Appendix to

The block criterion for multiscale inference about a density, with applications to other multiscale problems

by Kaspar Rufibach and Guenther Walther

Published in Journal of Computational and Graphical Statistics

9 Proofs

Proof of Theorem 2: For $\mathcal{I}_{jk} \in \boldsymbol{\mathcal{I}}_{app}(l)$ there are not more than $\lceil n/d_l \rceil$ possible values of j due to condition (a) in Table 1. For each such j there are not more than $\lceil m_l/d_l \rceil$ possible values of k due to (a) and (b). Hence $\#\boldsymbol{\mathcal{I}}_{app}(l) \leq \lceil n/d_l \rceil \lceil m_l/d_l \rceil \leq 4nM/D^2$. The assertion about the cardinality of $\boldsymbol{\mathcal{I}}_{app}$ now follows as $l_{max} \leq \log_2(n/M)$.

To prove the second claim we first first establish an analogous result to Lemma 7.5 in Dümbgen and Walther (2008):

Lemma 1 *Let f be a density, I be a bounded open interval and set $\delta := \int_I f$. Suppose*

1. $\inf_I f' \, |I|^2/\sqrt{\delta} \geq C\sqrt{\log(e/\delta)/n}$
2. $C \geq \frac{\sqrt{24}}{(1-\epsilon)^2\sqrt{1-\gamma-2/(n\delta)}}\Big(1 + \frac{\kappa_n(\alpha)+\eta}{\sqrt{2\log(e/\delta)}} + \frac{\gamma+2/(n\delta)}{2\log(e/\delta)}\Big)$
   for certain numbers $\epsilon \in (0,1)$, $\gamma \in (0, 1/2]$ and $\eta > 0$.
3. $C\epsilon^2\sqrt{\log(e/\delta)} \geq \frac{32D}{\sqrt{M(1-\gamma)}}$

Then

$$\begin{aligned}
&\mathbb{P}\Big(\mathcal{D}^+_{app}(\alpha) \textit{ contains no interval } J \subset I\Big) \\
&\leq \quad 2\exp(-n\delta\gamma^2/2) + 2\exp\big(-C(1-\gamma)\sqrt{n\delta\log(e/\delta)}\,\epsilon^2/64\big) + \exp(-\eta^2/2)\,.
\end{aligned}$$

The main difference to Lemma 7.5 in Dümbgen and Walther (2008) is the additional condition 3 which is necessitated by the reduced number of intervals under consideration.

Proof of Lemma 1. Write $I = (a, b)$ and define $I_{(\ell)} := (a, a+\epsilon|I|/2]$, $I_{(r)} := [b-\epsilon|I|/2, b)$. It was shown in the proof of Lemma 7.5 in Dümbgen and Walther (2008) that

$$\frac{F(I_{(r)})}{F(I)} \geq \frac{F(I_{(\ell)})}{F(I)} \geq C\sqrt{\frac{\log(e/\delta)}{n\delta}}\,\epsilon^2/8\,. \tag{6}$$

Set $N := \#\{i : X_i \in I\}$. If $N \geq M + 2D$ then we can define l to be the smallest integer such that $m_l + 2d_l \leq N$. This implies $l_{max} - l \leq \log_2(N/M)$ and hence

$$d_l \leq \text{round}(D(N/M)^{1/2})\,. \tag{7}$$

Furthermore we can conclude that there is an interval $\mathcal{I}_{jk} \in \boldsymbol{\mathcal{I}}_{app}$ with $\mathcal{I}_{jk} \subset I$ and $X_{(j)}$ is the first or second point on the d_l-spacing that falls into I, and $X_{(k)}$ is the last or second-to-last observation on the d_l-spacing that falls into I. Hence for $\gamma \in (0, 1/2]$ we have on the set $\{N \geq (1-\gamma)n\delta\}$:

$$\begin{aligned} & \mathrm{IP}(|\mathcal{I}_{jk}|/|I| \leq 1 - \epsilon | N) \\ \leq \; & \mathrm{IP}(\text{less than } 2d_l \text{ observations in } I_{(\ell)}|N) + \mathrm{IP}(\text{less than } 2d_l \text{ observations in } I_{(r)}|N) \end{aligned} \tag{8}$$

It follows from the general Hoeffding inequality, see A.6.1 in van der Vaart and Wellner (1996), that $\mathrm{IP}(N \leq (1-\gamma)n\delta) \leq \exp(-n\delta\gamma^2/2)$ and that (8) is not larger than

$$2\exp\Bigl(-\frac{NF(I_{(\ell)})}{8F(I)}\Bigr) \quad \text{provided that} \quad \frac{2d_lF(I)}{NF(I_{(\ell)})} \; < \; 1/2\,.$$

But (6), (7) and $N \geq (1-\gamma)n\delta$ imply

$$\frac{2d_lF(I)}{NF(I_{(\ell)})} \; \leq \; \frac{16D}{C\epsilon^2\sqrt{M(1-\gamma)\log(e/\delta)}}\,.$$

Hence (6) gives

$$\begin{aligned} & \mathrm{IP}(|\mathcal{I}_{jk}|/|I| \leq 1 - \epsilon) \\ \leq \;\; & 2\exp\bigl(-C(1-\gamma)\sqrt{n\delta\log(e/\delta)}\,\epsilon^2/64\bigr) + \exp(-n\delta\gamma^2/2) \end{aligned}$$

provided that Assumption 3 of the lemma holds. Now the assertion of the lemma obtains by following the proof of Lemma 7.5 in Dümbgen and Walther (2008) and by using the fact that the $\kappa_n(\alpha)$ used for $\mathcal{D}^+_{app}(\alpha)$ is not larger than the $\kappa_n(\alpha)$ used for $\mathcal{D}^+(\alpha)$ in (2) as $\boldsymbol{\mathcal{I}}_{app} \subset \{\mathcal{I}_{jk}, 1 \leq j < k \leq n\}$. □

To prove Theorem 2 we can now proceed similarly as in the proof of Theorem 4.2(iii) of

Dümbgen and Walther (2008): Set

$$\begin{aligned}\epsilon = \epsilon_n &:= \sqrt{32/C_n}(\log(e/\delta_n))^{-1/4}\,,\\ \gamma = \gamma_n &:= \sqrt{2D}(n\delta_n)^{-1/2}\,,\\ \eta = \eta_n &:= b_n\end{aligned}$$

where C_n and b_n are given in the statement of the theorem, $\delta_n := F_n(I_n)$, and $D > 0$ is an arbitrary fixed number. One easily verifies that $\epsilon_n \in (0, 1/4)$ for n large enough. It is shown in the proof of Lemma 7.5 in Dümbgen and Walther (2008) that $n\delta_n \geq \log n$. Hence $\gamma_n \leq 1/2$ for $n \geq \exp(8D)$ and the bound given by Lemma 1 is not larger than

$$2\exp(-D) \;+\; 2\exp(-\sqrt{\log n}/4) \;+\; \exp(-b_n^2/2)$$

provided that

$$C_n\,\epsilon_n^2 \;\geq\; \frac{32}{\sqrt{\log(e/\delta_n)}}$$

and

$$C_n \;\geq\; \sqrt{24}\Big(1 + k_1\epsilon_n + k_2(\gamma_n + \log(n)^{-1})\Big)\Big(1 + \gamma_n + \log(n)^{-1} + \frac{\kappa_n(\alpha) + \eta_n}{\sqrt{2\log(e/\delta_n)}}\Big)$$

for certain universal constants k_1 and k_2. With our choices for C_n, ϵ_n, γ_n and η_n, these conditions are satisfied for sufficiently large n. □

For the proof Theorem 3 we need the following result:

Proposition 1 *Let $\kappa > 0$. Then there is a $\delta_0 = \delta_0(\kappa) \in (0, 1)$ so that for all $\delta^2 \in (0, \delta_0^2]$*

$$\mathbb{P}\Big(\max_{1\leq j<k\leq n:\, \frac{k-j-1}{n}\in(\frac{\delta^2}{2},\delta^2]} |T_{jk}(\boldsymbol{U})| \;>\; \sqrt{2\log\frac{en}{k-j}} + \kappa\Big) \;\leq\; \exp\Big(-\sqrt{\log\frac{1}{\delta^2}}\,\kappa/2\Big)$$

Proof of Proposition 1: For $\delta \in (0, 1)$ we will consider $\mathcal{T}_n(\delta) := \Big\{(\frac{j}{n}, \frac{k}{n}) : \frac{k-j-1}{n} \in (\frac{\delta^2}{2}, \delta^2],\ 1 \leq j < k \leq n\Big\}$ with the distance $\rho\Big((s_1, s_2), (t_1, t_2)\Big) := (|t_1 - s_1| + |t_2 - s_2|)^{1/2}$

and define $\mathcal{T}_n^*(\delta)$ to be a maximal subset of $\mathcal{T}_n(\delta)$ such that $\rho(s,t) > \delta(\log(e/\delta))^{-2}$ for different $s, t \in \mathcal{T}_n^*(\delta)$. Thus for every $t \in \mathcal{T}_n(\delta)$ there exists a $t' \in \mathcal{T}_n^*(\delta)$ with $\rho(t,t') \le \delta(\log(e/\delta))^{-2}$. Hence the claim follows once we show that for some $\delta_0 = \delta_0(\kappa) \in (0,1)$ and all $\delta \in (0, \delta_0]$:

$$(9) \quad \mathbb{P}\Big(\mathcal{A} := \Big\{ \max_{(\frac{j}{n},\frac{k}{n}) \in \mathcal{T}_n^*(\delta)} |T_{jk}(\boldsymbol{U})| \; > \; \sqrt{2\log\frac{en}{k-j}} + \kappa/2 \Big\}\Big) \; \le \; \frac{1}{3}\exp\Big(-\sqrt{\log\frac{1}{\delta^2}}\,\kappa/2\Big)$$

and, writing $t := (\frac{j}{n}, \frac{k}{n})$ and $t' := (\frac{j'}{n}, \frac{k'}{n})$,

(10)
$$\mathbb{P}\Big(\max_{t \in \mathcal{T}_n(\delta), t' \in \mathcal{T}_n^*(\delta) :\, \rho(t,t') \le \frac{\delta}{(\log(e/\delta))^2}} |T_{jk}(\boldsymbol{U}) - T_{j'k'}(\boldsymbol{U})| \; > \; \kappa/2 \Big) \; \le \; \frac{2}{3}\exp\Big(-\sqrt{\log\frac{1}{\delta^2}}\,\kappa/2\Big).$$

If $\delta^2 < 1/n$, then (9) holds as $\mathcal{T}_n(\delta)$ is empty. For $\delta^2 \ge 1/n$ the arguments used in the proof of Lemma 11 in Dümbgen and Walther (2007) show that $\#\mathcal{T}_n^*(\delta) \le 24(\log(e/\delta))^8\delta^{-2}$ and that $T_{jk}(\boldsymbol{U})$ has subgaussian tails with scale factor 1. Hence

$$\begin{aligned}
\mathbb{P}(\mathcal{A}) \;&\le\; 24(\log(e/\delta))^8\delta^{-2}\,\mathbb{P}\Big(|T_{jk}(\boldsymbol{U})| > \sqrt{2\log\frac{e}{\delta^2 + 1/n}} + \kappa/2\Big) \\
&\le\; 48(\log(e/\delta))^8\delta^{-2}\exp\Big(-\Big(\sqrt{2\log\frac{e}{2\delta^2}} + \kappa/2\Big)^2/2\Big) \\
&\le\; 36(\log(e/\delta))^8\exp\Big(-\sqrt{2\log\frac{1}{\delta^2}}\,\kappa/2\Big) \\
&\le\; \frac{1}{3}\exp\Big(-\sqrt{\log\frac{1}{\delta^2}}\,\kappa/2\Big)
\end{aligned}$$

for $\delta \le \delta_0(\kappa)$. As for (10), it follows from Theorem 7 in Dümbgen and Walther (2007) together with the subsequent Remarks 2 and 3 and Lemma 11 that

$$\mathbb{P}\Big(\mathcal{B} := \Big\{ \sup_{t,t' \in \mathcal{T}_n(\delta) : \rho(t,t') \le \frac{\delta}{(\log(e/\delta))^2}} \frac{\Big|\sqrt{\frac{k-j-1}{n}}T_{jk}(\boldsymbol{U}) - \sqrt{\frac{k'-j'-1}{n}}T_{j'k'}(\boldsymbol{U})\Big|}{\rho(t,t')\log(e/\rho(t,t'))} > Q \Big\}\Big) \le K\frac{\delta}{(log(e/\delta))^2}$$

for certain constants $K, Q > 0$. The latter bound is not larger than $\frac{1}{3}\exp\Big(-\sqrt{\log\frac{1}{\delta^2}}\,\kappa/2\Big)$ for δ small enough, depending only on K and Q. Elementary calculations show $\Big|\sqrt{\frac{k-j-1}{n}} -$

$\sqrt{\frac{k'-j'-1}{n}}\Big| \le \rho(t,t')$. Hence on $\mathcal{A}^c \cap \mathcal{B}^c$:

$$\begin{aligned}
\Big|T_{jk}(\boldsymbol{U}) - T_{j'k'}(\boldsymbol{U})\Big| \;\le\; & \frac{\Big|\sqrt{\frac{k-j-1}{n}}T_{jk}(\boldsymbol{U}) - \sqrt{\frac{k'-j'-1}{n}}T_{j'k'}(\boldsymbol{U})\Big|}{\sqrt{\frac{k-j-1}{n}}} + \\
& + \Big|T_{j'k'}(\boldsymbol{U})\Big| \times \frac{\Big|\sqrt{\frac{k-j-1}{n}} - \sqrt{\frac{k'-j'-1}{n}}\Big|}{\sqrt{\frac{k-j-1}{n}}} \\
\le\; & \frac{Q\rho(t,t')\log(e/\rho(t,t'))}{\sqrt{\frac{k-j-1}{n}}} + \Big(\sqrt{2\log\frac{en}{k-j}} + \frac{\kappa}{2}\Big)\frac{\rho(t,t')}{\sqrt{\frac{k-j-1}{n}}}.
\end{aligned}$$

Employing $\frac{k-j-1}{n} \ge \delta^2/2$ and $\rho(t,t') \le \delta(\log(e/\delta))^{-2}$ one sees that this sum is smaller than $\kappa/2$ for δ small enough, depending on Q and κ. □

Proof of Theorem 3: The probability in (3) is not larger than $\sum_{l=1}^{l_{max}} \frac{\tilde{\alpha}}{(A+l)^2} \le \tilde{\alpha}\pi^2/6$. Thus $\tilde{\alpha} \ge 6\alpha/\pi^2$ by the definition of $\tilde{\alpha}$. Thus the probability of rejection in the l-th block is at least $\frac{6\alpha}{\pi^2(A+l)^2}$. First we consider small scales:

Let $\mathcal{I}_{jk} \in \boldsymbol{\mathcal{I}}_{app}(l)$ with $l \ge l_0$, where l_0 will be specified below. Thus $\frac{k-j-1}{n} \in (\delta^2/2, \delta^2]$, where $\delta^2 := (2m_l - 1)/n$. Together with $l_{max} \le \log_2(n/M)$ we obtain $l \le \log_2(1/\delta^2) + 1$. If we require $l_0 \ge \log_2(2/\delta_0^2)$, where δ_0^2 is given by Proposition 1 with $\kappa := \liminf_n \kappa_n(\alpha) > 0$, then we obtain $\delta^2 \le \delta_0^2$. Thus said Proposition gives

$$\begin{aligned}
\mathbb{P}\Big(\max_{1\le j<k\le n:\,\frac{k-j-1}{n}\in(\delta^2/2,\delta^2]} |T_{jk}(\boldsymbol{U})| \ge \sqrt{2\log\frac{en}{k-j}} + \kappa_n(\alpha)\Big) & \\
\le \exp\Big(-\sqrt{\log(1/\delta^2)}\,\kappa/2\Big) & \\
\le \exp\Big(-\sqrt{l-1}\,\kappa/4\Big) & \\
\le \frac{6\alpha}{\pi^2(A+l)^2} & \\
\le \frac{\tilde{\alpha}}{(A+l)^2} &
\end{aligned}$$

for $l \ge l_0(\delta_0, A, \kappa)$. Hence for intervals $\mathcal{I}_{jk} \in \boldsymbol{\mathcal{I}}_{app}(l)$ with $l \ge l_0$, the critical value $q_l\Big(\tilde{\alpha}/(A+l)^2\Big)$ is not larger than $\sqrt{2\log\frac{en}{k-j}} + \kappa_n(\alpha)$ and thus the power is not smaller than in the case of Theorem 2.

Next we consider large scales $l < l_0$, i.e. intervals I with $\int_I f \ge \delta_0 > 0$ for some δ_0. In that case the claim of Theorem 3 follows because $\max_{l<l_0} q_l\Big(\tilde{\alpha}/(A+l)^2\Big)$ stays bounded in n. In more detail: Lemma 1 continues to hold for such intervals with $\mathcal{D}^+_{block}$ in place of $\mathcal{D}^+_{app}$ if we replace the term $\exp(-\eta^2/2)$ by $\exp(-\eta^2/50)$ and Condition 2 by

$$2'.\ C \ \geq\ \frac{\sqrt{24}\Big(\max_{l<l_0} q_l(\tilde{\alpha}/(A+l)^2)+\eta/5\Big)}{(1-\epsilon)^2\sqrt{1-\gamma-2/(n\delta)}\sqrt{2\log(e/\delta)}}.$$

Thus, as in the proof of Theorem 2, we need to check that for the choices of ϵ_n, γ_n and η_n there, C_n satisfies

$$C_n \ \geq\ \sqrt{24}\Big(1+k_1\epsilon_n+k_2(\gamma_n+\log(n)^{-1})\Big)\frac{\max_{l<l_0} q_l(\tilde{\alpha}/(A+l)^2)+\eta_n/5}{\sqrt{2\log(e/F_n(I_n))}}$$

for certain universal constants k_1 and k_2. As $\eta_n = b_n$ this inequality holds for large enough n, provided that $\max_{l<l_0} q_l(\tilde{\alpha}/(A+l)^2)$ stays bounded in n. But that follows from Theorem 3.1 in Dümbgen and Walther (2008) together with $\frac{\tilde{\alpha}}{(A+l)^2} \geq \frac{6\alpha}{\pi^2(A+l_0)^2}$ and the fact that $\frac{n}{k-j}$ stays bounded for $l < l_0$. $\square$

References

DÜMBGEN, L. AND WALTHER, G. (2007). Multiscale inference about a density. University of Bern, IMSV, Technical Report 56.

DÜMBGEN, L. AND WALTHER, G. (2008). Multiscale inference about a density. *Ann. Statist.* ***36***, 1758–1785.

VAN DER VAART, A.W. AND WELLNER, J.A. (1996). *Weak Convergence and Empirical Processes.* Springer, New York