ulm university universität uulm

Dr. Kirsten Schorning
Dipl.-Math. Stefan Roth

WS 2016/17
7. November 2016

Angewandte Stochastik 2 - Übungsblatt 4

Besprechung: 14. November in der Übung.

Aufgabe 1 (2 + 2 + 3 Punkte)

Sei $\widehat{f}_h$ ein Kerndichteschätzer der Dichtefunktion f mit einer fest gewählten Bandbreite h und Kernfunktion $K : \mathbb{R} \to \mathbb{R}$. Dann ist $\widehat{f}_h(x)$ eine Zufallsvariable für alle $x \in \mathbb{R}$ und man definiert

- den mittleren quadratischen Fehler von $\widehat{f}_h$ an der Stelle $x_0 \in \mathbb{R}$ durch

$$\mathrm{MSE}\left(\widehat{f}_h(x_0)\right) = \mathbb{E}\left(\widehat{f}_h(x_0) - f(x_0)\right)^2,$$

- den integrierten mittleren quadratischen Fehler von $\widehat{f}_h(x)$ durch

$$\mathrm{MISE}\left(\widehat{f}_h\right) = \mathbb{E}\int_{\mathbb{R}}\left(\widehat{f}_h(x) - f(x)\right)^2 \mathrm{d}x.$$

a) Zeige: $\mathrm{MISE}\left(\widehat{f}_h\right) = \int_{\mathbb{R}} \mathrm{MSE}\left(\widehat{f}_h(x)\right) \mathrm{d}x$.

b) Zeige:[1] $\mathrm{BIAS}\left(\widehat{f}_h(x)\right) = \int_{\mathbb{R}} K(t)(f(x+ht) - f(x))\mathrm{d}t$.

c) Folgere: $\mathrm{MISE}\left(\widehat{f}_h\right) = \int_{\mathbb{R}} \mathsf{Var}\widehat{f}_h(x)\mathrm{d}x + \int_{\mathbb{R}}\left(\int_{\mathbb{R}} K(t)(f(x+ht) - f(x))\mathrm{d}t\right)^2 \mathrm{d}x$.

Aufgabe 2 (5 Punkte)

Betrachte den Standardschätzer

$$\widehat{l} = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}_{\{X_i > \gamma\}}$$

für $l = \mathbb{P}(X > \gamma)$ für eine Zufallsvariable X und einen Schwellwert γ. Gib mithilfe der Tschebscheff-Ungleichung an, wie groß n mindestens gewählt werden muss, damit zu 99% der Schätzer $\widehat{l}$ nur um 0.1 von l abweicht?

Aufgabe 3 (6 Punkte)

Es sei $(X_1, \ldots, X_n)$ eine i.i.d. Zufallsstichprobe zur Verteilung F. Die i-te Ordnungsstatistik $X_{(i)}$ von $(X_1, \ldots, X_n)$ ist definiert durch

$$X_{(i)} = \min\{X_j :\ \#\{k :\ X_k \le X_j\} \ge i\}, \quad i = 1, \ldots, n,$$

d.h. die Zufallsvariablen $X_{(1)} \le \cdots \le X_{(n)}$ sind die (punktweise) aufsteigend sortierten Werte der Stichprobe $(X_1, \ldots, X_n)$. Zeige: Die Verteilungsfunktion $F_{X_{(i)}}$ von $X_{(i)}$ ist gegeben durch

$$F_{X_{(i)}}(t) = \sum_{k=i}^{n} \binom{n}{k} (F(t))^k (1 - F(t))^{n-k}, \quad i = 1, \ldots, n.$$

[1] Der *Bias* (oder auch die *Verzerrung*) eines Schätzers $\hat{\theta}$ für θ ist definiert durch $\mathrm{BIAS}(\hat{\theta}) = \mathbb{E}(\hat{\theta} - \theta)$