Online Supplement to: ‘A Similarity Based Model for Ordered Categorical Data’

Gabi Gayer, Offer Lieberman and Omer Yaffe*

June 2, 2016

1 Summary

This Supplement provides: (i) Appendix B conatining proofs of consistency and asymptotic normality of the MLE; (ii) Tables and Figures relating to the simulations section of the paper.

2 Appendix B: Consistency and Asymptotic Normality

Proof of Theorem 3: The proof can be made by either checking the conditions of Proposition 7.5 of Hayashi (2000), Theorem 2.7 of Newey and McFadden (1994), or by directly verifying Wu's (1981) criterion. For any $\delta_1 > 0$, denote by $B_{\delta_1}(\theta_0)$ the ball $\{\theta \in \Theta : ||\theta - \theta_0|| \leq \delta_1\}$ and by $B^c_{\delta_1}(\theta_0)$ the complement of $B_{\delta_1}(\theta_0)$ in Θ. For any $\theta \in \Theta$, let

$$D_n(\theta_0, \theta_1) = \frac{1}{n}\left(l_n(\theta_0) - l_n(\theta_1)\right).$$

To establish consistency, we must prove that $\forall \delta_1 > 0$,

$$\lim_{n\to\infty}\inf \inf_{B^c_{\delta_1}(\theta_0)} D_n(\theta_0, \theta_1) \quad (1)$$

*Department of Economics, Bar-Ilan University

is strictly positive in probability. See, for instance, Wu (1981).

Let

$$l_{n,j}(\theta) \equiv \frac{1}{n}\sum_{t=1}^{n} 1\{Y_t = j\} \ln \Delta_{t,j}(\theta).$$

By Assumption A1, all the μ_j's are different from each other and therefore,

$$0 < \Delta_{t,j}(\theta) < 1$$

for all t, j and θ. Hence, there exists an L_1 satisfying

$$-\infty < L_1 < \ln \Delta_{t,j}(\theta) < 0.$$

The series $\{l_{n,j}(\theta)\}$ is evidently nonpositive and uniformly bounded from below and as a consequence of Theorem 1(ii), it is convergent $w.p.1$. We shall denote this limit by $l_j(\theta)$. This implies that $\forall\theta \in \Theta$, $n^{-1}l_n(\theta) \to_{a.s.} \sum_{j=1}^{M} l_j(\theta) \equiv l(\theta)$. Using Jensen's inequality and the fact that $\sum_{j=1}^{M}\Delta_{t,j}(\theta_0) = 1$,

$$\begin{aligned} E_{\theta_0}(D_n(\theta_1,\theta_0)) &= \frac{1}{n}E_{\theta_0}\sum_{t=1}^{n}\sum_{j=1}^{M}E_{\theta_0}\left(1\{Y_t = j\}\ln\frac{\Delta_{t,j}(\theta_1)}{\Delta_{t,j}(\theta_0)}|\mathcal{F}_{t-1}\right) \\ &= \frac{1}{n}E_{\theta_0}\sum_{t=1}^{n}\sum_{j=1}^{M}\Delta_{t,j}(\theta_0)\ln\frac{\Delta_{t,j}(\theta_1)}{\Delta_{t,j}(\theta_0)} \\ &\leq \frac{1}{n}E_{\theta_0}\sum_{t=1}^{n}\ln(1) \qquad (2) \\ &= 0. \end{aligned}$$

If $\mu_0 \neq \mu_1$, $\Delta_t(\mu_0, w, \rho) \neq \Delta_t(\mu_1, w, \rho)$, $\forall t$ and if $w_0 \neq w_1$, $\bar{Y}^s_{t-1} \neq \bar{Y}^s_{t-1}$ with positive probability $\forall t$ under Assumption A0, which also implies $\Delta_t(\mu, w_0) \neq \Delta_t(\mu, w_1)$ with positive probability $\forall t$. Furthermore, if $\rho_0 \neq \rho_1$, $\Delta_t(\mu, w, \rho_0) \neq \Delta_t(\mu, w, \rho_1)$, $\forall t$. Hence, as $n \to \infty$, equality in (2) holds iff $\theta_0 = \theta$ and the proof of the Theorem is completed. ■

In order to prove Theorem 4, we shall require the following lemmas.

Lemma 1 *Under Assumptions A0-A2,* $z_n(\theta_0) \xrightarrow{d} N(0, V(\theta_0))$.

Proof of Lemma 1: Let

$$f_{t,k}(\theta) = \phi\left(\mu_k - \bar{Y}^s_{t-1}\right),$$

where ϕ is the standard normal PDF. As

$$\dot{\Delta}^{\mu_k}_{t,j}(\theta) \equiv \frac{\partial \Delta_{t,j}(\theta)}{\partial \mu_k} = f_{t,k}(\theta)\left(1\{j=k\} - 1\{j=k+1\}\right),$$

we have

$$z_{n,\mu_k}(\theta) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} G^{\mu_k}_t(\theta), \tag{3}$$

where

$$G^{\mu_k}_t(\theta) = f_{t,k}(\theta)\left(\frac{1\{Y_t = k\}}{\Delta_{t,k}} - \frac{1\{Y_t = k+1\}}{\Delta_{t,k+1}}\right). \tag{4}$$

We notice that

$$E_{\theta_0}\left(G^{\mu_k}_t(\theta)\,|\mathcal{F}_{t-1}\right) = E_{\theta_0} E_{\theta_0}\left(G^{\mu_k}_t(\theta)\,|X_t, \mathcal{F}_{t-1}\right) = 0$$

so that $G^{\mu_k}_t$ is an m.d.s.. Furthermore,

$$\dot{\Delta}^{w_k}_{t,j}(\theta) \equiv \frac{\partial \Delta_{t,j}(\theta)}{\partial w_k} = -\delta_{t,j}(\theta)\,\dot{h}^{w_k}_t(\theta)$$

where

$$\delta_{t,j}(\theta) = f_{t,j}(\theta) - f_{t,j-1}(\theta)$$

and

$$\dot{h}^{w_k}_t(\theta) = \frac{\partial}{\partial w_k} \bar{Y}^s_{t-1}.$$

Thus,

$$z_{n,w_k}(\theta) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} G^{w_k}_t(\theta), \tag{5}$$

where

$$G^{w_k}_t(\theta) = -\dot{h}^{w_k}_t(\theta) \sum_{j=1}^{M} 1\{Y_t = j\} \frac{\delta_{t,j}(\theta)}{\Delta_{t,j}(\theta)}. \tag{6}$$

We have,

$$\begin{aligned} E_{\theta_0}\left(G_t^{w_k}(\theta)\,|\mathcal{F}_{t-1}\right) &= -\dot{h}_{t,k}^{w}(\theta)\sum_{j=1}^{M}\delta_{t,j}(\theta) \\ &= -\dot{h}_{t,k}^{w}(\theta)\left(f_{t,M}(\theta)-f_{t,0}(\theta)\right) \\ &= 0, \end{aligned}$$

so that $G_t^{w_k}(\theta)$ is also an m.d.s.. Finally,

$$z_{n,\rho}(\theta)=\frac{1}{\sqrt{n}}\sum_{t=1}^{n}G_t^{\rho}(\theta),$$

where

$$G_t^{\rho}(\theta)=-\rho^{-1}\bar{Y}_{t-1}^{s}\sum_{j=1}^{M}1\left\{Y_t=j\right\}\frac{\delta_{t,j}(\theta)}{\Delta_{t,j}(\theta)}, \tag{7}$$

which is also an m.d.s.. For asymptotic normality of the score function, it will thus be sufficient to verify conditions (2.3) of McLeish (1974). Let $\sigma_n^{i_k}(\theta)^2=\sum_{t=1}^{n}\left(G_t^{i_k}(\theta)\right)^2$, $i=\mu$ with $k=1,...,M-1$, $i=w$ with $k=1,...K$, or $i=\rho$ with the k-index suppressed. We need to show that for each $\theta\in\Theta$,

$$\frac{\sigma_n^{i_k}(\theta)^2}{n}\xrightarrow{p}V^{i_k}(\theta)<\infty \tag{8}$$

and that $\forall\varepsilon>0$, i and k,

$$\frac{1}{\sigma_n^{i_k}(\theta)^2}\sum_{t=1}^{n}\left(G_t^{i_k}(\theta)\right)^2 1\left\{\left|G_t^{i_k}(\theta)\right|>\varepsilon\sigma_n^{i_k}(\theta)\right\}\xrightarrow{p}0. \tag{9}$$

As $f_{t,k}(\theta)<\infty$, uniformly in n,k and Θ,

$$\frac{1}{n}\sum_{t=1}^{n}\left(G_t^{\mu_k}(\theta)\right)^2=\frac{1}{n}\sum_{t=1}^{n}f_{t,k}(\theta)^2\left(\frac{1\left\{Y_t=k\right\}}{\Delta_{t,k}}-\frac{1\left\{Y_t=k+1\right\}}{\Delta_{t,k+1}}\right)^2<\bar{K} \tag{10}$$

and convergence is assured by ergodicity, the limit of which is denoted by $V^{\mu_k}(\theta)$. Also, because $G_t^{\mu_k}(\theta)$ is uniformly bounded and $\sigma_n^{\mu_k}(\theta)$ behaves as $\sqrt{n}$ in probability, condition (9) trivially holds and we are done for $z_{n,\mu_k}(\theta)$.

For $G_t^{w_k}(\theta)$, observe that

$$\dot{h}_t^{w_k}(\theta) = \rho\left(\frac{\sum_{i<t}\dot{s}_{w_k}(X_i,X_t;w)Y_i}{\sum_{i<t}s(X_i,X_t;w)} - \frac{\sum_{i<t}s(X_i,X_t;w)Y_i\sum_{i<t}\dot{s}_{w_k}(X_i,X_t;w)}{\left(\sum_{i<t}s(X_i,X_t;w)\right)^2}\right), \tag{11}$$

where $\dot{s}_{w_k}(X_i,X_t;w) = \partial s(X_i,X_t;w)/\partial w_k$. It follows from (11) that under Assumptions A1-A2,

$$\sup_{t,k,\Theta}\left|\dot{h}_t^{w_k}(\theta)\right| < 2|\rho|\bar{K}M.$$

In view of (6) and the last inequality

$$\sup_{t,k,\Theta}\left|G_t^{w_k}(\theta)\right| < \bar{K},$$

so that, together with ergodicity,

$$\frac{1}{n}\sum_{t=1}^{n}\left(G_t^{w_k}(\theta)\right)^2 \xrightarrow{p} V^{w_k}(\theta) < \infty.$$

Condition (9) also holds because $G_t^{w_k}(\theta)$ is uniformly bounded and $\sigma_n^{w_k}(\theta)$ behaves as $\sqrt{n}$ in probability. Similar reasoning follows for $G_t^{\rho}(\theta)$ and the proof of the Lemma 1 is therefore completed.■

Lemma 2 *Under Assumptions A0-A4,* $\forall\theta\in\Theta$,

$$\lim_{n\to\infty}E_\theta\left(\left(H_{n,\theta_j,\theta_k}(\theta)\right)_{1\le j,k\le K+M}\right)$$

is finite and nonsingular.

Proof of Lemma 2: We have

$$\begin{aligned}
\frac{\partial^2 l_n(\theta)}{\partial \mu_j \partial \mu_k} &= \sum_{t=1}^{n} \dot{f}_{t,k}(\theta)\left(\frac{1\{Y_t = k\}}{\Delta_{t,k}} - \frac{1\{Y_t = k+1\}}{\Delta_{t,k+1}}\right) 1\{j = k\} \\
&- \sum_{t=1}^{n} f_{t,k}^2(\theta) \frac{1\{Y_t = k\}}{\Delta_{t,k}^2} 1\{j = k\} \\
&+ \sum_{t=1}^{n} f_{t,j}(\theta) f_{t,j+1}(\theta) \frac{1\{Y_t = j+1\}}{\Delta_{t,j+1}^2} 1\{j = k-1\} \\
&+ \sum_{t=1}^{n} f_{t,k}(\theta) f_{t,k+1}(\theta) \frac{1\{Y_t = k+1\}}{\Delta_{t,k+1}^2} 1\{j = k+1\} \\
&- \sum_{t=1}^{n} f_{t,k}^2(\theta) \frac{1\{Y_t = k+1\}}{\Delta_{t,k+1}^2} 1\{j = k\},
\end{aligned}$$

with $\dot{f}_{t,j}(\theta) = \partial f_{t,j}(x;\theta)/\partial x$. Hence,

$$\begin{aligned}
E_\theta\left(\frac{\partial^2 l_n(\theta)}{\partial \mu_j \partial \mu_k} | \mathcal{F}_{t-1}\right) &= -\sum_{t=1}^{n} f_{t,k}^2(\theta)\left(\frac{1}{\Delta_{t,k}} + \frac{1}{\Delta_{t,k+1}}\right) 1\{j = k\} \\
&+ \sum_{t=1}^{n} f_{t,j}(\theta) f_{t,j+1}(\theta) \frac{1\{j = k-1\}}{\Delta_{t,j+1}} \\
&+ \sum_{t=1}^{n} f_{t,k}(\theta) f_{t,k+1}(\theta) \frac{1\{j = k+1\}}{\Delta_{t,k+1}}.
\end{aligned}$$

In view of (6) and under Assumption A3,

$$\begin{aligned}
\frac{\partial^2 l_n(\theta)}{\partial w_l \partial w_k} &= -\sum_{t=1}^{n} \ddot{h}_t^{w_k,w_l}(\theta) \sum_{j=1}^{M} 1\{Y_t = j\} \frac{\delta_{t,j}(\theta)}{\Delta_{t,j}(\theta)} \\
&- \sum_{t=1}^{n} \dot{h}_t^{w_k}(\theta) \sum_{j=1}^{M} 1\{Y_t = j\}\left(\frac{\dot{\delta}_{t,j,l}(\theta)}{\Delta_{t,j}(\theta)} + \frac{\delta_{t,j}^2(\theta)\, \dot{h}_t^{w_l}(\theta)}{\Delta_{t,j}^2(\theta)}\right),
\end{aligned}$$

where

$$\dot{\delta}_{t,j,l}(\theta) = \frac{\partial \delta_{t,j}(\theta)}{\partial w_l} = -\left(\dot{f}_{t,j}(\theta) - \dot{f}_{t,j-1}(\theta)\right) \dot{h}_t^{w_l}(\theta) = -\rho_{t,j}(\theta)\, \dot{h}_t^{w_l}(\theta),$$

say. We have,

$$E_{\theta_0}\left(\frac{\partial^2 l_n(\theta)}{\partial w_k \partial w_l}|\mathcal{F}_{t-1}\right) = -\sum_{t=1}^{n} \dot{h}_t^{w_k}(\theta)\,\dot{h}_t^{w_l}(\theta)\sum_{j=1}^{M}\left(-\rho_{t,j}(\theta) + \frac{\delta_{t,j}^2(\theta)}{\Delta_{t,j}(\theta)}\right).$$

For the normal distribution, $\dot{\phi}(x) = -x\phi(x)$ so that $\sum_{j=1}^{M}\rho_{t,j}(\theta) = 0$ and we are left with

$$E_{\theta}\left(\frac{\partial^2 l_n(\theta)}{\partial w_k \partial w_l}|\mathcal{F}_{t-1}\right) = -\sum_{t=1}^{n} \dot{h}_t^{w_k}(\theta)\,\dot{h}_t^{w_l}(\theta)\sum_{j=1}^{M}\frac{\delta_{t,j}^2(\theta)}{\Delta_{t,j}(\theta)}.$$

Similarly, with $\dot{\delta}_{t,j,\rho}(x;\theta) = \partial\delta_{t,j}(\theta)/\partial\rho = -\rho_{t,j}(\theta)\,\rho^{-1}\bar{Y}_{t-1}^{s}$,

$$\begin{aligned}
\frac{\partial^2 l_n(\theta)}{\partial\rho^2} &= -\sum_{t=1}^{n}\rho^{-1}\bar{Y}_{t-1}^{s}\sum_{j=1}^{M}1\{Y_t=j\}\left(\frac{\dot{\delta}_{t,j}(\theta)}{\Delta_{t,j}(\theta)} + \frac{\delta_{t,j}^2(\theta)\,\rho^{-1}\bar{Y}_{t-1}^{s}}{\Delta_{t,j}^2(\theta)}\right)\\
&= -\sum_{t=1}^{n}\left(\rho^{-1}\bar{Y}_{t-1}^{s}\right)^2\sum_{j=1}^{M}1\{Y_t=j\}\left(-\frac{\rho_{t,j}(\theta)}{\Delta_{t,j}(\theta)} + \frac{\delta_{t,j}^2(\theta)}{\Delta_{t,j}^2(\theta)}\right),
\end{aligned}$$

giving

$$E_{\theta}\left(\frac{\partial^2 l_n(\theta)}{\partial\rho^2}|\mathcal{F}_{t-1}\right) = -\sum_{t=1}^{n}\left(\rho^{-1}\bar{Y}_{t-1}^{s}\right)^2\sum_{j=1}^{M}\frac{\delta_{t,j}^2(\theta)}{\Delta_{t,j}(\theta)}.$$

Because

$$\frac{\partial l_n(\theta)}{\partial w_k} = -\sum_{t=1}^{n}\dot{h}_t^{w_k}(\theta)\sum_{j=1}^{M}1\{Y_t=j\}\frac{\delta_{t,j}(\theta)}{\Delta_{t,j}(\theta)},$$

$$\begin{aligned}
\frac{\partial^2 l_n(\theta)}{\partial w_k\partial\mu_l} &= -\sum_{t=1}^{n}\dot{h}_t^{w_k}(\theta)\sum_{j=1}^{M}1\{Y_t=j\}\frac{1\{j=l\}-1\{j=l+1\}}{\Delta_{t,j}(\theta)}\\
&\quad\times\left(\dot{f}_{t,l}(\theta) - \frac{\delta_{t,j}(\theta)\,f_{t,l}(\theta)}{\Delta_{t,j}(\theta)}\right).
\end{aligned}$$

Thus,

$$E_\theta\left(\frac{\partial^2 l_n(\theta)}{\partial w_k \partial \mu_l}|\mathcal{F}_{t-1}\right)=\sum_{t=1}^{n}\dot{h}_t^{w_k} f_{t,l}(\theta)\left(\frac{\delta_{t,l}(\theta)}{\Delta_{t,l}(\theta)}-\frac{\delta_{t,l+1}(\theta)}{\Delta_{t,l+1}(\theta)}\right).$$

Also,

$$\begin{aligned}\frac{\partial^2 l_n(\theta)}{\partial w_k \partial \rho} &= -\sum_{t=1}^{n}\ddot{h}_t^{w_k,\rho}(\theta)\sum_{j=1}^{M}1\{Y_t=j\}\frac{\delta_{t,j}(\theta)}{\Delta_{t,j}(\theta)}\\ &\quad -\sum_{t=1}^{n}\dot{h}_t^{w_k}(\theta)\sum_{j=1}^{M}1\{Y_t=j\}\left(\frac{\dot{\delta}_{t,j,\rho}(\theta)}{\Delta_{t,j}(\theta)}+\frac{\delta_{t,j}^2(\theta)\rho^{-1}\bar{Y}_{t-1}^s}{\Delta_{t,j}^2(\theta)}\right),\end{aligned}$$

where $\ddot{h}_t^{w_k,\rho}(\theta)=\partial\dot{h}_t^{w_k}/\partial\rho$ and therefore,

$$E_\theta\left(\frac{\partial^2 l_n(\theta)}{\partial w_k \partial \rho}|\mathcal{F}_{t-1}\right)=-\rho^{-1}\sum_{t=1}^{n}\dot{h}_t^{w_k}(\theta)\bar{Y}_{t-1}^s\sum_{j=1}^{M}\frac{\delta_{t,j}^2(\theta)}{\Delta_{t,j}(\theta)}.$$

Finally,

$$\begin{aligned}\frac{\partial^2 l_n(\theta)}{\partial \mu_l \partial \rho} &= -\rho^{-1}\sum_{t=1}^{n}\bar{Y}_{t-1}^s\dot{f}_{t,l}(\theta)\left(\frac{1\{Y_t=l\}}{\Delta_{t,l}}-\frac{1\{Y_t=l+1\}}{\Delta_{t,l+1}}\right)\\ &\quad +\rho^{-1}\sum_{t=1}^{n}\bar{Y}_{t-1}^s f_{t,l}(\theta)\left(\frac{1\{Y_t=l\}}{\Delta_{t,l}^2}\delta_{t,l}-\frac{1\{Y_t=l+1\}}{\Delta_{t,l+1}^2}\delta_{t,l+1}\right).\end{aligned}$$

and

$$E_\theta\left(\frac{\partial^2 l_n(\theta)}{\partial \rho \partial \mu_l}|\mathcal{F}_{t-1}\right)=\rho^{-1}\sum_{t=1}^{n}\bar{Y}_{t-1}^s f_{t,l}(\theta)\left(\frac{\delta_{t,l}(\theta)}{\Delta_{t,l}(\theta)}-\frac{\delta_{t,l+1}(\theta)}{\Delta_{t,l+1}(\theta)}\right).$$

It is obvious that for any θ_k, θ_l, all the second-order derivatives may be written as

$$H_{n,\theta_j,\theta_k}(\theta)=\frac{1}{n}\sum_{t=1}^{n}z_t(\theta),$$

where, under Assumptions A1-A2, $z_t(\theta)$ are uniformly bounded. By Theorem 1(ii), $H_{n,\theta_j,\theta_k}(\theta)$ converges $w.p.1$ to a nonstochastic function, say $H_{\theta_j,\theta_k}(\theta)$.

Moreover, the Cauchy Schwartz inequality implies that the determinant of $E_\theta\left(H_{n,\theta_j,\theta_k}(\theta)\right)$ is non-negative for all n, θ_j, θ_k, with equality holding iff the terms in $\left(\dot{h}_t^{w_k}\right)_{1\leq k\leq K}$ are linearly dependent. This possibility is precluded by Assumption A4 and thus, the proof of Lemma 2 is complete.■

Proof of Theorem 4: It is straightforward to verify that the second-order Bartlett identity holds for all the second-order partial derivatives. The result $\text{plim}\left(H_{n,\theta_j,\theta_k}\left(\hat{\theta}_n\right)\right) = H_{\theta_j,\theta_k}(\theta_0)$ is a consequence of Theorem 4.1.5 of Amemiya (1985) because: (i) Theorem 1(ii) implies that the process is ergodic stationary; (ii) $\hat{\theta}_n \to_p \theta_0$; and (iii) $H_{\theta_j,\theta_k}(\theta)$ is continuous and uniformly bounded. This, together with Lemmas 1 and 2 and the mean value Theorem, as in equation (7.3.7) of Hayashi (2000), completes the proof. ■

3 Tables and Figures for the Simulations Section

Table 1. Simulated MLE point estimates for $\mu_0 = 0.5$ and $\lambda = 2$.

	n	250		500		1000		2000	
		$\hat{w}$	$\hat{\mu}$	$\hat{w}$	$\hat{\mu}$	$\hat{w}$	$\hat{\mu}$	$\hat{w}$	$\hat{\mu}$
	Mean	4.299	0.498	1.531	0.499	1.129	0.499	1.043	0.500
	Std	25.187	0.083	5.967	0.059	0.905	0.041	0.364	0.029
	Trim	1.605	0.498	1.177	0.500	1.077	0.499	1.031	0.500
$w_0 = 1$	Std Trim	2.014	0.073	0.774	0.051	0.484	0.035	0.305	0.025
	Median	1.033	0.498	0.998	0.499	0.995	0.499	1.000	0.500
	Q_1	0.488	0.441	0.609	0.459	0.706	0.473	0.792	0.480
	Q_3	1.983	0.557	1.581	0.540	1.384	0.525	1.235	0.519
	Mean	12.624	0.500	6.916	0.500	3.980	0.500	3.363	0.500
	Std	41.375	0.083	23.756	0.058	7.520	0.041	2.676	0.029
	Trim	6.640	0.500	4.019	0.500	3.438	0.500	3.199	0.500
$w_0 = 3$	Std Trim	13.148	0.073	3.588	0.051	1.629	0.036	0.980	0.025
	Median	2.914	0.501	3.041	0.501	3.097	0.501	3.063	0.499
	Q_1	1.546	0.444	1.932	0.459	2.230	0.472	2.440	0.481
	Q_3	5.971	0.557	4.727	0.540	4.256	0.528	3.775	0.518
	Mean	16.587	0.499	13.074	0.501	7.995	0.500	5.916	0.500
	Std	44.573	0.074	35.555	0.058	17.764	0.041	5.138	0.029
	Trim	10.481	0.499	8.097	0.501	6.146	0.500	5.463	0.500
$w_0 = 5$	Std Trim	20.133	0.063	10.217	0.050	3.696	0.035	1.999	0.025
	Median	4.815	0.500	5.118	0.502	5.177	0.500	5.037	0.500
	Q_1	2.593	0.451	3.117	0.463	3.636	0.473	3.967	0.480
	Q_3	9.791	0.546	8.946	0.538	7.429	0.527	6.550	0.520

Note: λ is the lag-length; Std is the standard deviation of the mean; Trim is the trimmed mean with 5% symmetric trimming; Std Trim is the standard deviation of the trimmed mean; Q_1 and Q_3 are the first and third quartiles, respectively.

Table 2. Simulated MLE point estimates for $\mu_0 = 0.5$ and $\lambda = 5$.

	n	250		500		1000		2000	
		$\hat{w}$	$\hat{\mu}$	$\hat{w}$	$\hat{\mu}$	$\hat{w}$	$\hat{\mu}$	$\hat{w}$	$\hat{\mu}$
	Mean	1.601	0.500	1.171	0.500	1.072	0.501	1.033	0.500
	Std	5.120	0.082	0.968	0.057	0.579	0.040	0.381	0.029
	Trim	1.253	0.500	1.111	0.500	1.048	0.501	1.024	0.500
$w_0 = 1$	Std Trim	1.218	0.071	0.713	0.050	0.462	0.035	0.324	0.025
	Median	0.977	0.500	0.992	0.500	0.995	0.500	0.997	0.500
	Q_1	0.396	0.445	0.584	0.463	0.693	0.474	0.777	0.481
	Q_3	1.790	0.555	1.533	0.537	1.351	0.527	1.248	0.519
	Mean	4.986	0.499	3.699	0.500	3.246	0.500	3.098	0.500
	Std	14.116	0.081	5.228	0.057	1.390	0.040	0.814	0.028
	Trim	3.733	0.499	3.339	0.500	3.173	0.500	3.078	0.500
$w_0 = 3$	Std Trim	2.990	0.071	1.661	0.050	1.045	0.035	0.698	0.024
	Median	2.915	0.499	2.986	0.500	3.028	0.501	3.004	0.500
	Q_1	1.745	0.444	2.098	0.461	2.373	0.473	2.520	0.482
	Q_3	4.843	0.555	4.219	0.539	3.881	0.526	3.586	0.518
	Mean	9.723	0.500	6.539	0.501	5.457	0.500	5.147	0.500
	Std	32.849	0.083	14.780	0.059	2.576	0.040	1.438	0.029
	Trim	6.529	0.500	5.621	0.501	5.296	0.500	5.096	0.500
$w_0 = 5$	Std Trim	5.770	0.072	2.934	0.051	1.845	0.034	1.201	0.025
	Median	4.858	0.502	4.963	0.501	4.913	0.500	4.928	0.500
	Q_1	2.990	0.444	3.442	0.462	3.838	0.473	4.151	0.481
	Q_3	8.037	0.555	7.048	0.540	6.437	0.526	5.943	0.520

Note: λ is the lag-length; Std is the standard deviation of the mean; Trim is the trimmed mean with 5% symmetric trimming; Std Trim is the standard deviation of the trimmed mean; Q_1 and Q_3 are the first and third quartiles, respectively.

Table 3. Simulated MLE point estimates for $\mu_0 = 0.5$ and $\lambda = 10$.

	n	250		500		1000		2000	
		$\hat{w}$	$\hat{\mu}$	$\hat{w}$	$\hat{\mu}$	$\hat{w}$	$\hat{\mu}$	$\hat{w}$	$\hat{\mu}$
	Mean	1.697	0.500	1.269	0.500	1.091	0.499	1.035	0.499
	Std	3.620	0.081	1.505	0.057	0.794	0.040	0.504	0.028
	Trim	1.362	0.500	1.155	0.500	1.052	0.499	1.019	0.499
$w_0 = 1$	Std Trim	1.731	0.071	0.965	0.050	0.602	0.035	0.411	0.024
	Median	0.953	0.498	0.986	0.499	0.986	0.498	0.975	0.500
	Q_1	0.219	0.445	0.433	0.461	0.587	0.471	0.701	0.480
	Q_3	1.991	0.554	1.729	0.539	1.454	0.526	1.297	0.518
	Mean	4.766	0.500	3.579	0.500	3.265	0.499	3.086	0.499
	Std	10.682	0.081	3.009	0.056	1.678	0.040	1.016	0.028
	Trim	3.878	0.500	3.345	0.499	3.178	0.499	3.058	0.499
$w_0 = 3$	Std Trim	3.852	0.071	2.035	0.049	1.297	0.035	0.868	0.024
	Median	2.796	0.500	2.905	0.500	3.005	0.498	2.940	0.499
	Q_1	1.328	0.446	1.827	0.462	2.203	0.471	2.381	0.480
	Q_3	5.216	0.555	4.513	0.538	4.019	0.526	3.663	0.519
	Mean	7.608	0.500	5.824	0.499	5.399	0.498	5.116	0.499
	Std	15.915	0.081	4.416	0.0579	2.524	0.041	1.542	0.028
	Trim	6.182	0.499	5.506	0.499	5.280	0.498	5.068	0.499
$w_0 = 5$	Std Trim	5.574	0.070	3.117	0.050	2.014	0.035	1.289	0.025
	Median	4.654	0.498	4.852	0.499	4.984	0.498	4.947	0.499
	Q_1	2.234	0.445	3.215	0.462	3.728	0.471	4.056	0.480
	Q_3	8.383	0.554	7.151	0.538	6.5777	0.525	5.967	0.518

Note: λ is the lag-length; Std is the standard deviation of the mean; Trim is the trimmed mean with 5% symmetric trimming; Std Trim is the standard deviation of the trimmed mean; Q_1 and Q_3 are the first and third quartiles, respectively.

Table 4. Simulated MLE point estimates for $\mu_0 = 0.3$ and $\lambda = 2$.

	n	250		500		1000		2000	
		$\hat{w}$	$\hat{\mu}$	$\hat{w}$	$\hat{\mu}$	$\hat{w}$	$\hat{\mu}$	$\hat{w}$	$\hat{\mu}$
	Mean	3.572	0.297	2.110	0.298	1.143	0.300	1.059	0.300
	Std	21.488	0.085	12.308	0.060	0.755	0.042	0.409	0.030
	Trim	1.603	0.297	1.215	0.298	1.090	0.300	1.043	0.300
$w_0 = 1$	Std Trim	2.115	0.074	0.868	0.052	0.494	0.037	0.326	0.026
	Median	1.010	0.298	1.022	0.299	0.998	0.301	1.003	0.299
	Q_1	0.483	0.238	0.608	0.256	0.707	0.271	0.791	0.280
	Q_3	1.893	0.354	1.561	0.337	1.394	0.329	1.267	0.320
	Mean	11.653	0.295	6.481	0.298	4.281	0.300	3.338	0.300
	Std	38.079	0.085	20.200	0.059	10.253	0.042	1.929	0.030
	Trim	6.241	0.295	4.203	0.298	3.483	0.300	3.204	0.300
$w_0 = 3$	Std Trim	10.761	0.074	3.914	0.052	1.890	0.037	1.078	0.026
	Median	3.008	0.297	2.982	0.299	2.990	0.300	2.997	0.301
	Q_1	1.620	0.238	1.942	0.258	2.177	0.272	2.352	0.280
	Q_3	6.039	0.353	4.882	0.339	4.316	0.329	3.849	0.321
	Mean	20.497	0.298	12.716	0.301	8.237	0.300	5.975	0.300
	Std	53.766	0.0842	32.829	0.059	18.844	0.041	7.266	0.030
	Trim	14.160	0.298	8.172	0.301	6.144	0.300	5.451	0.300
$w_0 = 5$	Std Trim	34.347	0.073	10.250	0.051	3.982	0.036	2.023	0.026
	Median	4.808	0.301	4.961	0.302	5.007	0.299	5.013	0.300
	Q_1	2.327	0.242	3.076	0.261	3.593	0.273	3.978	0.281
	Q_3	10.661	0.356	8.712	0.339	7.306	0.328	6.481	0.320

Note: λ is the lag-length; Std is the standard deviation of the mean; Trim is the trimmed mean with 5% symmetric trimming; Std Trim is the standard deviation of the trimmed mean; Q_1 and Q_3 are the first and third quartiles, respectively.

Table 5. Simulated MLE point estimates for $\mu_0 = 0.3$ and $\lambda = 5$.

	n	250		500		1000		2000	
		$\hat{w}$	$\hat{\mu}$	$\hat{w}$	$\hat{\mu}$	$\hat{w}$	$\hat{\mu}$	$\hat{w}$	$\hat{\mu}$
	Mean	2.125	0.299	1.196	0.300	1.086	0.300	1.036	0.300
	Std	17.826	0.084	1.407	0.059	0.614	0.041	0.374	0.029
	Trim	1.235	0.299	1.109	0.300	1.057	0.300	1.028	0.300
$w_0 = 1$	Std Trim	1.211	0.073	0.713	0.052	0.462	0.036	0.320	0.026
	Median	0.952	0.299	0.988	0.300	0.999	0.300	1.011	0.300
	Q_1	0.413	0.244	0.586	0.260	0.704	0.272	0.776	0.281
	Q_3	1.762	0.356	1.516	0.341	1.369	0.329	1.263	0.320
	Mean	6.215	0.297	3.922	0.299	3.244	0.299	3.111	0.299
	Std	24.331	0.084	9.208	0.059	1.410	0.041	0.883	0.029
	Trim	3.928	0.297	3.407	0.300	3.179	0.299	3.083	0.299
$w_0 = 3$	Std Trim	3.457	0.072	1.836	0.051	1.129	0.035	0.734	0.025
	Median	2.990	0.299	3.001	0.299	2.994	0.299	2.982	0.299
	Q_1	1.682	0.243	2.047	0.260	2.307	0.272	2.504	0.279
	Q_3	4.977	0.353	4.413	0.338	3.917	0.326	3.621	0.319
	Mean	9.562	0.298	6.755	0.300	5.385	0.301	5.201	0.301
	Std	29.605	0.084	16.343	0.059	2.531	0.041	1.504	0.029
	Trim	6.369	0.298	5.658	0.300	5.222	0.301	5.139	0.301
$w_0 = 5$	Std Trim	5.269	0.074	3.076	0.051	1.776	0.035	1.192	0.025
	Median	4.990	0.300	4.960	0.302	4.928	0.301	5.001	0.301
	Q_1	2.931	0.239	3.392	0.260	3.892	0.274	4.207	0.282
	Q_3	8.091	0.355	7.241	0.340	6.340	0.327	5.991	0.321

Note: λ is the lag-length; Std is the standard deviation of the mean; Trim is the trimmed mean with 5% symmetric trimming; Std Trim is the standard deviation of the trimmed mean; Q_1 and Q_3 are the first and third quartiles, respectively.

Table 6. Simulated MLE point estimates for $\mu_0 = 0.3$ and $\lambda = 10$.

	n	250		500		1000		2000	
		$\hat{w}$	$\hat{\mu}$	$\hat{w}$	$\hat{\mu}$	$\hat{w}$	$\hat{\mu}$	$\hat{w}$	$\hat{\mu}$
	Mean	2.119	0.296	1.308	0.297	1.129	0.298	1.048	0.299
	Std	9.604	0.082	1.612	0.056	0.846	0.041	0.526	0.029
	Trim	1.460	0.296	1.188	0.298	1.090	0.298	1.033	0.299
$w_0 = 1$	Std Trim	1.949	0.071	1.055	0.049	0.661	0.035	0.441	0.025
	Median	0.980	0.296	1.007	0.298	1.011	0.298	0.987	0.299
	Q_1	0.164	0.242	0.416	0.259	0.582	0.272	0.695	0.280
	Q_3	2.086	0.350	1.751	0.336	1.525	0.325	1.340	0.318
	Mean	5.723	0.297	3.597	0.298	3.233	0.299	3.093	0.299
	Std	18.085	0.081	3.128	0.056	1.659	0.040	1.029	0.029
	Trim	4.085	0.297	3.351	0.299	3.162	0.299	3.069	0.299
$w_0 = 3$	Std Trim	4.199	0.071	2.086	0.049	1.331	0.035	0.885	0.025
	Median	2.835	0.296	2.916	0.299	2.966	0.299	2.976	0.299
	Q_1	1.354	0.245	1.825	0.260	2.120	0.272	2.364	0.280
	Q_3	5.545	0.351	4.463	0.337	4.038	0.327	3.731	0.317
	Mean	8.488	0.297	5.931	0.299	5.400	0.299	5.141	0.299
	Std	21.108	0.081	4.701	0.056	2.602	0.041	1.588	0.029
	Trim	6.438	0.297	5.574	0.299	5.272	0.299	5.096	0.299
$w_0 = 5$	Std Trim	6.131	0.071	3.254	0.049	2.050	0.035	1.330	0.025
	Median	4.682	0.297	4.905	0.300	4.975	0.299	5.031	0.299
	Q_1	2.201	0.244	3.139	0.261	3.676	0.272	4.033	0.279
	Q_3	8.578	0.351	7.360	0.337	6.605	0.3267	6.019	0.318

Note: λ is the lag-length; Std is the standard deviation of the mean; Trim is the trimmed mean with 5% symmetric trimming; Std Trim is the standard deviation of the trimmed mean; Q_1 and Q_3 are the first and third quartiles, respectively.

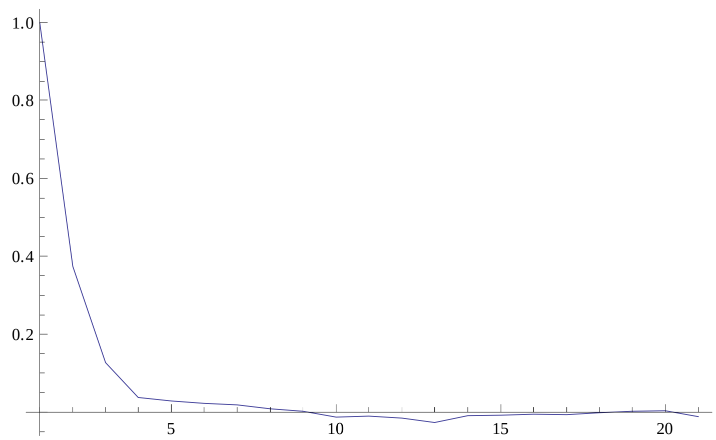

Figure 1. Correlogram of the process in the case $\lambda = 1$, $M = 2$, $n = 10000$.

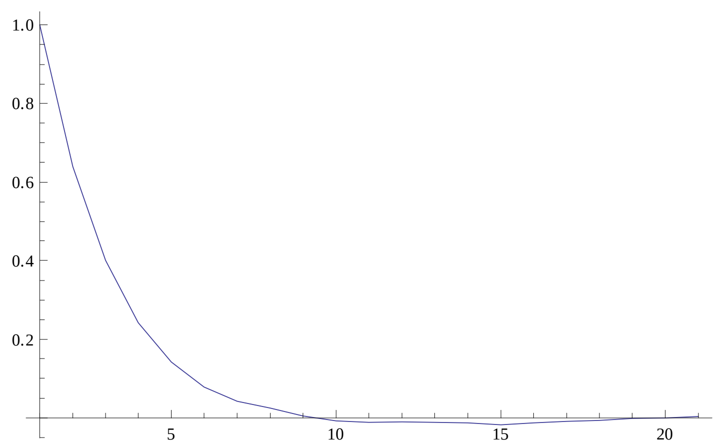

Figure 2. Correlogram of the process in the case $\lambda = 1$, $M = 3$, $n = 10000$.

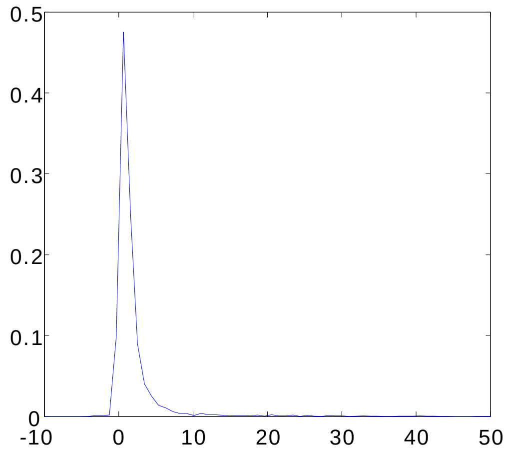

Figure 3. Kernel density estimate for $\hat{w}$, $w_0 = 1$, $\mu_0 = 0.3$, $\lambda = 2$, $n = 250$.

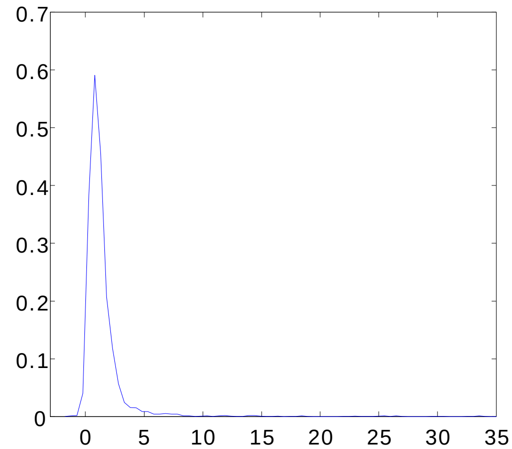

Figure 4. Kernel density estimate for $\hat{w}$, $w_0 = 1$, $\mu_0 = 0.3$, $\lambda = 2$, $n = 500$.

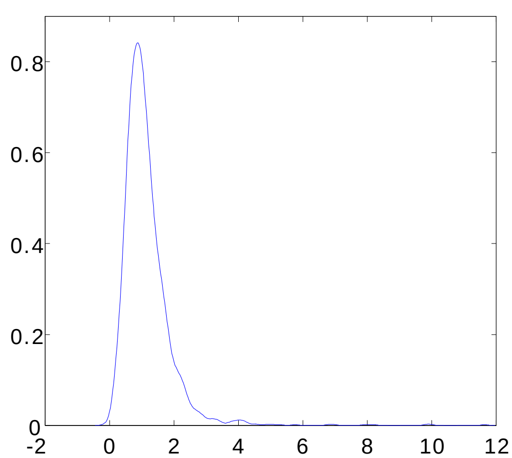

Figure 5. Kernel density estimate for $\hat{w}$, $w_0 = 1$, $\mu_0 = 0.3$, $\lambda = 2$, $n = 1000$.

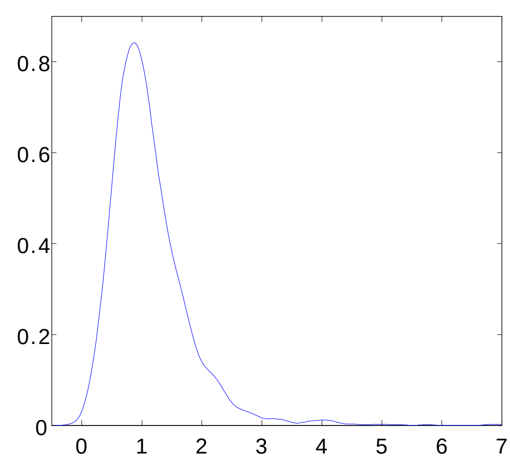

Figure 6. Kernel density estimate for $\hat{w}$, $w_0 = 1$, $\mu_0 = 0.3$, $\lambda = 2$, $n = 2000$.

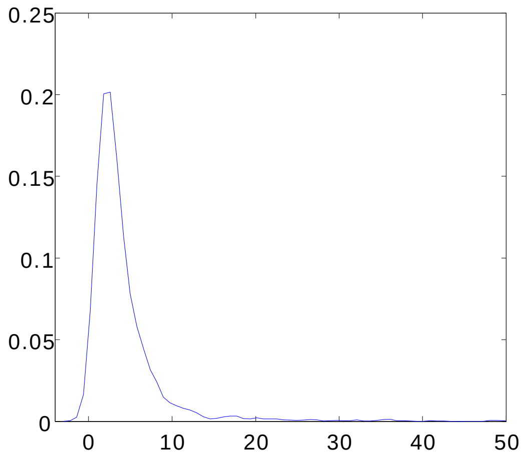

Figure 7. Kernel density estimate for $\hat{w}$, $w_0 = 3$, $\mu_0 = 0.5$, $\lambda = 5$, $n = 250$.

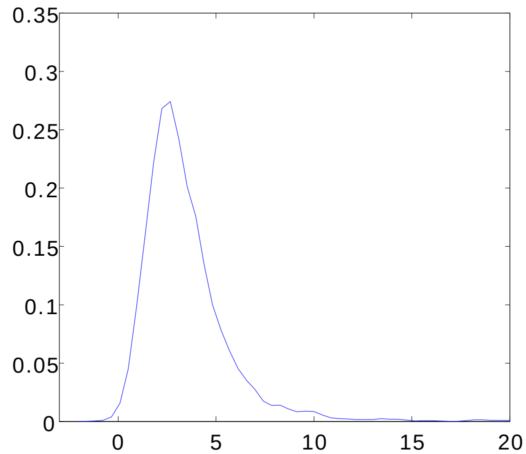

Figure 8. Kernel density estimate for $\hat{w}$, $w_0 = 3$, $\mu_0 = 0.5$, $\lambda = 5$, $n = 500$.

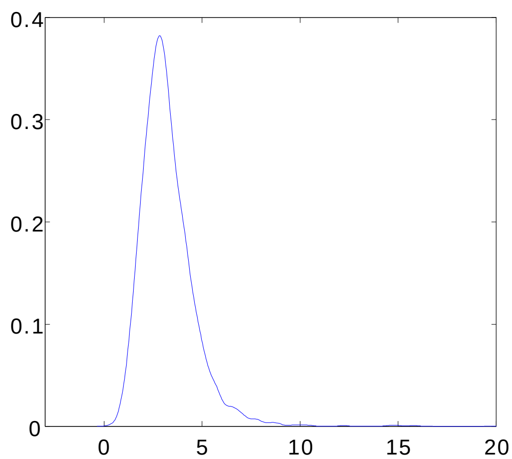

Figure 9. Kernel density estimate for $\hat{w}$, $w_0 = 3$, $\mu_0 = 0.5$, $\lambda = 5$, $n = 1000$.

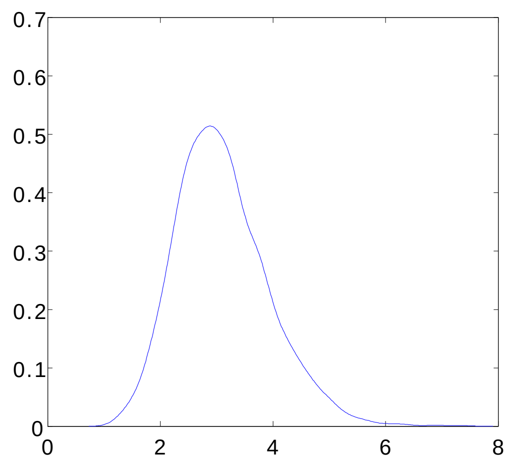

Figure 10. Kernel density estimate for $\hat{w}$, $w_0 = 3$, $\mu_0 = 0.5$, $\lambda = 5$, $n = 2000$.

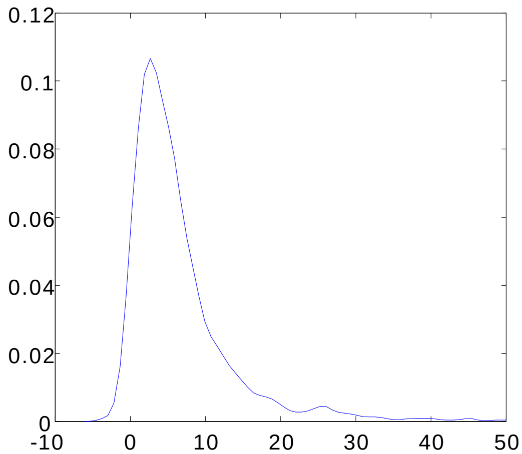

Figure 11. Kernel density estimate for $\hat{w}$, $w_0 = 5$, $\mu_0 = 0.5$, $\lambda = 10$, $n = 250$.

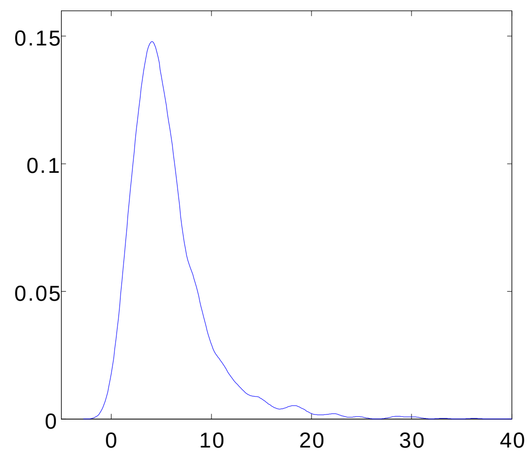

Figure 12. Kernel density estimate for $\hat{w}$, $w_0 = 5$, $\mu_0 = 0.5$, $\lambda = 10$, $n = 500$.

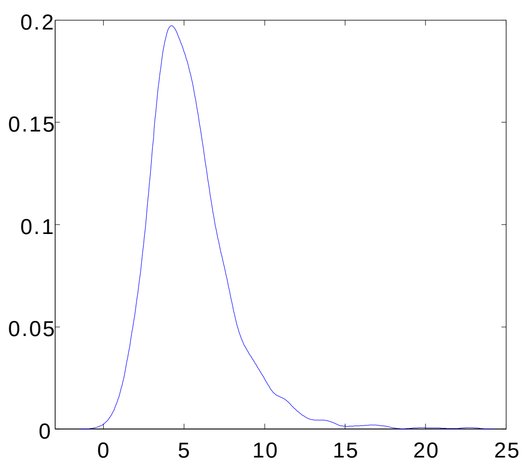

Figure 13. Kernel density estimate for $\hat{w}$, $w_0 = 5$, $\mu_0 = 0.5$, $\lambda = 10$, $n = 1000$.

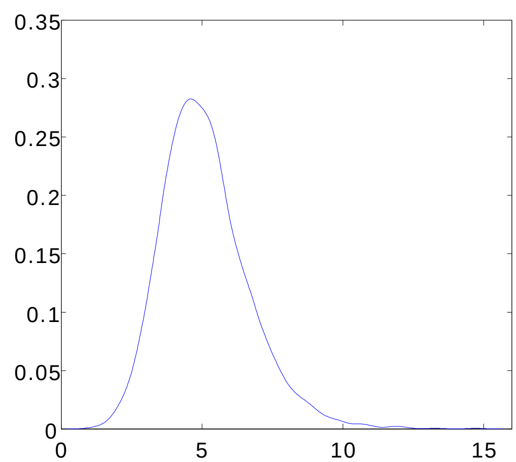

Figure 14. Kernel density estimate for $\hat{w}$, $w_0 = 5$, $\mu_0 = 0.5$, $\lambda = 10$, $n = 2000$.