4. Übung am 04. April 2022

UV Angewandte Statistik (405.170)
Link Ankreuzliste: siehe www.trutschnig.net/courses
Mit 'F' versehene Aufgaben sind freiwillig, mit * versehene Aufgaben haben einen erhöhten Schwierigkeitsgrad.
Alle Verweise beziehen sich auf das Statistik-Skriptum.

Übungsaufgabe 19. Beweisen Sie Satz 3.1.

Übungsaufgabe 20. Beweisen Sie, dass die Funktion f in Definition 3.5 tatsächlich eine Wahrscheinlichkeitsdichte auf $\mathbb{R}^d$ ist.

Übungsaufgabe 21. Konstruieren Sie (i) eine diskrete und (ii) eine absolut stetige Zufallsvariable X, die $\mathbb{P}^X = \mathbb{P}^{T(X)}$ mit $T(x) = \frac{1}{x}$ für jedes $x \in \mathbb{R} \setminus \{0\}$ erfüllt.

Übungsaufgabe 22. Angenommen X ist eine einfache Zufallsvariable. Aus der VO Mathematische Statistik wissen wir, dass der Erwartungswert $\mathbb{E}(X)$ die Abbildung $a \mapsto \|X - a\|_2^2 = \mathbb{E}(X - a)^2$ minimiert. Welche Größe minimiert die Abbildung $a \mapsto \|X - a\|_1 = \mathbb{E}|X - a|$?
Hinweis: Betrachten Sie zuerst eine einfache Zufallsvariable, die nur 2 Werte annehmen kann, und beantworten Sie die Fragen für diesen Fall.

Übungsaufgabe 23. Sei $X_1, X_2, \ldots$ eine Folge unabhängiger, $\mathcal{N}(\mu, \sigma^2)$-verteilter Zufallsvariable; $S_n^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2$ bezeichne die Stichprobenvarianz. Für jedes $n \geq 2$ sei die Zufallsvariable T_n definiert durch

$$T_n = \sqrt{n}\,\frac{\overline{X}_n - \mu}{\sqrt{S_n^2}}.$$

Beweisen Sie, dass T_n (für $n \to \infty$) schwach gegen $Z \sim \mathcal{N}(0,1)$ konvergiert und illustrieren Sie das Resultat mittels Simulationen in R.
Hinweis: Lesen Sie den Satz von Slutzky auf Wikipedia (English Version!) und wenden Sie ihn dann auf obige Situation an.

Übungsaufgabe 24. Gehen Sie den mehr oder weniger detailliert ausgeführten Beweis für die Implikationen 'aus 1. folgt 2.' des nachfolgenden Satzes Schritt für Schritt durch und ergänzen Sie etwaige Zwischenschritte. Inwiefern ist Satz 6.1 relevant für die Schätzung von Quantilen einer Verteilungsfunktion?

Satz 6.1 ([3]). *$G, G_1, G_2, \ldots$ seien Verteilungfunktionen. Dann sind die folgenden zwei Bedingungen äquivalent:*

1. $\lim_{n\to\infty} G_n(t) = G(t)$ *für jede Stetigkeitsstelle t von G (also $G_n \xrightarrow{weak} G$).*

2. $\lim_{n\to\infty} G_n^-(q) = G^-(q)$ *für jede Stetigkeitsstelle q von G^- (also $G_n^- \xrightarrow{weak} G^-$).*

Beweis: Wir zeigen, dass die erste Bedingung die zweite impliziert: Sei dafür $Z \sim \mathcal{N}(0,1)$ (oder irgendeine andere absolut stetige ZV, deren Dichte g auf ganz $\mathbb{R}$ ungleich 0 ist). Wir setzen $Y_n = G_n \circ Z$ und $Y = G \circ Z$; F_{Y_n} und F_Y bezeichne die entsprechenden Verteilungsfunktionen. Wegen $G_n \xrightarrow{weak} G$ und der Tatsache, dass G höchstens abzählbar viele Unstetigkeitsstellen hat[†], folgt sofort $Y_n \stackrel{[\mathbb{P}]}{\longrightarrow} Y$, also insbesondere $F_{Y_n} \stackrel{weak}{\longrightarrow} F_Y$. Für $q \in (0,1)$ erhalten wir unter Verwendung von Satz 4.12 (Statistik Skriptum)

$$\begin{aligned}\Phi(G_n^-(q)) &= \mathbb{P}\big(Z \leq G_n^-(q)\big) = \mathbb{P}\big(Z < G_n^-(q)\big) = \mathbb{P}\big(G_n \circ Z < q\big) = \mathbb{P}\big(Y_n < q\big) \\ &= \mathbb{P}\big(Y_n \leq q\big) - \mathbb{P}(Y_n = q) = F_{Y_n}(q) - \mathbb{P}(Y_n = q)\end{aligned}$$

Für jede Stetigkeitsstelle $q \in (0,1)$ von F_Y folgt $\lim_{n\to\infty} F_{Y_n}(q) = F_Y(q)$ sowie (sauber durchdenken!) $\lim_{n\to\infty} \mathbb{P}(Y_n = q) = 0$, und wir erhalten für ein solches q insgesamt

$$\begin{aligned}\lim_{n\to\infty} \Phi(G_n^-(q)) &= \lim_{n\to\infty} \big(F_{Y_n}(q) - \mathbb{P}(Y_n = q)\big) = F_Y(q) = \mathbb{P}(Y \leq q) \\ &= \mathbb{P}(G \circ Z < q) = \mathbb{P}(Z < G^-(q)) = \mathbb{P}(Z \leq G^-(q)) = \Phi(G^-(q)).\end{aligned}$$

Nachdem jede Stetigkeitsstelle von G^- auch Stetigkeitsstelle von F_Y ist (warum?) erhalten wir für jedes solche $q \in (0,1)$ also $\lim_{n\to\infty} \Phi(G_n^-(q)) = \Phi(G^-(q))$, aus der mittels Anwendung von Φ^- sofort $\lim_{n\to\infty} G_n^-(q) = G^-(q)$ folgt.

Für die Umkehrung skizzieren wir nur die Beweisidee: Sei $U \sim \mathcal{U}(0,1)$. Wegen $G_n^- \stackrel{weak}{\longrightarrow} G^-$ folgt für $Y_n = G_n^- \circ U$ und $Y = G^- \circ U$ sofort $Y_n \stackrel{[\mathbb{P}]}{\longrightarrow} Y$, woraus die gewünschte Eigenschaft $G_n \stackrel{weak}{\longrightarrow} G$ unschwer folgt.

[†] die Menge der Unstetigkeitsstellen also Lebesgue Mass 0 hat