Corrections and Additions to: Eggermont and LaRiccia, Maximum Penalized Likelihood Estimation. Volume I

(The “real” unattributed ones are courtesy of David Mason.)

Page 0: Throughout the text: double exponential $\rightarrow$ two-sided exponential

Page xvii, definition of $W^{m,p}(0,1)$: $L^p(\Omega)$ has a double right parenthesis. (These types of errors should really be discovered by computer programs, such as texmatch or ones that also check the sizes of parentheses, etc. We promise to do better on Volume II.)

Page 6, last line: § 5 should be § 6.

Page 15, Equation (3.11): The power of $\sigma$ should be $(1/p) - 1$, so

$$\| \varphi_\sigma - \psi_\sigma \|_p = \sigma^{(1/p)-1} \| \varphi - \psi \|_p.$$ 

Courtesy of Bin Wang, University of South Alabama (27/2/05).

Page 16, line 7 from below: Chapter 10 should be Chapter 9. Specifically, Exercise (9.1.22). (An extra hint: Cauchy-Schwarz.)

Page 148, Theorem (4.21) and proof: The inequality in line 12,

$$\lim_{n \to \infty} \sup \left( 2 s_n \log n \right)^{-1/2} \left| S_n - E[S_n] \right| \leq_{as} 1$$

$$\lim_{n \to \infty} \sup \left( 2 s_n \log n \right)^{-1/2} \left| S_n - E[S_n] \right| \leq_{as} 1$$
should be replaced by
\[ \limsup_{n \to \infty} \left( \frac{1}{2} s_n^2 \log n \right)^{-1/2} \left| S_n - \mathbb{E}[S_n] \right| \leqslant 1. \]

An analogous correction is to be made in line 18.

Page 149: The right hand side of the last inequality of Theorem (4.22) should be replaced by \( \sqrt{2} \).

Page 154–155: “where \( \varepsilon = (nh)^{-1} \| A \|_\infty \)” occurs twice.

Page 159 (§4.6): Theorem (6.8) on the nonparametric estimation of the entropy can be strengthened in that the smoothing parameter need not be constant on blocks, as follows.

(6.8') THEOREM. Suppose \( f \) satisfies (6.2). Let \( \alpha > 1 \). If \( f \) has a moment of order \( > \frac{1}{\alpha - 1} \), then for \( h \approx n^{-1/(2\alpha+1)} \) (deterministic)
\[ \delta_{nh} = \mathcal{O}\left( n^{-2/(2\alpha+1)} \right). \]

For the proof we need the monotonicity in \( h \) of the neg-entropy of the kernel estimator
\[ \mathcal{E}(f^{nh}) = \int_{\mathbb{R}} f^{nh}(x) \log f^{nh}(x) \, dx, \]
in which \( f^{nh}(x) = [2h \ast dF_n](x) \).

(6.8.1) LEMMA. Let \( h_o > 0 \). If \( \mathcal{E}(f^{nh}) < \infty \) for \( h = h_o \), then \( \mathcal{E}(f^{nh}) < \infty \) for all \( h > h_o \) and is a decreasing function of \( h \).

The proof is similar to that of Lemma (5.19).

PROOF OF THEOREM (6.8'). Recall that
\[ \delta_{nh} = \mathcal{E}(f^{nh}) - \frac{1}{n} \sum_{i=1}^{n} f(X_i), \]
and that the old Theorem (6.8) says that
\[ \delta_{nh} = \mathcal{O}\left( n^{-2/(2\alpha+1)} \right), \]
provided \( h \approx n^{-1/(2\alpha+1)} \) is deterministic and \( h \) is constant on blocks,
\[ h = h_k, \quad 2^{k-1} < n \leqslant 2^k. \]

Now, let \( h \approx n^{-1/(2\alpha+1)} \) (deterministic but not necessarily constant on blocks). By Lemma (6.8.1), \( \delta_{nh} \) is decreasing in \( h \), and so, with
\[ \lambda \equiv \lambda_k \overset{\text{def}}{=} \max \left\{ h_n : 2^{k-1} < n \leqslant 2^k \right\}, \]
we obtain
\[ \delta_{nh} \leq \delta_{n,\lambda}, \quad 2^{k-1} < n \leq 2^k. \]

but this is \( O\left(n^{-2/(2\alpha+1)}\right) \), almost surely.

Similarly, with
\[ \nu \equiv \nu_k \overset{\text{def}}{=} \min \left\{ h_n : 2^{k-1} < n \leq 2^k \right\}, \]

one gets
\[ \delta_{nh} \geq \delta_{n,\nu}, \quad 2^{k-1} < n \leq 2^k. \]

and this is also \( O\left(n^{-2/(2\alpha+1)}\right) \), almost surely. Q.e.d.

Finally, we note that this fills a gap in the proof of the Main Theorem 1 in our IEEE paper [IEEE Trans. Information Theory 45, 1321–1326 (1999)], since the quantity
\[ F_{n,h} = \int_{\mathbb{R}} \left( s_h \ast \log s_h \ast g - \log g \right) dG_n \]
in the notation of that paper does not appear to be monotone in \( h \). (Note that \( s_h \equiv \mathfrak{B}_h \)

Page 173 (§4.6. Asymptotic normality of the \( L^1 \) error) : The material in this section is very subtle, e.g., the application of the Sweeting (1977) Lemma in this form is suspect. It has all been cleaned up in a paper by Giné, Mason and Zaitsev (2003)

Page 207 : the section header “4. Roughness penalization · · ·” should not be the last line on the page.

Page 211, last displayed formula in the proof of Lemma (4.16): replace “\( \wedge \)” by “\( \vee \)” . (6/6/01)

Page 217, formula (5.20) : replace “9” by “(” . [ (x) vs 9x ] (5/12/01)

Page 231, Exercise (2.27) : The observant reader will notice that the minimization problem stated here is not a least-squares or even a quadratic problem. It should read
\[ \text{minimize } \int_0^\infty |f(x)|^2 \, dx - 2 \int_0^\infty f(x) \, dF_n(x) \]
subject to \( f \) is monotone .

(4/27/04)

Page 243, line 12 : Replace “\( \text{im}(\Phi) \)” by “\( \text{im}(\varphi) \)”.

Page 259, Lemma (6.34) : the bound should be \( (nh)^{-1/2} (\log(1/h))^{1/2} \). The power 1/2 on the logarithm is missing.

Page 333, Caption of Figure 4.4 : “acronyms” runs into the margin. We had fixed that in Figure 4.3 but not here. Aggh!!