Prof. Dr. W. Wefelmeyer
Dipl.-Math. Markus Schulz **Sommersemester 2010**

Übungen zur Asymptotischen Statistik
Serie 5
Abgabe: Dienstag, 18. Mai 2010, vor der Vorlesung

Es gelte $Y = X + Z$, und X und Z seien unabhängig. X besitze die Dichte $f \in L_2(\lambda)$, Z die Dichte g und Y die Dichte h, und g sei bekannt. Um f zu schätzen, setzt man an

$$T := \frac{1}{2\pi} \int e^{-itx} \frac{\varphi_{\hat{h}}(t)}{\varphi_g(t)}\, dt$$

mit dem Kernschätzer $\hat{h}(x) = \frac{1}{n}\sum_{j=1}^n K_b(x - Y_j)$, $\varphi_g(t) = \int e^{itx} g(x)dx$ und einem Kern $K \in L_2(\lambda)$, dessen Fourier-Transformierte φ_K einen kompakten Träger besitzt. Zudem gelte $\varphi_g(t) \neq 0$ für alle t.

Hinweis: Für geeignete Funktionen gilt $\frac{1}{2\pi}\int e^{-itx}\varphi_f(t)dt = f(x)$. Die Parseval-Identität $\frac{1}{2\pi}\int |\varphi_f(t)|^2 dt = \int f^2(x)dx$ könnte ebenfalls nützlich sein.

21. (3 Punkte)
a) Begründen Sie (heuristisch), warum der angegebene Schätzer geeignet ist, um die Dichte f von X zu schätzen.
b) Zeigen Sie, dass der Schätzer T die Darstellung $\frac{1}{n}\sum_{j=1}^n K_b^Z(x - Y_j, b)$ mit $K_b^Z(u,b) = b^{-1}K^Z(\frac{u}{b}, b)$ und $K^Z(u,b) = (2\pi)^{-1}\int e^{-itu}\varphi_K(t)\big(\varphi_g\big(\frac{t}{b}\big)\big)^{-1}dt$ besitzt.

22. (5 Punkte) Die Situation sei die gleiche wie in Aufgabe 21.
a) Zeigen Sie, dass der mittlere integrierte quadratische Fehler (MISE) gegeben ist durch

$$\frac{1}{2\pi n}\int |\varphi_K(tb)|^2\Big(\frac{1}{|\varphi_g(t)|^2} - |\varphi_f(t)|^2\Big)dt + \frac{1}{2\pi}\int |\varphi_K(tb) - 1|^2 |\varphi_f(t)|^2 dt.$$

b) Weisen Sie nach, dass sich dies auch in der Form

$$\frac{1}{nb}R(K^Z(.,b)) + \int f^2(x)dx + (1-n^{-1})\int (K_b * f)^2(x)dx - 2\int (K_b * f)(x)f(x)dx$$

mit $R(K^Z(.,b)) = \int (K^Z(u,b))^2 du$ schreiben lässt.

23. Sei (X, Y) ein zweidimensionaler Zufallsvektor mit gemeinsamer Dichte f. Sei f_1 die Dichte von X und $r(x) = E(Y|X = x)$ die Regressionsfunktion von Y auf X. Definiere den Nadaraya-Watson-Schätzer

$$\hat{r}(x) = \frac{\sum_{i=1}^{n} Y_i K_b(x - X_i)}{\sum_{i=1}^{n} K_b(x - X_i)}.$$

Berechnen Sie unter geeigneten Bedingungen die Konvergenzrate von $\hat{r}(x)$.

Voraussetzungen für die Aufgaben 24 und 25: Sei (X, Y) ein Zufallsvektor mit zweimal stetig differenzierbarer Regressionsfunktion g. Die zweite Ableitung von g sei beschränkt. Ferner besitze X die Dichte f mit Träger $[0, 1]$. f sei stetig auf $(0, 1)$ mit $f(0) = f(0+)$. Die Funktion

$$v(x) := \mathrm{Var}(Y|X = x) := E((Y - g(X))^2|X = x)$$

sei ebenfalls stetig. Um die Regressionsfunktion zu schätzen, wählen wir einen symmetrischen beschränkten Kern $K \geq 0$ mit Träger $[-1, 1]$ und $\int u^6 K(u)du < \infty$.

24. a) Zeigen Sie, dass für $j = 0, 1, 2, 3$ gilt

$$\frac{1}{n}\sum_{i=1}^{n} X_i^j K_b(X_i) = b^j f(0+) \int_0^1 u^j K(u)du + o(b^j).$$

b) Sei $\hat{g}_0(x)$ der aus der Vorlesung bekannte Nadaraya-Watson-Schätzer. Berechnen Sie für diesen den bedingten Bias $E(\hat{g}_0(0)|X_1, \ldots, X_n) - g(0)$ und die bedingte Varianz

$$\mathrm{Var}(\hat{g}_0(0)|X_1, \ldots, X_n) = E((\hat{g}_0(0) - E(\hat{g}_0(0)|X_1, \ldots, X_n))^2|X_1, \ldots, X_n).$$

25. a) Bestimmen Sie explizit den lokalen polynomialen Glätter $\hat{g}_1(x)$ der Ordnung $r = 1$ (sog. *lokaler linearer Glätter*).
b) Berechnen Sie den bedingten Bias $E(\hat{g}_1(0)|X_1, \ldots, X_n) - g(0)$ und die bedingte Varianz

$$\mathrm{Var}(\hat{g}_1(0)|X_1, \ldots, X_n) = E((\hat{g}_1(0) - E(\hat{g}_1(0)|X_1, \ldots, X_n))^2|X_1, \ldots, X_n).$$

c) Vergleichen Sie die in Aufgabe 24b) gefundenen Raten mit den Raten aus Aufgabenteil b).