Supplementary material for *Detecting heteroskedasticity in nonparametric regression using weighted empirical processes*

Justin Chown[1] and Ursula U. Müller[2]

The proof of Theorem 2 follows in the same spirit as the proof of Theorem 1 in the main article. The primary difference between the proofs is in how one handles the technical details regarding the estimated weights $\tilde{W}_j$, which are now more complicated due to the presence of the scale estimator $\hat{\sigma}$. Recall that the test statistic in Theorem 2 is given by

$$\tilde{T}_n = \sup_{t \in \mathbb{R}} \left| n^{-1/2} \sum_{j=1}^{n} \tilde{W}_j \mathbf{1}\left[\hat{\varepsilon}_j \le t\right] \right|$$

with estimated weights $\tilde{W}_j$ as specified in equation (2.4) in the article,

$$\tilde{W}_j = \left\{ \hat{\sigma}(X_j) - \overline{\hat{\sigma}} \right\} \Big/ \left[\frac{1}{n} \sum_{k=1}^{n} \left\{ \hat{\sigma}(X_k) - \overline{\hat{\sigma}} \right\}^2 \right]^{1/2} \tag{1}$$

for $j = 1, \ldots, n$, writing $\overline{\hat{\sigma}} = n^{-1} \sum_{j=1}^{n} \hat{\sigma}(X_j)$.

Sketch of proof of Theorem 2. Consider the statistic

$$\sup_{t \in \mathbb{R}} \left| n^{-1/2} \sum_{j=1}^{n} W_n(X_j) \Big\{ \mathbf{1}\left[\varepsilon_j \le t\right] - F(t) \Big\} \right|, \tag{2}$$

with weights $W_n(X_1), \ldots, W_n(X_n)$ specified in equation (5) below. Using similar arguments as in the proof of Theorem 1, we can show that this statistic and the test statistic $\tilde{T}_n$ are asymptotically equivalent,

$$\sup_{t \in \mathbb{R}} \left| \frac{1}{n} \sum_{j=1}^{n} \tilde{W}_j \mathbf{1}\left[\hat{\varepsilon}_j \le t\right] - \frac{1}{n} \sum_{j=1}^{n} W_n(X_j) \Big\{ \mathbf{1}[\varepsilon_j \le t] - F(t) \Big\} \right| = o_P(n^{-1/2}). \tag{3}$$

The term in absolute brackets in (2) converges in distribution to a standard Brownian bridge B_0, i.e. we have

$$\sup_{t \in \mathbb{R}} \left| n^{-1/2} \sum_{j=1}^{n} W_n(X_j) \Big\{ \mathbf{1}\left[\varepsilon_j \le t\right] - F(t) \Big\} \right| \xrightarrow{D} \sup_{t \in [0,1]} B_0(t) \quad (n \to \infty). \tag{4}$$

Combining equations (3) and (4) yields the desired limiting result for $\tilde{T}_n$.

Corresponding author: Justin Chown (justin.chown@ruhr-uni-bochum.de)
[1] *Ruhr-Universität Bochum, Fakultät für Mathematik, Lehrstuhl für Stochastik, 44780 Bochum, DE*
[2] *Department of Statistics, Texas A&M University, College Station, TX 77843-3143, USA.*

Before sketching the proofs of equations (3) and (4), let us explain the notation and how W_n emerges. We use local polynomial smoothers and properties of local polynomial smoothers as presented in Müller et al. (2009). For greater transparency, we therefore use the same notation as in that article. In particular we write $\hat{r}$ for the estimated regression function and $\hat{a}$ for the approximation of the difference $\hat{r} - r$. Our estimators $\hat{r}$ and $\hat{r}_2$ are explained in the article: $\hat{r}_2$ has the same form as $\hat{r}$ with Y^2 in place of Y; $\hat{a}$ and $\hat{a}_2$ approximate the differences associated with $\hat{r}$ and $\hat{r}_2$.

Our weights W_n in the equivalent statistic (2) are

$$W_n(\cdot) = \frac{(nc_n^m)^{1/2} W_{n,1}(\cdot)}{\Sigma_{\hat{\sigma}}^{1/2}} \tag{5}$$

with

$$W_{n,1}(\cdot) = V_{n,1}(\cdot) - \int_{[0,\,1]^m} V_{n,1}(x) g(x)\,dx, \quad V_{n,1}(x) = \frac{1}{2\sigma_0}\{\hat{a}_2(x) - 2r(x)\hat{a}(x)\}$$

and, writing $\mu_4 = E[\varepsilon^4]$ and $\Psi = (\psi_i)_{i \in I(d)}$,

$$\Sigma_{\hat{\sigma}} = \frac{\mu_4 - \sigma_0^4}{4\sigma_0^2} \int_{[0,\,1]^m} \int_{[-1,\,1]^m} \left\{e^T Q_*^{-1}(u)\Psi(v)\right\}^2 w^2(v) g^2(u)\,dv\,du;$$

see the article and Müller et al. (2009) for notation. The $W_{n,1}$ in the definition of W_n comes from the numerator of $\tilde{W}_j$ in equation (1), namely from approximating $\hat{\sigma}(x) - \sigma_0$ by

$$\frac{1}{2\sigma_0}\{\hat{a}_2(x) - r(x)\hat{a}(x)\};$$

the factor $(nc_n^m/\Sigma_{\hat{\sigma}})^{1/2}$ is an approximation of the denominator of $\tilde{W}_j$,

$$\left|\frac{1}{n}\sum_{j=1}^n \left\{\hat{\sigma}(X_j) - \bar{\hat{\sigma}}\right\}^2 - \frac{\Sigma_{\hat{\sigma}}}{nc_n^m}\right| = o_P(1). \tag{6}$$

Proceeding as in Dette (2002), who uses Corollary 3.1 of Hall and Heyde (1980), and using equicontinuity arguments derived from Theorem 1 of Bae et al. (2014), we obtain that

$$(nc_n^{m/2})\frac{1}{n}\sum_{j=1}^n W_{n,1}(X_j)\Big\{\mathbf{1}[\varepsilon_j \le \cdot] - F(\cdot)\Big\} = c_n^{m/2}\sum_{j=1}^n W_{n,1}(X_j)\Big\{\mathbf{1}[\varepsilon_j \le \cdot] - F(\cdot)\Big\}$$

converges in distribution to $\Sigma_{\hat{\sigma}}^{1/2} B_0 \circ F$. A similar argument can also be found on pages 549-551 of Koul and Ossiander (1994) in the proof of Lemma 2.1.

This yields for the scaled statistic

$$\begin{aligned}
&\sup_{t\in\mathbb{R}}\left|c_n^{m/2}\sum_{j=1}^n \frac{W_{n,1}(X_j)}{\Sigma_{\hat{\sigma}}^{1/2}}\Big\{\mathbf{1}\big[\varepsilon_j \le t\big] - F(t)\Big\}\right| \\
&= \sup_{t\in\mathbb{R}}\left|n^{-1/2}\sum_{j=1}^n \frac{(nc_n^m)^{1/2}W_{n,1}(X_j)}{\Sigma_{\hat{\sigma}}^{1/2}}\Big\{\mathbf{1}\big[\varepsilon_j \le t\big] - F(t)\Big\}\right| \\
&\xrightarrow{D} \sup_{t\in[0,1]} B_0(t),
\end{aligned}$$

i.e., in view of (5), the desired limiting result (4).

To see that (3) holds true write $\overline{W}_n = n^{-1}\sum_{j=1}^n W_n(X_j)$ and

$$V^* = \big(\Sigma_{\hat\sigma}/(nc_n^m)\big)^{1/2}\Big/\Big[\frac{1}{n}\sum_{k=1}^n\Big\{\hat\sigma(X_k)-\overline{\hat\sigma}\Big\}^2\Big]^{1/2}.$$

Define the remainder terms

$$R_1^* = \frac{1}{n}\sum_{j=1}^n W_n(X_j)\Big\{\mathbf{1}\big[\varepsilon_j \le t+\hat r(X_j)-r(X_j)\big] - F\big(t+\hat r(X_j)-r(X_j)\big) - \mathbf{1}[\varepsilon_j\le t] + F(t)\Big\},$$

$$R_2^* = V^*\frac{1}{n}\sum_{j=1}^n W_n(X_j)\Big\{F\big(t+\hat r(X_j)-r(X_j)\big) - F(t)\Big\},$$

$$R_3^* = \big(V^*-1\big)\Big(\frac{1}{n}\sum_{j=1}^n W_n(X_j)\Big\{\mathbf{1}\big[\varepsilon_j \le t+\hat r(X_j)-r(X_j)\big] - F\big(t+\hat r(X_j)-r(X_j)\big)\Big\}\Big),$$

$$R_4^* = V^*\overline{W}_n\Big(\frac{1}{n}\sum_{j=1}^n\Big\{1\big[\varepsilon_j \le t+\hat r(X_j)-r(X_j)\big] - F\big(t+\hat r(X_j)-r(X_j)\big)\Big\}\Big),$$

$$R_5^* = V^*\overline{W}_n\Big(\frac{1}{n}\sum_{j=1}^n\Big\{F\big(t+\hat r(X_j)-r(X_j)\big) - F(t)\Big\}\Big).$$

The difference of the two statistics in (3) is equal to $R_1^*+R_2^*+R_3^*-R_4^*-R_5^*+o_P(n^{-1/2})$. Metric entropy and equicontinuity arguments as in the proof of Lemma 2 yield $\sup_{t\in\mathbb{R}}|R_1^*| = o_P(n^{-1/2})$ and

$$\sup_{t\in\mathbb{R}}\Big|\frac{1}{n}\sum_{j=1}^n W_n(X_j)\Big\{F\big(t+\hat r(X_j)-r(X_j)\big)-F(t)\Big\}\Big| = o_P(n^{-1/2}).$$

By equation (6), the quantity V^* consistently estimates one. This and the latter statement give $\sup_{t\in\mathbb{R}}|R_2^*| = o_P(n^{-1/2})$. Proceeding exactly as in Dette (2002), we obtain that the weights W_n satisfy $|\overline{W}_n| = O_P(n^{-1/2})$. Using this and the consistency of V^* it follows that $\sup_{t\in\mathbb{R}}|R_i^*| = o_P(n^{-1/2})$, $i=3,4,5$. □

References

[1] Bae, J., Jun, D. and Levental, S. (2014). The second central limit theorem for martingale difference arrays. *Bull. Korean Math. Soc.* **51**, 317-328.

[2] Dette, H. (2002). A consistent test for heteroscedasticity in nonparametric regression based on the kernel method. *J. Stat. Plann. Inference* **103**, 311-329.

[3] Koul, H.L. and Ossiander, M. (1994). Weak convergence of randomly weighted dependent residual empiricals with applications to autoregression. *Ann. Statist.* **22**, 540-562.

[4] Hall, P. and Heyde, C.C. (1980). *Martingale limit theory and its application.* Academic Press, Inc., New York.

[5] Müller, U.U., Schick, A. and Wefelmeyer, W. (2009). Estimating the error distribution function in nonparametric regression with multivariate covariates. *Statist. Probab. Lett.*, **79**, 957-964.