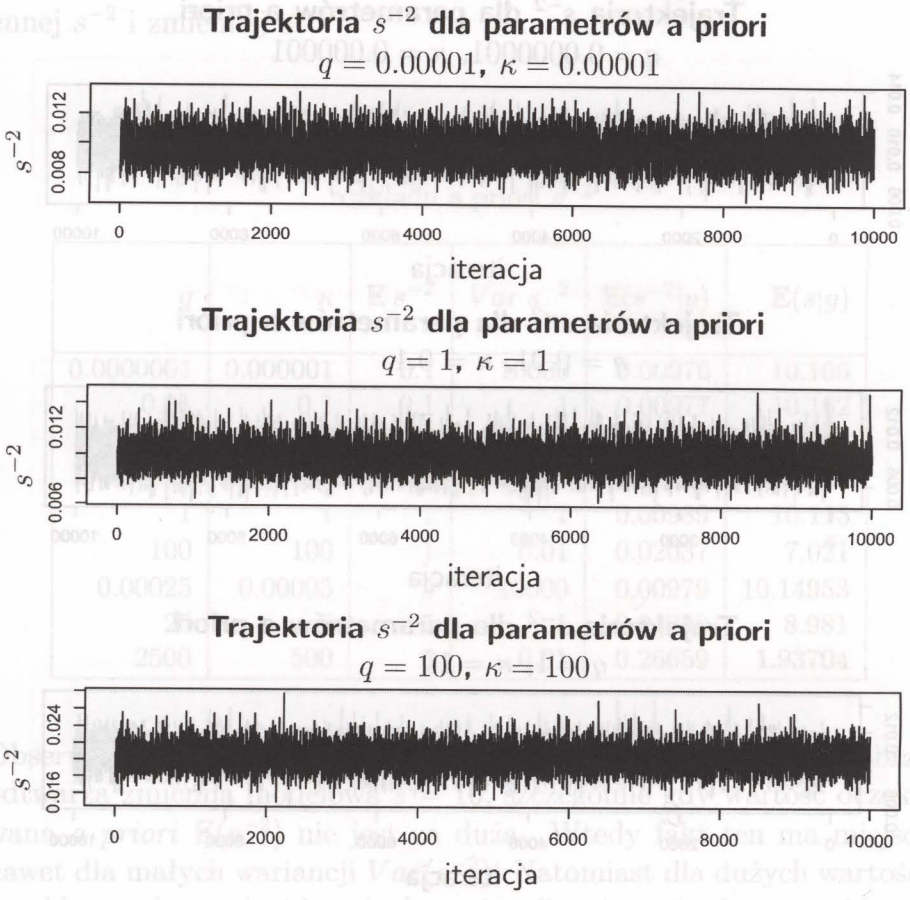
Obserwujemy, że estymator w wielu sytuacjach stosunkowo dobrze odtwarza zmienną modelową $s = 10$, szczególnie gdy wartość oczekiwana *a priori* $\mathbb{E}(s^{-2})$ nie jest za duża. Wtedy fakt ten ma miejsce nawet dla małych wariancji $Var(s^{-2})$. Natomiast dla dużych wartości oczekiwanych *a priori* i małych wariancji, estymator bayesowski wyraźnie nie radzi sobie ze zbliżaniem się do wartości modelowych.

Wykr. 44, 45 i 46 przedstawiają trajektorie estymatorów zmiennej s dla wszystkich rozważanych kombinacji rozkładu *a priori* dla s^{-2}.

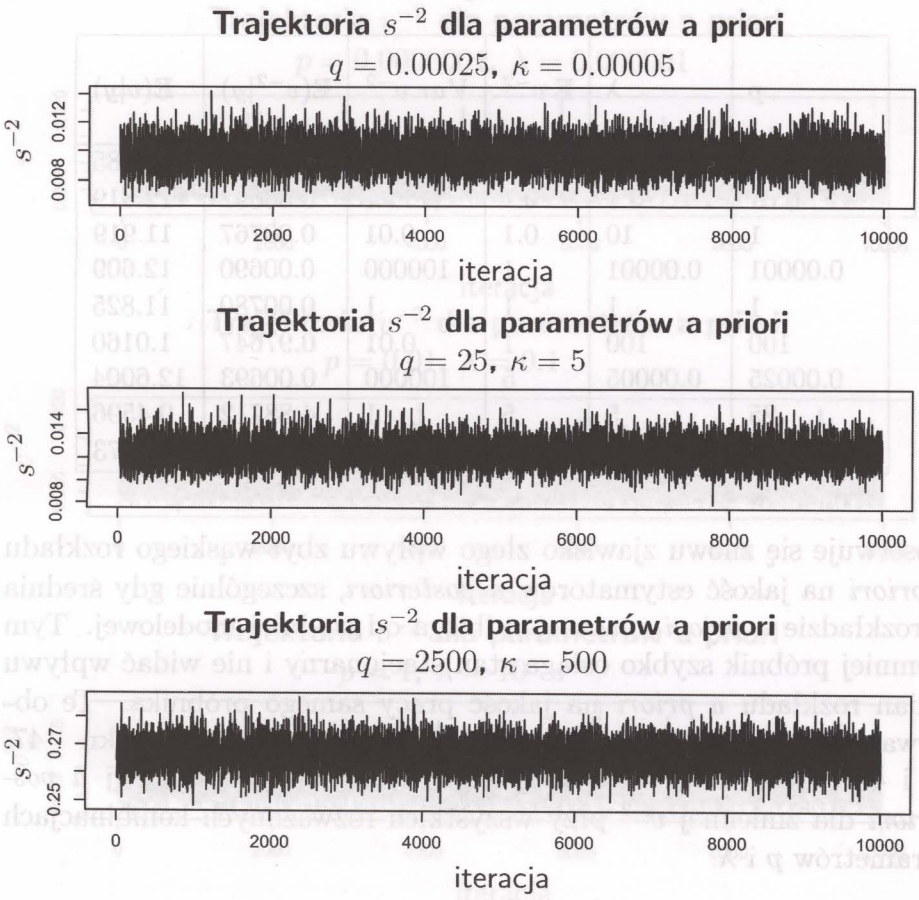
Wykr. 44. Trajektorie estymatorów zmiennej s^{-2} dla pierwszych trzech kombinacji parametrów (q, κ)



Wykr. 45. Trajektorie estymatorów zmiennej s dla kolejnych trzech kombinacji parametrów (q, κ)



Wykr. 46. Trajektorie estymatorów zmiennej s^{-2} dla ostatnich trzech kombinacji parametrów (q, κ)



Zaobserwować można, że próbnik osiąga stan stacjonarny w zakresie zmiennej s^{-2} dla wszystkich rozważanych kombinacji parametrów q i κ. Jednak zbyt małe wartości wariancji *a priori* mogą wymuszać niepożądane własności objawiające się tym, że odpowiednie rozkłady *a posteriori* znacząco odbiegają od wartości modelowych.

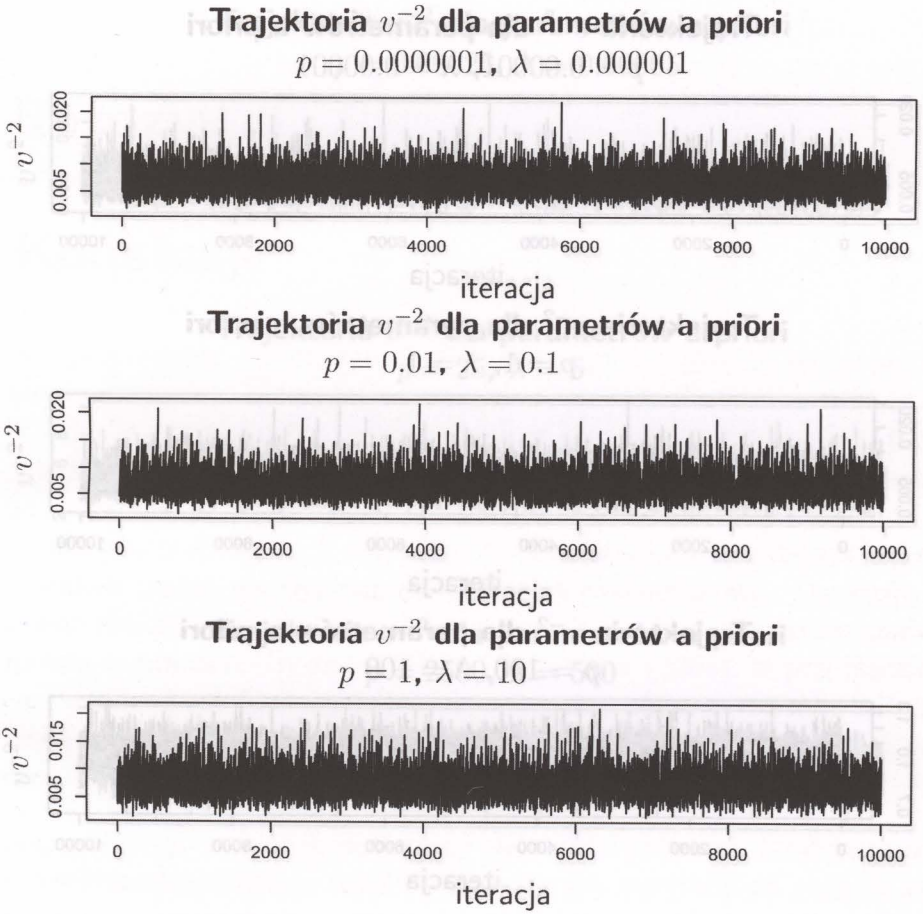
Trzecia para modyfikowanych parametrów to parametry rozkładu *a priori* zmiennej $v^{-2} \sim \mathcal{G}(p, \lambda)$. Rozważane były te same kombinacje parametrów p i λ, co wcześniej dla parametrów q i κ dla zmiennej s^{-2}. Wyniki estymacji podane są w tabl. 8. Ponownie, przyjęto wartości zerowe dla pozostałych parametrów m, τ^{-2}, q i κ.

Tabl. 8. Wartości estymatorów średniej a posteriori $\mathbb{E}(v|Y)$ (oraz $\mathbb{E}(v^{-2}|Y)$) dla różnych kombinacji parametrów rozkładu a priori v^{-2}

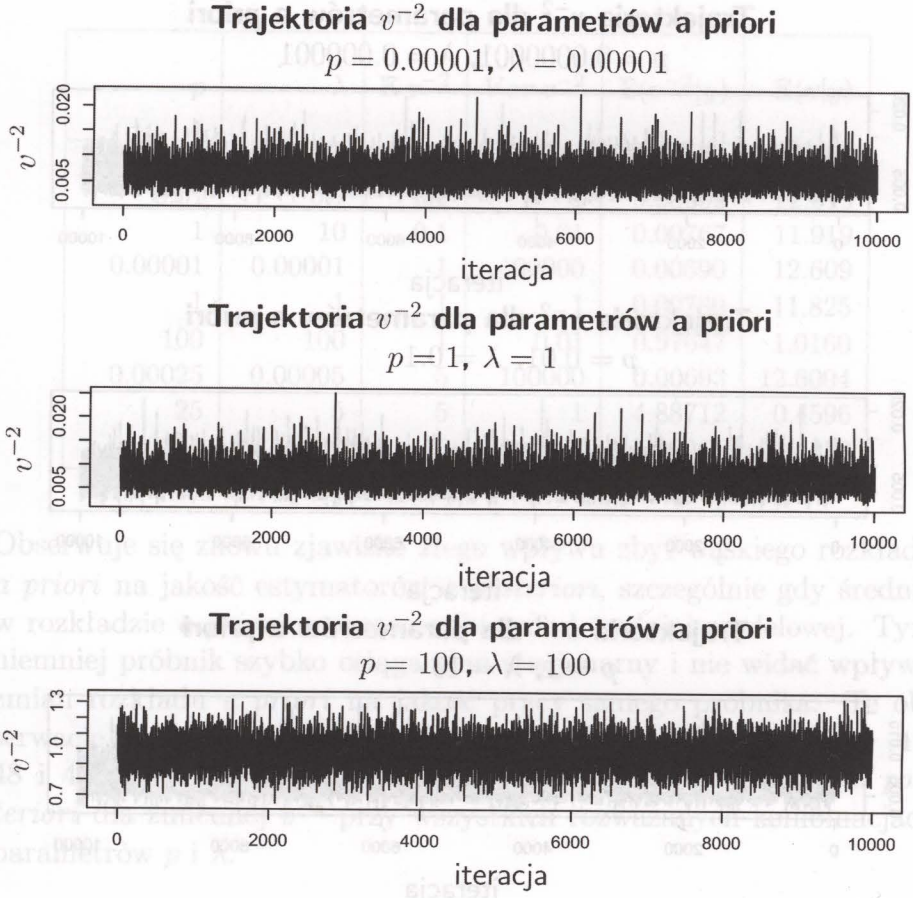
p	λ	$\mathbb{E}\, v^{-2}$	$Var\, v^{-2}$	$\mathbb{E}(v^{-2}\|y)$	$\mathbb{E}(v\|y)$
0.0000001	0.000001	0.1	100000	0.00694	12.585
0.01	0.1	0.1	1	0.00692	12.619
1	10	0.1	0.01	0.00767	11.919
0.00001	0.00001	1	100000	0.00690	12.609
1	1	1	1	0.00780	11.825
100	100	1	0.01	0.97647	1.0160
0.00025	0.00005	5	100000	0.00693	12.6004
25	5	5	1	4.88712	0.4596
2500	500	5	0.01	4.99965	0.4473

Obserwuje się znowu zjawisko złego wpływu zbyt wąskiego rozkładu *a priori* na jakość estymatorów *a posteriori*, szczególnie gdy średnia w rozkładzie *a priori* znacząco odbiega od średniej modelowej. Tym niemniej próbnik szybko osiąga stan stacjonarny i nie widać wpływu zmian rozkładu *a priori* na jakość pracy samego próbnika. Te obserwacje są również wnioskiem z kolejnych wykresach (wykr. 47, 48 i 49) przedstawiających trajektorie estymatorów średniej *a posteriori* dla zmiennej v^{-2} przy wszystkich rozważanych kombinacjach parametrów p i λ.

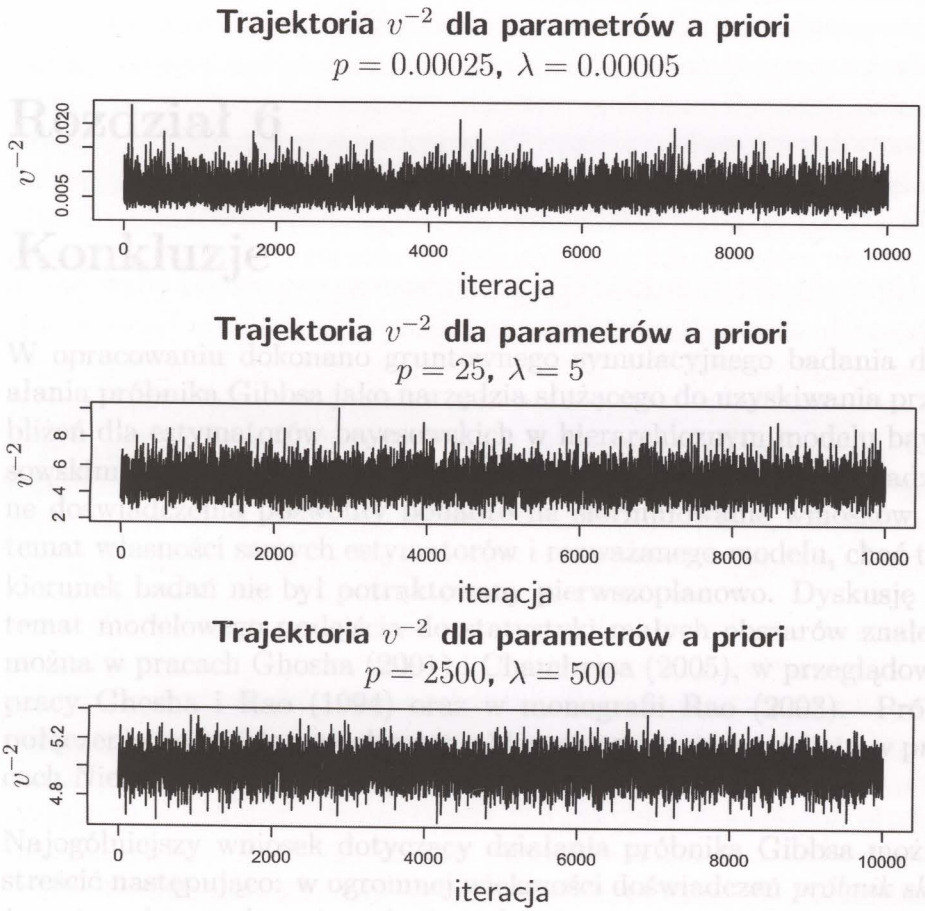
Wykr. 47. Trajektorie estymatorów zmiennej v^{-2} dla pierwszych trzech kombinacji parametrów (p, λ)



Wykr. 48. Trajektorie estymatorów zmiennej v^{-2} dla kolejnych trzech kombinacji parametrów (p, λ)



Wykr. 49. Trajektorie estymatorów zmiennej v^{-2} dla ostatnich trzech kombinacji parametrów (p, λ)



Rozdział 6

Konkluzje

W opracowaniu dokonano gruntownego symulacyjnego badania działania próbnika Gibbsa jako narzędzia służącego do uzyskiwania przybliżeń dla estymatorów bayesowskich w hierarchicznym modelu bayesowskim typowym dla statystyki małych obszarów. Przeprowadzone doświadczenia pozwoliły ponadto na sformułowanie wniosków na temat własności samych estymatorów i rozważanego modelu, choć ten kierunek badań nie był potraktowany pierwszoplanowo. Dyskusję na temat modelowego podejścia do statystyki małych obszarów znaleźć można w pracach Ghosha (2001) i Chambersa (2005), w przeglądowej pracy Ghosha i Rao (1994) oraz w monografii Rao (2003). Próba połączenia podejścia modelowego z klasycznym została podjęta w pracach Niemiry i Wesołowskiego (2008, 2009).

Najogólniejszy wniosek dotyczący działania próbnika Gibbsa można streścić następująco: w ogromnej większości doświadczeń *próbnik skutecznie wykonywał swoje zadanie*, polegające na generowaniu próbek z rozkładu zbliżonego do stacjonarnego i aproksymowaniu estymatorów bayesowskich.

Do tego wniosku prowadzą wyniki zebrane w rozdziałach 2, 3, 4 i 5. Jedyna seria doświadczeń, która mogła nasunąć pewne wątpliwości co do stabilizacji próbnika, opisana w rozdziale 2, dotyczy niezróżnicowanych obszarów w połączeniu z nieinformacyjnymi rozkładami *a priori*. Jest to sytuacja skrajnie niekorzystna, w której sam sens zastosowania hierarchicznego modelu bayesowskiego wydaje się wątpliwy. Jedynie wtedy próbnik miał trudności z ustabilizowaniem swojego przebiegu. We wszystkich innych, bardziej realistycznych przykładach,

próbnik stabilizował się zazwyczaj zaskakująco szybko osiągając stan stacjonarny. W szczególności świadczą o tym wyniki z rozdziałów 3 i 4, dotyczące obciążenia i zależności od punktu startowego. Na podstawie tych wyników można zaryzykować wniosek, że faza *burnin* przebiegu próbnika, potrzebna do redukcji obciążenia, może być bardzo krótka. W kilku doświadczeniach w rozdziale 4 punkt startowy został celowo dobrany bardzo niekorzystnie. W najgorszym przypadku wydłużyło to okres oczekiwania na stabilizację do ok. 300 kroków. W większości doświadczeń przebieg stabilizował się po zaledwie *kilku* krokach!

Szybkość zbieżności do rozkładu stacjonarnego jest przedmiotem badań w pracach wielu matematyków, np. Douc i in. (2004), Fort (2002), Roberts i Rosenthal (2004) oraz Rosenthal (2004). W szczególności należy tu zwrócić uwagę na prace Rosenthala (1995), Hoberta i Geyera (1998), Jonesa i Hoberta (2004) oraz Johnsona i Jonesa (2008), gdzie rozpatrywana jest zbieżność w sensie normy „pełnego wahania" dla hierarchicznego modelu komponentów wariancyjnych, w istocie identycznego z tu rozważanym.

Kwestia ta ma niewątpliwie znaczenie teoretyczne, ale nie wydaje się, aby przedstawiała poważny problem w praktycznym stosowaniu próbnika Gibbsa, przynajmniej dla rozważanego modelu małych obszarów. Poważniejszy problem stanowi zachowanie średnich wzdłuż trajektorii, czyli przybliżeń estymatorów bayesowskich. Zbieżność tych średnich do docelowych wartości jest konsekwencją stabilizacji przebiegu próbnika (po stosunkowo krótkim okresie *burnin*). Zaobserwowano tę zbieżność we właściwie wszystkich seriach doświadczeń i systematycznie przeanalizowano w rozdziale 3. Wyniki świadczą, że można traktować próbnik Gibbsa jako godną zaufania metodę obliczeniową Monte Carlo.

Ciekawym wnioskiem z przeprowadzonych eksperymentów jest stwierdzenie zdecydowanej przewagi estymatorów Rao-Blackwella nad prostymi średnimi wzdłuż trajektorii. Na podstawie literatury (np. Asmussen i Glynn, 2007) można było przewidzieć, że warunkowe estymatory są lepsze, ale dopiero doświadczenia symulacyjne pokazują na ile poprawiają jakość przybliżeń.

Jeśli chodzi o oszacowanie błędu aproksymacji Monte Carlo, to

z jednej strony niniejsze opracowanie potwierdza użyteczność wyników asymptotycznych, opartych na Centralnym Twierdzeniu Granicznym dla łańcuchów Markowa, por. Roberts i Rosenthal (2004), Jones i in. (2006). Symulacje relacjonowane w rozdziale 3 pokazują, że zachowanie błędu średniokwadratowego jest zaskakująco zgodne z teorią asymptotyczną i że obciążenie jest, też zgodnie z tą teorią, zaniedbywalne. Fakt ten wskazuje na praktyczną ważność problemu estymacji lub/i szacowania z góry błędu średniokwadratowego próbnika Gibbsa. Niestety, ten problem, nawet w ogólniejszym kontekście metod MCMC jest znacznie słabiej zbadany, niż problem zbieżności do rozkładu stacjonarnego. Estymacja jest rozważana w pracy Jonesa i in. (2006). Oszacowania nieasymptotyczne użyteczne w zastosowaniach do statystyki małych obszarów są bardzo trudne. Wstępne zaledwie wyniki tego typu są zawarte w pracach Łatuszyńskiego (2008) oraz Łatuszyńskiego i Niemiry (2006, 2009) oraz Łatuszyńskiego i in. (2009). Widać tu dużą lukę pomiędzy potrzebami praktyki i możliwościami teorii. W tym kontekscie niniejsze opracowanie może sugerować kierunek dalszych badań, por. także Chambers i in. (2006).

Własności estymatorów i modelu bayesowskiego nie były co prawda głównym celem opisanych tutaj badań, ale na podstawie przeprowadzonych symulacji można sformułować również na ten temat wnioski. Wstępne symulacje wykazały, że przy dużej zmienności między małymi obszarami estymatory działały bardzo dobrze. Jak już wspomniano, próbnik szybko się stabilizuje i dobrze odtwarza rozkłady *a posteriori*. Najpierw rozważono sytuację, gdy rozkład *a priori* ma charakter nieinformacyjny. Konkretnie zastosowano niewłaściwy rozkład *a priori* o nieskończonej wariancji. Wtedy estymatory bayesowskie mogą w zasadzie polegać jedynie na danych. Jednakże przy małej zmienności między małymi obszarami i nieinformacyjnym rozkładzie *a priori* estymatory nie potrafią odróżniać małych obszarów. Wartości estymatorów są odległe od wartości modelowych. Łączy się to z gorszym działaniem próbnika Gibbsa. Jego stabilizacja następuje znacznie wolniej. Aproksymacje Monte Carlo mają duże wariancje, nawet w sytuacji, gdy punktem startu są wartości modelowe.

Nakładają się dwa niekorzystne efekty: zła jakość estytmatora *a posteriori* i kiepskie działanie próbnika. Sytuację można poprawić przyj-

mując rozkład *a priori*, który lepiej przybliża prawdziwe wartości parametrów. Symulacje opisane w trzeciej części rozdziału 2 pokazują, że mimo niezróżnicowanych rozkładów w małych obszarach, podanie odpowiedniego, informacyjnego rozkładu *a priori* znacznie poprawia działanie próbnika. Okres stabilizacji jest tylko nieznacznie dłuższy niż w sytuacji gdy obszary są zróżnicowane. Tym niemniej same wartości estymatorów *a posteriori* są odległe od wartości modelowych. Przeważa efekt niewielkiego zróżnicowania małych obszarów w stosunku do jednostkowego błędu losowego.

W rozdziale 5 zrelacjonowano badania wpływu rozkładu *a priori* na estymatory i działanie próbnika. Wskazują one, że niezależnie od tego rozkładu próbnik szybko osiąga stan stacjonarny. Potwierdza to wspomniane konkluzje na temat satysfakcjonującego działania próbnika Gibbsa jako metody obliczeniowej. Jednakże zbyt małe wartości wariancji rozkładów *a priori* prowadzą do niepożądanych własności rozkładów *a posteriori*. W szczególności ich średnie, czyli estymatory bayesowskie mogą znacząco odbiegać od przyjętych wartości w modelu. Ten właśnie efekt jest obserwowany w symulacjach omówionych w rozdziale 5.

Podsumowując, rozważany tu model bayesowski dla statystyki małych obszarów i wynikające z tego modelu estymatory można uznać za użyteczne i satysfakcjonujące gdy zmienność pomiędzy małymi obszarami jest dostatecznie wyraźna. Wtedy wybór rozkładów *a priori* nie ma decydującego znaczenia dla wartości estymatorów. Jeśli natomiast zmienność pomiędzy obszarami jest mała (w porównaniu ze zmiennością wewnątrz małych obszarów), to tylko bardzo konretna i wiarogodna informacja *a priori* umożliwia obliczenie użytecznie dokładnych estymatorów *a postriori* za pomocą próbnika Gibbsa.

Literatura

Asmussen R., Glynn P. W. (2007), *Stochastic Simulation: Algorithms and Analysis*, Springer, New York

Chambers R. (2005), *An introduction to model-based survey sampling*, [w:] *Workshop on Survey Sampling Theory and Methodology*, D. Krapavickaite, A. Plikusas, eds. Statistics Lithuania, Vilinius, 9–51

Chambers R., van den Brakel J., Hedlin D., Lehtonen R., Zhang L.-Ch. (2006), *Future Challenges of Small Area Estimation*, „Statistics in Transition", vol. 7, No. 4, 759–769

Douc R., Moulines E., Rosenthal J. R. (2004), *Quantitative bounds on convergence of time-inhomogeneous Markov Chains*, „Annals of Applied Probability", vol. 14, 1643–1665

Fort G. (2002), *Computable bounds for V-geometric ergodicity of Markov transition kernels*, Preprint

Ghosh M. (2001) *Model dependent small area estimation - theory and practice*, [w:] *Lecture Notes on Estimation for Population Domains and Small Areas*, R. Lehtonen, K. Djerf, eds. Statistics Finland, Helsinki 2001, 51–108

Ghosh M., Rao J. N. K. (1994), *Small Area Estimation: an appraisal*, „Statistical Science", vol. 9, 55–76

Hobert J. P., Geyer C. J. (1998), *Geometric ergodicity of Gibbs and block Gibbs samplers for Hierarchical Random Effects Model*, „Journal of Multivariate Analysis", vol 67, 414–439

Jones G. L., Haran M., Caffo B. S., Neath, R. (2006), *Fixed-width output analysis for Markov chain Monte Carlo*, „Journal of American Statistical Association", vol. 101, 1537–1547

Jones G.J., Hobert J.P. (2004), *Sufficient burn-in for Gibbs samplers for a hierarchical random effects model*, „Annals of Statistics", vol. 32, 784–817

Łatuszyński K. (2008), *MCMC $\varepsilon - \alpha$ approximation under drift condition with application to Gibbs samplers for a hierarchical random effects model*, praca doktorska, Wydział Matematyki, Informatyki i Mechaniki, Uniwersytet Warszawski (arXiv:0907.4716)

Łatuszyński K., Niemiro W. (2006), *$(\varepsilon - \alpha)$-MCMC approximation under drift condition*, [w:] *Proceedings of the 6th International Workshop on Rare Event Simulation*, RESIM

Łatuszyński K., Miasojedow B., Niemiro W. (2009), *Non-assymptotic bounds on the estimation error for regenerative MCMC algorithms*, praca złożona do druku, arXiv:0907.4915v1

Łatuszyński K., Niemiro W. (2009), *Rigorous confidence bounds for MCMC under a geometric drift condition*, praca złożona do druku, arXiv:0908.2098v1

Niemiro W., Wesołowski J. (2008), *Synthetic and composite estimation under a superpopulation model*, „Statistical Papers", DOI 10.1007/s00362-008-0136-1

Niemiro W., Wesołowski J. (2009), *Linear estimation and prediction under model-design approach with small area effects*, praca złożona do druku

Rao J. N. K. (2003), *Small Area Estimation*, Wiley, Hoboken

Roberts G. O., Rosenthal J. S. (2004), *General state space Markov chains and MCMC algorithms*, „Probability Surveys", vol. 1, 20–71

Rosenthal J. S. (2004), *Quantitative convergence rates of Markov chains: a simple account*, „Electronic Communications in Probability", vol. 7, 123–128

Rosenthal J. S. (1995), *Rates of convergence for Gibbs sampling for variance component models*, „Annals of Statistics", vol. 23, 740–761

A/53398

183751

A/53398

183751

183751

ISBN 978-83-7027-447-4

CENTRALNA BIBLIOTEKA
STATYSTYCZNA - GUS
066456