Mathematische Statistik 2, Lösungen, Blatt 9

33.

a) Für den Histogramm-Schätzer $\hat{f}_n(x) = \frac{1}{nh}\sum_{i=1}^n 1_{I(x)}(X_i)$ der Dichte $f(x)$, mit $1_{I(x)}(X_i) = 1$, falls x und X_i in dem selben Intervall I_j ($= 0$ sonst), rechnet man [I_j, $j \in \mathbb{Z}$, disjunkte Zerlegung von $\mathbb{R}$; $\hat{f}_n(x)$ konstant auf jedem I_j]

$$\int_{-\infty}^{\infty} \hat{f}_n(x)\,dx = \sum_{j\in\mathbb{Z}} \int_{I_j} \hat{f}_n(x)\,dx = \sum_{j\in\mathbb{Z}} |I_j| \cdot \hat{f}_n(x)|_{x\in I_j}$$
$$= \frac{1}{n\,h}\,h \sum_{i=1}^n \sum_{j\in\mathbb{Z}} 1_{I_j}(X_i) = \frac{n\,h}{n\,h} = 1\,.$$

Alternativ: $$\int_{-\infty}^{\infty} \hat{f}_n(x)\,dx = \frac{1}{n\,h}\sum_{i=1}^n \int_{-\infty}^{\infty} 1_{I(x)}(X_i)\,dx$$
$$= \frac{1}{n\,h}\sum_{i=1}^n \int_{-\infty}^{\infty} 1_{I(X_i)}(x)\,dx = \frac{1}{n\,h}\sum_{i=1}^n \int_{I(X_i)} 1\,dx = \frac{n\,h}{n\,h} = 1\,.$$

b) Sei $x \in \mathbb{R}$. Die $1_{I(x)}(X_i)$, $i = 1, 2, \ldots$, sind unabhängig und identisch verteilt, jedes X_i mit Dichte f.

$$\mathbb{E}\big(\hat{f}_n(x)\big) = \frac{1}{n\,h}\sum_{i=1}^n \mathbb{E}\big(1_{I(x)}(X_i)\big) = \frac{1}{n\,h}\,n \int_{I(x)} f(t)\,dt = \frac{1}{h}\int_{I(x)} f(t)\,dt \equiv \frac{1}{h}\,p_h(x)\,.$$

$$\hat{f}_n(x) = \frac{1}{h}\frac{1}{n}\sum_{i=1}^n 1_{I(x)}(X_i) \longrightarrow_{GdgZ} \frac{1}{h}\mathbb{E}\big(1_{I(x)}(X_1)\big) = \frac{1}{h}\int_{I(x)} f(t)\,dt \quad [\mathbb{P}\,f.s.]\,.$$

c) Da die $1_{I(x)}(X_i)$, $i = 1, 2, \ldots$, unabhängig und $B(1, p_h(x))$-verteilt sind, gilt

$$\mathrm{Var}\,\big(\hat{f}_n(x)\big) = \frac{1}{(n\,h)^2}\sum_{i=1}^n \mathrm{Var}\,\big(1_{I(x)}(X_i)\big) = \frac{1}{n\,h^2}\cdot p_h(x)\cdot(1 - p_h(x))\,.$$

34.

a) Wegen des MWS der Integralrechnung gibt es ein $x_h \in I(x)$ mit

$$\mathbb{E}\big(\hat{f}_n(x)\big) =_{33b} \frac{1}{h}\int_{I(x)} f(t)\,dt =_{MWS} \frac{1}{h}|I(x)|\cdot f(x_h) = f(x_h) \longrightarrow f(x)$$

bei $n \to \infty$ (also bei $h \to 0$ bzw. $x_h \to x$). Dabei haben wir die Stetigkeit von f ausgenützt.

b) Sei $x_h \in I(x)$. Wegen $x_h \to x$ (bei $n \to \infty$) und der Stetigkeit von f folgt

(i) $f(x_h) \;=\; f(x) + \eta_h$, mit $\;\eta_h \to 0\;$ (bei $h \to 0$)
(ii) $p_h(x) \;=_{MWS}\; h\,f(x_h) \longrightarrow 0\;$ (bei $h \to 0$). Also

$$\begin{aligned}\mathrm{Var}\left(\hat{f}_n(x)\right) &=_{33c} \frac{1}{n\,h^2}\cdot p_h(x)\cdot(1-p_h(x)) \;=\; \left(\frac{1}{n\,h}\,f(x) + \frac{1}{n\,h}\,\eta_h\right)\cdot(1-p_h(x)) \\ &= \frac{1}{n\,h}\,f(x) + \frac{1}{n\,h}\left[-\;f(x)\,p_h(x) + \eta_h\,(1-p_h(x))\right] \;=\; \frac{1}{n\,h}\,f(x) + o\Big(\frac{1}{nh}\Big).\end{aligned}$$

c) Wenn x der Mittelpunkt des Intervalls $I(x)$ der Länge h ist, dann gilt $\int_{I(x)}(t-x)\,dt \;=\; 0$ und $\int_{I(x)}(t-x)^2\,dt \;=\; \frac{1}{12}\,h^3$. Also mittels Taylorentwicklung von $f(t)$ an der Stelle x

$$\begin{aligned}\mathbb{E}\big(\hat{f}_n(x)\big) - f(x) &\underset{33b}{=} \frac{1}{h}\int_{I(x)}[f(t)-f(x)]dt \underset{T.E.}{=} \frac{1}{h}\int_{I(x)}\left[f'(x)\,(t-x) + \frac{1}{2}f''(x^*)(t-x)^2\right]dt \\ &=_{MWS\ I} \frac{1}{h}\Big[0 + \frac{1}{2}\,f''(x^{**})\,\frac{1}{12}\,h^3\Big] \;=\; \frac{1}{24}\left[h^2 f''(x) \;+\; h^2\left(f''(x^{**}) - f''(x)\right)\right] \\ &= \frac{1}{24}\,h^2 f''(x) + o(h^2) \qquad [x^* = x^*(t) \text{ und } x^{**} \in I(x),\ f''(x^{**}) \to f''(x) \text{ bei } h \to 0].\end{aligned}$$

35.

a) Für Y und $f(X)$ aus L_2 ist $I_f < \infty$. Ferner ist $m \in L_2^X$, d. h. $m(X) \in L_2$. In der Tat, aufgrund der Jensenschen Ungleichung für bedingte Erwartungen gilt mit der konvexen Funktion $\phi(x) = x^2$

$$\mathbb{E}\big(\phi(m(X))\big) \;=_{Def.}\; \mathbb{E}\big(\phi(\mathbb{E}(Y|X))\big) \;\leq\; \mathbb{E}\big(\mathbb{E}(\phi(Y)|X)\big) \;=\; \mathbb{E}\big(\phi(Y)\big) \;<\; \infty\,,$$

Letzteres, weil $Y \in L_2$.

b) $\mathbb{E}(m(X)) \;=_{Def.}\; \mathbb{E}\big(\mathbb{E}(Y|X)\big) \;=\; \mathbb{E}(Y)$. Unter Verwendung der Ungleichung in a) und der Verschiebungsformel erhalten wir für die Varianzen.

$$\begin{aligned}\mathrm{Var}(m(X)) \;&=\; \mathbb{E}\big(m(X) - \mathbb{E}(m(X))\big)^2 \;=_{\phi=x^2}\; \mathbb{E}\big(\phi(m(X))\big) - \big(\mathbb{E}(m(X))\big)^2 \\ &\leq\; \mathbb{E}(\phi(Y)) - \big(\mathbb{E}(Y)\big)^2 \;=\; \mathbb{E}(Y^2) - \big(\mathbb{E}(Y)\big)^2 \;=\; \mathrm{Var}(Y)\,.\end{aligned}$$

Genauer rechnet man: $\mathrm{Var}\left(m(X)\right) = \mathrm{Var}(Y) - \mathbb{E}\big(\,\mathrm{Var}(Y|X)\big)$.

c) $I_f \;=\; \mathbb{E}\big(Y - m(X) + m(X) - f(X)\big)^2 \;=\; \mathbb{E}\big(Y - m(X)\big)^2 + \mathbb{E}\big(m(X) - f(X)\big)^2 \equiv I_m + B$. Dabei verschwindet der gemischte Term wegen

$$\begin{aligned}\mathbb{E}\big[(Y - m(X))\cdot(m(X) - f(X))\big] \;&=\; \mathbb{E}\big\{\mathbb{E}\big[(...)\cdot(...)|X\big]\big\} \\ &=\; (m(X) - f(X))\cdot\mathbb{E}\big[Y - m(X)|X\big] \;=\; 0\,.\end{aligned}$$

Da $B \geq 0$, folgt $\;I_f \geq I_m\;$ für alle $f \in L_2^X$, mit Gleicheitszeichen genau dann, wenn $B = 0$, also genau dann, wenn $m(X) = f(X)\;$ $\mathbb{P}$-f.s (d. h. $m = f\;$ $\mathbb{P}_X$-f.s).

36.

a) Ausgehend von $\;Y_{t+1} = \nu + \alpha\cdot Y_t + e_{t+1},\;$ $t \in \mathbb{N}_0,\;$ berechnet man für die Regressionsfunktion m_1 im Fall $(X,Y) = (Y_t, Y_{t+1})$

$$\hat{Y}_{t,1} \equiv \mathbb{E}(Y_{t+1}|Y_t) \;=\; \mathbb{E}\big(\nu + \alpha\cdot Y_t + e_{t+1}|Y_t\big) \;=\; \nu + \alpha\cdot Y_t$$

(also: $m_1(y) = \nu + \alpha\, y$), denn $\mathbb{E}(Y_t|Y_t) = Y_t$, $\mathbb{E}(e_{t+1}|Y_t) = \mathbb{E}(e_{t+1}) = 0$, wegen der Unabhängigkeit von e_{t+1} und Y_t.
Für die Regressionsfunktion m_2 im Fall $(X,Y) = (Y_t, Y_{t+2})$ erhält man, mit der Prognose $\hat{Y}_{t,1}$ von eben

$$\hat{Y}_{t,2} \equiv \mathbb{E}(Y_{t+2}|Y_t) \;=\; \mathbb{E}\big(\nu + \alpha \cdot Y_{t+1} + e_{t+2}|Y_t\big) \;=\; \nu + \alpha \cdot \hat{Y}_{t,1} \;=\; \nu + \alpha\,\nu + \alpha^2\, Y_t$$

(also: $m_2(y) = \nu + \alpha\,\nu + \alpha^2\, y$), denn wiederum ist $\mathbb{E}(e_{t+2}|Y_t) = \mathbb{E}(e_{t+2}) = 0$.
Rekursiv folgt $\hat{Y}_{t,l} \equiv \mathbb{E}(Y_{t+l}|Y_t) \;=\; \nu + \alpha \cdot \hat{Y}_{t,l-1}$, also

$$m_l(y) \;=\; \nu\,\big(1 + \alpha + \ldots + \alpha^{l-1}\big) + \alpha^l\, y\,.$$

b) Wegen $\alpha^l\, Y_t \longrightarrow 0$ (f. s., bei $l \to \infty$) haben wir den f.s. Limes

$$\hat{Y}_{t,l} \;\longrightarrow\; \nu\,(1 + \alpha + \alpha^2 + \ldots) \;=\; \nu\,\frac{1}{1-\alpha} \;=\; \mu \;\equiv\; \mathbb{E}(Y_t)\,,\ \text{für alle } t \in \mathbb{N}.$$

Weit ausgreifende Prognosen reproduzieren den Mittelwert der (bisher bekannten) Beoachtungen.