Übungen zur Vorlesung „Stochastik für die Informatik“

Abgabe der Lösungen zu den S-Aufgaben: Freitag, 17. Januar 2020, vor der Vorlesung (12:10-12:15 im Magnus HS)

37.S. X sei uniform verteilt auf $[-1, 1]$, Z sei $\mathrm{N}(0, 1)$-verteilt und unabhängig von X. Es sei $Y := X^3 + \sigma Z$, mit $\sigma > 0$.
a) Finden Sie die im Sinn des erwarteten quadratischen Fehlers
i) beste ii) beste affin lineare
Prognose von Y auf der Basis von X,
d.h. jeweis eine Funktion $e : [-1, 1] \to \mathbb{R}$, für die $\mathbf{E}[(Y - e(X))^2]$ minimal wird, wobei in ii) die Form $e(x) = \beta_0 + \beta_1 x$ verlangt wird.
b) Was ist jeweils der erwartete quadratische Prognosefehler $\mathbf{E}[(Y - e(X))^2]$ und der erwartete quadratische Bias? (Siehe dazu Folie 11 in V9b.)

38. Aus der Vereinigung von drei disjunkten Populationen, deren Größen im Verhältnis 1 : 2 : 3 stehen, wird rein zufällig ein Individuum J gewählt und eine reelles Merkmal $h(J)$ beobachtet. In den drei Populationen ist das Merkmal der Individuen jeweils uniform verteilt auf einem Intervall, und zwar in der kleinsten Population (d.h. der mit den wenigsten Individuen) auf dem Intervall $[20, 50]$, in der größten Population auf $[30, 40]$, und in der zweitgrößten Population auf $[30, 80]$.
a) Was ist (i) der bedingte Erwartungswert, (ii) die bedingte Varianz,
von $h(J)$, gegeben dass J aus der Population mit den wenigsten Individuen gewählt wurde?
b) Berechnen Sie $\mathbf{Var}[h(J)]$, also die Varianz des Merkmals in der Gesamtpopulation.

39. S (Frei nach dem Eingangsbeispiel im 2. Vortrag der Ringvorlesung 2018/19)
https://www.mathe-uni-ffm.de/ringvorlesung/algorithmen-maschinelles-lernen-quantencomputing:
Jemand führt einen Münzwurf vor. Aus gewissen Gründen kommt nur in Frage, dass er entweder die ganze Zeit eine faire 01-Münze verwendet, oder eine mit $p = 0.9$. Bevor er beginnt, schätzen Sie die Wahrscheinlichkeit, dass er eine faire Münze verwendet, mit 0.8 ein. Wie aktualisieren Sie die Wahrscheinlichkeit, dass es sich um eine faire Münze handelt, nachdem
(i) beim ersten Wurf eine Eins
(ii) bei den ersten drei Würfen eine Eins
geworfen wurde?

40. a) U_1 und U_2 seien unabhängige, auf $[0, 1]$ uniform verteilte Zufallsvariable. Bestimmen Sie
(i) die Verteilungsfunktion (ii) die Dichte
von $Y := \max(U_1, U_2)$.
b) X_1, X_2 und X_3 seien unabhängig und Exp(1)-verteilt. Begründen Sie, warum die Zufallsvariable $X_1 + X_2$ gegeben $\{X_1 + X_2 + X_3 = 1\}$ so verteilt ist wie Y aus a). Argumentieren Sie
(i) anschaulich über einen Münzwurf mit kleiner Erfolgswahrscheinlichkeit durch Betrachten der Zeitpunkte der ersten drei Erfolge (vgl. V10a Abschnitte 2 und 3)
sowie
(ii) durch eine Rechnung analog zu der im Beispiel in V10 Abschnitt 4.

Zusatzaufgabe.[1] U_0, U_1, U_2, U_3 seien unabhängig und uniform auf $[0, 1]$ verteilt.
(a) Für $i = 0, 1, 2, 3$ setzen wir $R(i) := \sum_{j \in \{0,1,2,3\} \setminus \{i\}} I_{\{U_j < U_i\}}$ Begründen Sie, warum $(R(0), R(1), R(2), R(3))$ eine rein zufällige Permutation von 0,1,2,3 ist.
b) Bestimmen Sie $\mathbf{P}(U_1 < U_0, U_2 < U_0)$ und $\mathbf{P}(U_1 < U_0, U_2 < U_0, U_3 \geq U_0)$.
c) Wir definieren $Z_1 := I_{\{U_1 < U_0\}}, Z_2 := I_{\{U_2 < U_0\}}, Z_3 := I_{\{U_3 < U_0\}}$.
(i) Berechnen Sie $\mathbf{P}(Z_2 = 1 | Z_1 = 1)$ und $\mathbf{P}(Z_3 = 0 | Z_1 = 1, Z_2 = 1)$.
(ii) Tragen Sie die Gewichte der Übergangsverteilungen des durch (Z_1, Z_2, Z_3) beschriebenen 3-stufigen Experiments in einen Baum der Tiefe 3 ein.

Gesegnete Weihnachten und ein glückliches Jahr 2020!

[1] Ein Weihnachtsbonus: Ein schriftliches Bearbeiten dieser Aufgabe ist optional. Es wird wie auch die anderen S-Aufgaben auf die Bonuspunkte angerechnet.