

Übungen zur Vorlesung „Stochastik für die Informatik“

Abgabe der Lösungen zu den S-Aufgaben: Dienstag, 19. Januar 2016, vor der Vorlesung (10:05-10:15 im Magnus HS)

41 a) Für $n \in \mathbb{N}$ sei S_1 die Menge der Permutationen von $1, 2, \dots, n$ und S_2 die Menge der Permutationen von $1, 2, \dots, n+1$. Zwei Elemente $a_1 \in S_1$ und $a_2 \in S_2$ heißen *benachbart*, wenn die Zyklendarstellung von a_1 durch Wegstreichen des Elementes $n+1$ aus der Zyklendarstellung von a_2 entsteht. (Ein Beispiel für $n = 6$: Zwei der insgesamt 7 zu $a_1 := (1\ 5\ 3)(2\ 6)(4)$ benachbarten Elemente in S_2 sind $(1\ 5\ 7\ 3)(2\ 6)(4)$ und $(1\ 5\ 3)(2\ 6)(4)(7)$. Zur Notation der Zyklendarstellung vgl. Aufgabe 5.) X_1 sei uniform auf S_1 verteilt, und gegeben $\{X_1 = a_1\}$ sei X_2 uniform verteilt auf den zu a_1 benachbarten Elementen. Wie ist X_2 verteilt?

b) S_1 und S_2 seien gleich $S := \{0, 1\}^3$, der Menge der Ecken des 3-dimensionalen Einheitswürfels. Zwei Elemente a_1, a_2 von S heißen *benachbart*, wenn sie sich genau in einer der drei Komponenten unterscheiden (und in den anderen beiden übereinstimmen).

X_1 sei uniform auf S_1 verteilt, und gegeben $\{X_1 = a_1\}$ sei X_2 uniform verteilt auf den zu a_1 benachbarten Elementen. Wie ist X_2 verteilt? Wie ist (X_1, X_2) verteilt? Wie ist (X_2, X_1) verteilt?

42. a) X_1 sei $\text{Pois}(\lambda)$ -verteilt, und gegeben $X_1 = n$ sei $X_2 = (Z_1, \dots, Z_r)$ multinomialverteilt zu den Parametern $n, p_1, \dots, p_r$. Dann sind $Z_1, \dots, Z_r$ unabhängig, und Z_j ist $\text{Pois}(p_j \lambda)$ -verteilt, $j = 1, \dots, r$.

Begründen Sie dies, indem Sie zuerst (für $k_j \geq 0, k_1 + \dots + k_r = n$) die folgende Gleichheit nachvollziehen und sie dann als Produktformel für die Verteilungsgewichte der Z_j interpretieren:

$$\frac{1}{n!} \lambda^n e^{-\lambda} \frac{n!}{k_1! \dots k_r!} \prod_{j=1}^r p_j^{k_j} = \prod_{j=1}^r \left(\frac{1}{k_j!} (p_j \lambda)^{k_j} e^{-\lambda p_j} \right).$$

b) 10 Listen (nummeriert mit $j = 1, 2, \dots, 10$) haben die zufälligen Längen $Z_1, \dots, Z_{10}$. Anders gesagt: Die Anzahl der in Liste j eingetragenen Namen ist Z_j , und die Namen stehen in den Tiefen $0, \dots, Z_j - 1$. Die Z_j seien unabhängig und $\text{Pois}(j)$ -verteilt, $j = 1, \dots, 10$. Berechnen Sie mithilfe von Teil a) und nach dem in Vorlesung 10a vorgestellten Muster die erwartete Tiefe eines rein zufällig (aus der Gesamtheit der Namen in allen 10 Listen) gewählten Namens. (Verwenden Sie als erste Stufe jetzt eine $\text{Pois}(1 + 2 + \dots + 10)$ -verteilte Zufallsvariable X_1 , und nehmen Sie als Übergangswahrscheinlichkeiten von der ersten zur zweiten Stufe die Multinomialverteilungen mit Parameter $n, 1/55, \dots, 10/55$.)

43. S. X_1 sei uniform verteilt auf $[-1, 1]$; gegeben $X_1 = a_1$ habe X_2 die Verteilung $N(a_1^2, a_1^2/4)$. Finden Sie

a) die (im Sinn des mittleren quadratischen Abstandes) beste Prognose von X_2 , basierend auf X_1 , d.h. diejenige Funktion h , für die $\mathbf{E}[(X_2 - h(X_1))^2]$ minimal ist,

b) für eben dieses h den erwarteten quadratischen Prognosefehler $\mathbf{E}[(X_2 - h(X_1))^2]$,

c) die Varianz von X_2 .

Dabei dürfen Sie verwenden, dass alle in Vorlesung 10b für diskrete Zufallsvariable hergeleiteten Ergebnisse auch für kontinuierlich verteilte Zufallsvariable (mit endlichen Varianzen) gelten.

44. S. In drei Bevölkerungsgruppen H, N und U, die zusammen die Gesamtbevölkerung ausmachen, sind der Mittelwert und die Standardabweichung eines bestimmten individuellen Merkmals gegeben durch

$$\text{H: } 120 \pm 10 \quad \text{N: } 80 \pm 7 \quad \text{U: } 70 \pm 6$$

Die drei Gruppen H, N und U haben die Anteile 5%, 85% und 10% an der Gesamtpopulation. Berechnen Sie den Erwartungswert und die Varianz des Merkmals eines rein zufällig aus der Gesamtpopulation herausgegriffenen Individuums.

Gesegnete Weihnachten und ein glückliches 2016 !