Übungen zur Vorlesung „Stochastik für die Informatik"

Abgabe der Lösungen zu den S-Aufgaben: Freitag, 12. Januar 2018, vor der Vorlesung (12:10-12:15 im Magnus HS)

37. 15 Namen sind in 5 Listen einsortiert. Die Längen $Z_1, \dots, Z_5$ der Listen sind identisch verteilt und haben Varianz 16. Die Suchtiefen der in Liste j einsortierten Namen sind $0, 1, \dots, Z_j - 1$.
a) Finden Sie $\mathbf{E}[Z_j]$.
b) Aus den 15 Namen wird rein zufällig einer gewählt. Berechnen Sie den Erwartungswert seiner Suchtiefe.

38. S a) X sei Bernoulli(1/3)-verteilt. Gegeben $\{X = 0\}$ sei Y Exp(3)-verteilt, und gegeben $\{X = 1\}$ sei $-Y$ Exp(5)-verteilt. Berechnen Sie $\mathbf{E}[Y]$ und $\mathbf{Var}[Y]$.
b) Y sei uniform verteilt auf $[0, 1] \cup [10, 20]$. Berechnen Sie $\mathbf{E}[Y]$ und $\mathbf{Var}[Y]$.
Verwenden Sie sowohl in a) als auch in b) den Satz von der Zerlegung der Varianz. Was war nochmal die Varianz einer auf dem Intervall $[0, 1]$ *uniform verteilten ZV'en, und wie verhält sich die Varianz unter Verschiebung und Streckung?*

39. S a) Y sei N(3,1)-verteilt und N sei N(2,1)-verteilt; Y und N seien unabhängig.
(i) Wie ist $Y + N$ verteilt?
(ii) Finden Sie die bedingte Dichte von Y gegeben $\{Y + N = 5\}$.
Hinweis: Für die gemeinsame Dichtefunktion gilt
$f_{Y,Y+N}(a, 5) = f_{Y,N}(a, 5 - a) = \frac{1}{2\pi} e^{-\frac{(a-3)^2}{2}} e^{-\frac{(5-a-2)^2}{2}} = \cdots$. *Gefragt ist nach* $\frac{f_{Y,Y+N}(a, 5)}{f_{Y+N}(5)}\, da$.
(iii) Bestimmen Sie den bedingten Erwartungswert und die bedingte Varianz von Y, gegeben $\{Y + N = 5\}$.
(iv) Bestimmen Sie die (im Sinn des erwarteten quadratischen Fehlers) beste *lineare* Prognose von Y auf der Basis von $X := Y + N$. Welchen Wert nimmt sie an, wenn X auf den Ausgang $b = 5$ fällt?

b) Finden Sie ein einfaches Beispiel von zwei (diskreten) Zufallsvariablen X, Y, für die die beste *lineare* Prognose von Y auf der Basis von X verschieden ist von der besten Prognose von Y auf der Basis von X.
(Hier und auch im folgenden Teil c) ist das Gütekriterium jeweils die Kleinheit des erwarteten quadratischen Fehlers.)

c) (für Zusatzpunkte) Finden Sie ein Beispiel von zwei nicht unabhängigen Zufallsvariablen X, Y mit einer gemeinsamen Dichte, für die die beste *lineare* Prognose von Y auf der Basis von X und die beste Prognose von Y auf der Basis von X übereinstimmen.

40. Im Fachbereich A einer Universität war die Durchfallquote bei der Aufnahmsprüfung 0.4 bei den Männern und 0.3 bei den Frauen, im Fachbereich B war sie 0.8 bei den Männern und 0.7 bei den Frauen. (In beiden Fachbereichen haben Frauen damit jeweils besser abgeschnitten als Männer.) Paradoxerweise war die Durchfallsquote bei den Männern insgesamt (mit den beiden Fachbereichen zusammengenommen) niedriger, nämlich 0.5, im Gegensatz zu 0.6 bei den Frauen. Wie kann das sein? Berechnen Sie aus den Angaben, ein wie großer Prozentsatz der Männer (Frauen) im Fachbereich B angetreten ist, und formulieren Sie mit Blick darauf eine auch für alle Ihre Freunde verständliche Erklärung des Paradoxons. [1]

Gesegnete Weihnachten und ein glückliches Jahr 2018 !

[1] Die Zahlen hier sind erfunden. Eine ähnliche Geschichte hat sich allerdings tatsächlich einmal an der Universität Berkeley zugetragen und führte sogar zu einer Diskriminierungsklage, siehe https://de.wikipedia.org/wiki/Simpson-Paradoxon