Technische Universität Berlin

Fakultät II – Institut für Mathematik SS 06
Prof. Dr. J-D Deuschel 20. Juli 2006

Juli – Klausur

Stochastik für Informatiker

Name: .... Vorname: ....

Matr.–Nr.: .... Studiengang: ....

Als Hilfsmittel ist allein ein handbeschriebenes A4 Blatt mit Notizen zugelassen! Eine **Tabelle der Normalverteilungsfunktion ist auf Seite 4 abgedruckt**. Die Lösungen sind in **Reinschrift** auf A4 Blättern abzugeben. Mit Bleistift geschriebene Klausuren können **nicht** gewertet werden.

Geben Sie immer den **vollständigen Rechenweg** an.

Die Bearbeitungszeit beträgt **zweieinhalb Stunden**.

Korrektur

1	2	3	4	5	6	Σ

1. Aufgabe 6 Punkte

Im Seminarraum befinden sich 3 Sitzreihen mit je 8 Sitzplätzen. Unter den 15 (unterscheidbaren!) Studierenden gibt es 5, die sich immer in die erste Reihe setzen, und vier, die stets ganz hinten Platz nehmen. Wieviele verschiedene Sitzordnungen sind damit insgesamt noch möglich?

Loesung:
Die Anzahl der moeglichen Sitzverteilungen ist

$$\underbrace{\frac{8!}{(8-5)!}}_{\text{f. Studierende in Reihe 1}} \cdot \underbrace{\frac{8!}{(8-4)!}}_{\text{f. Studierende in Reihe 3}} \cdot \underbrace{\frac{(24-9)!}{((24-9)!-(15-9))!}}_{\text{f. restl. Studierende auf restl. Plaetze}}$$

2. Aufgabe 6 Punkte

Die Bevölkerung der Stadt S setze sich zusammen aus 20 % Studenten, 30% Schülern und 50% sonstigen. 80% aller Schüler koennen einen Web-Browser bedienen, von den Studenten sind es 60% und von den sonstigen 30%. Sie gehen in S in ein Internet-Cafe und treffen dort auf eine Person, die einen Web-Browser bedienen kann. Wie wahrscheinlich ist diese Person Schüler/Schülerin?

Loesung:
Wir betrachten die Ereignisse Stu, Sch, So und Web. Gefragt ist nach

$$P(\text{Sch}|\text{Web}) = \frac{P(\text{Sch};\text{Web})}{P(\text{Web})}$$

Mit den Zahlen aus der Aufgabenstellung ist

$$P(\text{Sch};\text{Web}) = P(\text{Web}|\text{Sch}) \cdot P(\text{Sch}) = 0.8 * 0.3 = 0.24$$

und

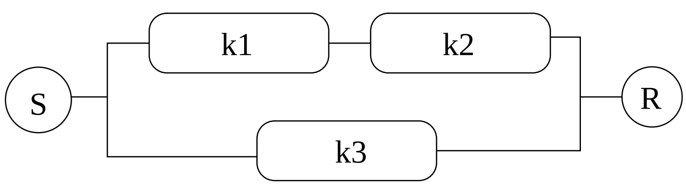
$$\begin{aligned} P(\text{Web}) &= P(\text{Web};\text{Sch}) + P(\text{Web};\text{Stu}) + P(\text{Web};\text{So}) \\ &= 0.8 * 0.3 + 0.6 * 0.2 + 0.3 * 0.5 = 0.51 \end{aligned}$$

Also

$$P(\text{Sch}|\text{Web}) = \frac{0.24}{0.51} = \frac{8}{17}.$$

3. Aufgabe 6 Punkte

Die Knoten k_1, k_2, k_3 des untenstehenden Datennetzes fallen unabhängig voneinander exponentialverteilt mit Parametern $\lambda_1, \lambda_2, \lambda_3$ aus. Berechnen Sie die Verteilung und den Erwartungswert der Ausfallzeit für die Datenübertragung von S nach R.



Loesung:
Die Ausfallzeit fuer die Datenuebertragung im Netz ist gegeben durch

$$\tau = \max(T_3, \min(T_2, T_1)),$$

wobei T_i die Ausfallzeiten der komponenten k_i seien. Es ist

$$P(\min(T_2, T_1) > t) = P(T_2 > t)P(T_1 > t) = e^{-\lambda_2 t}e^{-\lambda_1 t} = e^{-(\lambda_1+\lambda_2)t}$$

und damit

$$\begin{aligned} F_\tau(t) = P(\tau \le t) &= P(T_3 \le t)P(\min(T_2, T_1) \le t) \\ &= (1 - e^{-\lambda_3 t})(1 - e^{-(\lambda_1+\lambda_2)t}) \\ &= 1 - e^{-\lambda_3 t} - e^{-(\lambda_1+\lambda_2)t} + e^{-(\lambda_1+\lambda_2+\lambda_3)t}, \quad t \ge 0, \end{aligned}$$

sowie $F_\tau(t) = 0$ fals $t < 0$. Damit ist

$$G_\tau(t) = 1 - F_\tau(t) = e^{-\lambda_3 t} + e^{-(\lambda_1+\lambda_2)t} - e^{-(\lambda_1+\lambda_2+\lambda_3)t}$$

so dass

$$E(\tau) = \int_0^\tau G_\tau(s)ds = \frac{1}{\lambda_3} + \frac{1}{\lambda_1+\lambda_2} - \frac{1}{\lambda_1+\lambda_2+\lambda_3}.$$

4. Aufgabe 6 Punkte

Es sei X uniform auf $[-1,1]$ verteilt. Berechnen Sie die Wahrscheinlichkeitsdichte f_Y und den Erwartungswert $E(Y)$ von

$$Y := \mathrm{sign}(X)X^2,$$

wobei die Signumsfunktion $t \to \mathrm{sign}(t) \in \{-1,1\}$ die Vorzeichenfunktion auf $\mathbb{R}$ ist:

$$\mathrm{sign}(t) = -1 \text{ falls } t < 0 \text{ und } \mathrm{sign}(t) = 1, \text{ falls } t > 0, \text{ sowie } \mathrm{sign}(0) := 0.$$

Hinweis: Unterscheiden Sie die Fälle $X \geq 0$ und $X < 0$.

Loesung:
Es sei $t \geq 0$, dann ist

$$G_Y(t) := P(Y > t) := P(\mathrm{sign}(X)X^2 > t) = P(X > \sqrt{t}) = \frac{1-\sqrt{t}}{2}$$

Damit ist auf fuer $t \in]0,1[$

$$f_Y(t) = -G_Y'(t) = \frac{1}{2}\frac{1}{2\sqrt{t}} = \frac{1}{4\sqrt{t}}$$

Falls $t < 0$, dann ist

$$\begin{aligned} F_Y(t) &= P(Y \leq t) = P(-X^2 \leq t, X < 0) = P(|X| > \sqrt{-t}, X < 0) \\ &= P(X < -\sqrt{-t}, X < 0) = P(X < -\sqrt{-t}) = \frac{1-\sqrt{-t}}{2} \end{aligned}$$

Damit ist auf fuer $t \in]-1,0[$

$$f_Y(t) = F_Y'(t) = -\frac{1}{2}\frac{1}{2\sqrt{-t}} \cdot (-1) = \frac{1}{4\sqrt{-t}}$$

Wegen der Symmetrie von Y ist $E(Y) = 0$.

5. Aufgabe 6 Punkte

Es seien drei (reellwertige) Zufallsvariablen X_1, X_2, X_3 gegeben mit

$$E(X_i) = 0, \quad \forall\, i \in \{1,2,3\}$$

sowie

$$E(X_i \cdot X_j) = \frac{1}{i \cdot j} \quad \forall\, i,j \in \{1,2,3\}.$$

a) Berechnen Sie die Varianz von $Y = X_1 - \alpha X_2 - \beta X_3$, wobei die Parameter $\alpha, \beta \in \mathbb{R}$ beliebig aber fest gewählt seien.
b) Es sei $\alpha = 0$. Fuer welches β wird die Varianz von Y minimal?

Loesung:
a) Da die Erwartungswerte der X_i alle identisch null sind, ist auch

$$E(Y) = E(X_1 - \alpha X_2 - \beta X_3) = E(X_1) - \alpha E(X_2) - \beta E(X_3) = 0,$$

und damit

$$\begin{aligned} \mathrm{Var}(Y) &= E(Y^2) = E((X_1 - \alpha X_2 - \beta X_3)^2) \\ &= E\left[X_1^2 + \alpha^2 X_2^2 + \beta^2 X_3^2 - 2\left(\alpha X_1 \cdot X_2 + \beta X_1 \cdot X_3 - \alpha\beta X_2 \cdot X_3\right)\right] \\ &= 1 + \frac{\alpha^2}{4} + \frac{\beta^2}{9} - 2\left(\frac{\alpha}{2} + \frac{\beta}{3} - \frac{\alpha\beta}{6}\right) \end{aligned}$$

b) Falls $\alpha = 0$ ist

$$\text{Var}(Y) = 1 + \frac{\beta^2}{9} - \beta\frac{2}{3}$$

Als quadratische Funktion von β wird diese minimiert falls

$$\frac{2}{9}\beta = \frac{2}{3} \Leftrightarrow \beta = 3 \qquad \left(= \frac{\text{Kov}(X_1, X_3)}{\text{Var}(X_1)}\right)$$

6. Aufgabe 6 Punkte

Die Bearbeitung einer Datenbankabfrage dauert im Mittel 3ms ($=3{\cdot}10^{-3}$s), bei einer Streuung (Standardabweichung) von 2ms. Schätzen Sie die Wahrscheinlicheit, dass die Bearbeitung von 10000 (unabhängigen) aufeinanderfolgenden Abfragen länger als 30.1s dauert. (Verwenden Sie eine geeignete Approximation.)

Loesung:
Es sei X_i die Dauer der i-ten Abfrage, dann ist wegen der Unabhaengigkeit der X_i gemaess dem zentralen Grenzwertsatz die Gesamtdauer $S := \sum_{i=1}^{1000} X_i$ approximativ normalverteilt mit

$$E(S) = 10000 \cdot E(X_1) = 10000 * 3 \cdot 10^{-3}s = 30s$$

und Varianz

$$\begin{aligned}\text{Var}(S) &= \text{Var}(\sum_{i=1}^{1000} X_i) = 10000 \cdot \text{Var}(X_1) \\ &= 10000 \cdot \sigma^2(X_1) = 10000 * 4 \cdot 10^{-6}s^2 = 4 \cdot 10^{-2}s^2\end{aligned}$$

Also

$$\sigma(S) = \sqrt{\text{Var}(S)} = 2 \cdot 10^{-1} = 0.2s,$$

Damit ist die Standardisierung $\xi = (S - E(S))/\sigma(S)$ von S approximativ standardnormaverteilt, und somit

$$\begin{aligned}P(S > 30.1s) &= P(\frac{S - 30s}{0.2s} > \frac{0.1s}{0.2s}) \\ &= P(\xi > \frac{1}{2}) \\ &\sim P(\nu_{0,1} > \frac{1}{2}) = 1 - \Phi(0,5) = 30,85\%\end{aligned}$$