Wahrscheinlichkeitstheorie und Statistik

Serie 6

MC 6-1. Sei X eine Zufallsvariable mit Werten in $\{0, 1, 2, \dots\}$. Welche der folgenden Aussagen sind korrekt? (Mehrere richtige Antworten sind möglich.)

(a) $\mathbb{E}(X) = \sum_{k=0}^{\infty} \mathbb{P}(X = k)$

(b) $\mathbb{E}(X) = \sum_{k=0}^{\infty} k \cdot \mathbb{P}(X = k)$

(c) $\mathbb{E}(X) = \sum_{k=1}^{\infty} \mathbb{P}(X \geq k)$

(d) $\mathbb{E}(X) = \sum_{k=0}^{\infty} \mathbb{P}(X \geq k)$

MC 6-2. Sei X Bernoulli-verteilt mit Parameter $p \in [0, 1]$ und $Y = (2X - 1)^2$. Was ist die Varianz von Y? (Nur eine Antwort ist richtig.)

(a) 1

(b) p

(c) p^2

(d) 0

Aufgabe 6-3. Sie $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und seien $X_1, \dots, X_n : \Omega \to \mathbb{R}$ Zufallsvariablen mit Verteilung $P : \mathcal{B}(\mathbb{R}) \to [0, 1]$. Die empirische Verteilung von $X_1, \dots, X_n$ ist

$$P_n : \Omega \times \mathcal{B}(\mathbb{R}) \to [0, 1], \qquad P_n(\omega, A) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_A \left(X_i(\omega) \right)).$$

Für jedes Ereignis $A \in \mathcal{B}(\mathbb{R})$ ist also $P_n(A) : \Omega \to \mathbb{R}$ eine Zufallsvariable. Zeigen Sie, dass $P_n(A)$ im folgenden Sinn eine gute Näherung für $P(A)$ ist:

(a) $\mathbb{E}\left(P_n(A)\right) = P(A)$.

(b) $\mathbb{V}\mathrm{ar}\left(P_n(A)\right) \leq 1/n$, wenn $X_1, \dots, X_n$ unabhängig sind.

Hier sind $\mathbb{E}$ bzw. $\mathbb{V}\mathrm{ar}$ der Erwartungswert bzw. die Varianz unter $\mathbb{P}$.

Aufgabe 6-4. Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und P_n die empirische Verteilung von Zufallsvariablen $X_1, \dots, X_n : \Omega \to \mathbb{R}$. Berechnen Sie den Erwartungswert und die Varianz des Wahrscheinlichkeitsmasses $P_n(\omega) : \mathcal{B}(\mathbb{R}) \to [0, 1]$, für eine beliebige aber fixe Realisierung $\omega \in \Omega$.

Hilfestellung: Es geht also um die Berechnung von $E_n(X)$ bzw. $\mathrm{Var}_n(X)$, wobei $X : \mathbb{R} \to \mathbb{R}$ die Identitätsabbildung ist und E_n bzw. Var_n die Erwartung bzw. Varianz von P_n bezeichnen, jeweils für fixes $\omega \in \Omega$.

Aufgabe 6-5. Schreiben Sie ein Python Programm, welches die lineare Regression Z einer Zufallsvariable Y auf eine Zufallsvariable X berechnet. Dabei soll die unbekannte Verteilung von (X, Y) näherungsweise durch die empirischen Verteilung von $(X_1, Y_1), \ldots, (X_n, Y_n)$ ersetzt werden. Wenn wir E_n, Var_n und Cov_n für die Erwartung, Varianz und Kovarianz unter der empirischen Verteilung P_n schreiben, geht es also um die Berechnung von

$$Z = \beta_0 + \beta_1 X, \qquad \beta_0 = E_n(Y) - \frac{\mathrm{Cov}_n(X,Y)}{\mathrm{Var}_n(X)} E_n(X), \qquad \beta_1 = \frac{\mathrm{Cov}_n(X,Y)}{\mathrm{Var}_n(X)}$$

für einen gegebenen Datensatz $(X_1, Y_1), \ldots, (X_n, Y_n)$ von Paaren reeller Zahlen.