Technische Universität München
Institut für Informatik
Lehrstuhl für Effiziente Algorithmen
Prof. Dr. Ernst W. Mayr

SS 2001
Lösungsblatt Klausur
4. Aug. 2001

Diskrete Strukturen II

Aufgabe 1

(3 Punkte)
Berechnen Sie die Verteilungsfunktion einer Zufallsvariable X, die binomialverteilt ist mit Parametern $n = 2$ und $p = 1/3$. Skizzieren Sie das Ergebnis graphisch. Markieren Sie insbesondere Lage und Höhe der Sprungstellen. Wie groß ist die Wahrscheinlichkeit $\Pr[X \leq 1/7]$?

Lösung

Die Verteilungfunktion hat den Wert

$$F_X(t) = \begin{cases} 0 & t \text{ im Intervall}]-\infty, 0[, \\ (2/3)^2 & t \text{ im Intervall } [0, 1[, \\ (2/3)^2 + 2(2/3) \cdot 1/3 & t \text{ im Intervall } [1, 2[, \\ 1 & t \text{ im Intervall } [2, \infty[. \end{cases}$$

Daraus liest man sofort ab, daß $\Pr[X \leq 1/7] = F_X(1/7) = (2/3)^2$.

Aufgabe 2

(4 Punkte)

a) Vier Tetraeder, deren vier Seitenflächen jeweils mit den Zahlen $1, 1, 2, 4$ beschriftet sind, werden gleichzeitig auf den Tisch geworfen. Nehmen Sie an, daß jede Seite mit gleicher WS 1/4 unten zu liegen kommt. Bestimmen Sie Erwartungswert und Varianz der Summe $X := X_1 + X_2 + X_3 + X_4$ der Zahlen, die auf den unten liegenden Seiten der Tetraeder stehen.

b) Jetzt werfen Sie n Tetraeder wie oben. Normieren Sie die Summe $X^{(n)} := \sum_{i=1}^{n} X_i$, indem Sie den Erwartungswert $\mathbb{E}\left[X^{(n)}\right]$ abziehen und dann das Ergebnis durch die Standardabweichung $\sqrt{\mathrm{Var}\left[X^{(n)}\right]}$ teilen. Das Ergebnis sei mit $Y^{(n)}$ bezeichnet. Gegen welche Funktion konvergiert die Verteilungsfunktion von $Y^{(n)}$, wenn n gegen unendlich geht?

   Begründen Sie Ihre Antwort.

Lösung

zu a) Sei X_i die Augenzahl des i–ten Würfels, und $X = \sum_{i=1}^{4} X_i$. Es ist $\mathbb{E}[X] = 4\mathbb{E}[X_1] = 4(1/2 \cdot 1 + 1/4 \cdot 2 + 1/4 \cdot 4) = 8$, und $\mathrm{Var}[X] = 4\mathrm{Var}[X_1] = 4(1/2 \cdot (1-8/4)^2 + 1/4 \cdot (2-8/4)^2 + 1/4 \cdot (4-8/4)^2) = 6$.

zu b) Die Verteilungsfunktion $F_{Y^{(n)}}(\cdot)$ von $Y^{(n)}$ konvergiert punktweise gegen die Verteilungsfunktion $\Phi(\cdot, 0, 1)$ der Standardnormalverteilung. Dies ist die Aussage des Zentralen Grenzwertsatzes, der angewendet werden darf, weil die Varianzen der X_i endlich und nichttrivial sind.

Aufgabe 3

(2 Punkte) Betrachten Sie folgendes Spiel. Ein (schlechter) Schütze schießt gleichverteilt auf daß Intervall $[-1, 1]$. Je näher er der Null kommt, um so höher ist sein Preisgeld. Genauer gesagt, bekommt er $100 \cdot (1 - |x|)$ Mark, wenn er den Wert x trifft. Wieviele Mark beträgt sein erwarteter Gewinn?

Lösung

Sei $g(x) := 100 \cdot (1 - |x|)$, $\mu(x) \equiv 1/2$ ist die (kontinuierliche) Dichte der Zufallsvariable X, welche die Treffer beschreibt. Der erwarteter Gewinn ist

$$\mathbb{E}[g(X)] = \int_{-1}^{1} g(x) \cdot \mu(x)\, dx = \int_{-1}^{1} 100 \cdot (1 - |x|) \cdot 1/2\, dx = 2 \int_{0}^{1} 100 \cdot (1 - x) \cdot 1/2\, dx$$
$$= 100 - 100 \int_{0}^{1} x\, dx = 50.$$

Aufgabe 4

(1+2+2 Punkte)

Nehmen Sie an, daß die Dauer (gemessen in einer beliebigen Zeiteinheit) zwischen der Ankunft eines Telefongesprächs bei einer Telefonauskunft und der nächsten Ankunft exponentialverteilt ist.

a) Welche Eigenschaft der Exponentialverteilung macht dies zu einer sinnvollen Annahme?

b) Wenn wir uns eine beliebige Minute herausgreifen, dann sei die Wahrscheinlichkeit, daß in dieser Minute kein Anruf eingeht, gleich $e^{-5} = .0067379469(\ldots)$. Mit welcher Rate (Anrufe pro Minute im Mittel) rufen die Anrufer bei der Telefonauskunft an, und was ist der Parameter der obigen Exponentialverteilung?

c) Wie berechnet man die WS, daß in einer beliebig herausgegriffenen Minute genau 2 Anrufe eintreffen?

Lösung

zu a) Wir hatten in der Vorlesung gezeigt, daß die Exponentialverteilung dadurch charakterisiert werden kann, daß sie *gedächtnislos* ist. Wenn wir einen beliebigen Zeitpunkt wählen, dann ist die Wartezeit bis zum nächsten Anruf genauso verteilt, wie zu jedem anderen Zeitpunkt. Das ist wohl eine vernünftige Annahme, wenn wir mal davon absehen, daß reale Menschen zu gewissen Tageszeiten häufiger anrufen, als zu anderen.

zu b) Wir hatten besprochen, daß die Anzahl von Anrufern innerhalb eines festen Zeitintervalls, hier eine Minute, Poisson–verteilt ist mit Parameter μ, wenn die Abstände zwischen Anrufen exponentialverteilt sind mit Parameter $\lambda = 1/\mu$. Also ist $Po_\mu(0) = e^{-\mu} = e^{-5}$. Daraus folgt $\mu = 5$. Also ist der Parameter der Exponentialverteilung gleich $\lambda = 1/\mu = 1/5$.

zu c) Wir müssen $Po_\mu(2) = e^{-\mu}\mu^2/2!$ ausrechnen.

Bem.: Für $\mu = 5$ ist dies gleich $.0842243(\ldots)$.

Aufgabe 5

(3 Punkte)
Sei $f(x) = ce^{-2x}$ für $x > 0$ und $f(x) = 0$ sonst. Was muss für c gelten, damit f eine (kontinuierliche) Dichte ist? Bestimmen Sie $\Pr[X > 3]$.

Lösung

Es muss gelten $f(x) \geq 0$, was für alle $c \geq 0$ erfüllt ist, und

$$1 = \int_{-\infty}^{+\infty} f(x)dx = \int_0^{+\infty} c \cdot e^{-2x}dx = \left[-\tfrac{c}{2} \cdot e^{-2x}\right]_0^{+\infty} = \tfrac{c}{2}.$$

Also muss $c = 2$ sein. $\Pr[X > 3]$ berechnet sich durch Integration:

$$\Pr[X > 4] = \int_3^{+\infty} 2 \cdot e^{-2x}dx = [-e^{-2x}]_3^{+\infty} = e^{-6}.$$

Aufgabe 6

(3+3 Punkte)
a) Es seien zwei unabhängige Zufallsvariablen X und Y mit den (kontinuierlichen) Dichten f_X und f_Y gegeben; $W_X, W_Y = \mathbb{R}$. Ferner sei C eine von X und Y unabhängige Bernoulli-Zufallsvariable mit $\Pr[C = 1] = p$ und $\Pr[C = 0] = 1 - p$. Wir definieren

$$Z := \begin{cases} X & \text{falls } C = 1, \\ Y & \text{falls } C = 0. \end{cases}$$

Was ist der Wertebereich von Z? Geben Sie die (kontinuierliche) Dichte f_Z in Abhängigkeit von f_X und f_Y an.

b) Seien X und Y wie in a), aber wir betrachten nun $Z' := pX + (1 - p)Y$. Ist die Dichte von Z' gleich der von Z? Drücken Sie $\text{Var}[Z']$ durch $\text{Var}[X]$ und $\text{Var}[Y]$ aus.

Lösung

zu a) Es gilt

$$\begin{aligned} F_Z(t) &= \Pr[Z \le t] = \Pr[Z \le t \,|\, C = 1] \cdot \Pr[C = 1] \\ &\quad + \Pr[Z \le t \,|\, C = 0] \cdot \Pr[C = 0] \\ &= \Pr[X \le t] \cdot p + \Pr[Y \le t] \cdot (1-p) \\ &= p \cdot F_X(t) + (1-p) \cdot F_Y(t). \end{aligned}$$

Durch Differenzieren folgt $f_Z(t) = p \cdot f_X(t) + (1-p) \cdot f_Y(t)$.

zu b) Nein! Denn die Dichte von Z' bestimmt sich durch die Faltung der Dichten von pX und $(1-p)Y$, also $f_{Z'} = f_{pX} * f_{(1-p)Y} = p(1-p) f_X * f_Y$. Wegen der Unabhängigkeit von X und Y können wir jedoch die Varianz von Z' leicht berechnen:

$$\mathrm{Var}[Z'] = \mathrm{Var}[pX + (1-p)Y] = \mathrm{Var}[pX] + \mathrm{Var}[(1-p)Y] = p^2 \mathrm{Var}[X] + (1-p)^2 \mathrm{Var}[Y].$$

Aufgabe 7

(3 Punkte)
Wir wollen eine 'Punktmasse' an der Stelle x_0 durch eine kontinuierliche Dichte annähern. Dazu betrachten wir (für jedes $n \in \mathbb{N}$) eine Zufallsvariable $X^{(n)}$, die im Intervall $[x_0 - \frac{1}{2n}, x_0 + \frac{1}{2n}]$ gleichverteilt ist.

Geben Sie $f_{X^{(n)}}$ an. Für welche n nimmt $f_{X^{(n)}}$ Werte größer als 1 an?
Berechnen Sie die Verteilungsfunktion $F_{X^{(n)}}$, und berechnen Sie (für jedes $t \in \mathbb{R}$; Fallunterscheidung!) $\lim_{n\to\infty} F_{X^{(n)}}(t)$.
Ergibt $\lim_{n\to\infty} f_{X^{(n)}}$ eine kontinuierliche Dichte?

Lösung

Wegen $\int f_{X^{(n)}} \overset{!}{=} 1$, muß gelten

$$f_{X^{(n)}}(x) := \begin{cases} n & x \in [x_0 - \frac{1}{2n}, x_0 + \frac{1}{2n}] \\ 0 & \text{sonst} \end{cases}.$$

Das Maximum ist n, und dies kann beliebig groß werden; insbesondere nimmt die Dichte für alle $n > 1$ Werte größer als 1 an.
Die Verteilungsfunktion ist

$$F_{X^{(n)}}(x) := \begin{cases} 0 & x < x_0 - \frac{1}{2n} \\ n(x - x_0) + 1/2 & x \in [x_0 - \frac{1}{2n}, x_0 - \frac{1}{2n}] \\ 1 & x > x_0 + \frac{1}{2n} \end{cases}.$$

Dies konvergiert (punktweise) gegen

$$\lim_{n\to\infty} F_{X^{(n)}}(x) = \begin{cases} 0 & x < x_0 \\ 1/2 & x = x_0 \\ 1 & x > x_0 + \end{cases}.$$

Die (kontinuierliche) Dichte, die wir oben berechnet haben, konvergiert hingegen gegen keine (totale reellwertige) Funktion, weil sie an der Stelle x_0 divergiert.

Aufgabe 8

(3 Punkte)
Die Zufallsvariable X besitze die Dichte f_X. Drücken Sie die Verteilungsfunktion von $Y := aX + b$ durch f_X aus, wobei $a, b \in \mathbb{R}$. Wo möglich, drücken Sie auch die kontinuierliche Dichte von Y durch f_X aus. (Vollständige Fallunterscheidung!)

Lösung

Es gilt $F_Y(t) = \Pr[Y \leq t] = \Pr[aX \leq t - b]$. Nun müssen wir mehrere Fälle unterscheiden: Im Fall $a = 0$ gilt $F_Y(t) = 1$, wenn $t \geq b$ ist, und $F_Y(t) = 0$ sonst. Für $a > 0$ erhalten wir $F_Y(t) = \Pr[X \leq (t - b)/a] = F_X((t - b)/a)$. Durch Differenzieren folgt $f_Y(t) = (1/a) \cdot f_X((t - b)/a)$. Für den Fall $a < 0$ erhalten wir analog $F_Y(t) = 1 - F_X((t - b)/a)$ und somit $f_Y(t) = -(1/a) \cdot f_X((t - b)/a)$.

Aufgabe 9

(2 Punkte)
Ein Rechnernetz enthalte n Router, die jeder im Mittel t Zeiteinheiten zuverlässig laufen, bis es zu einem Absturz oder ähnlichen Problemen kommt. Wir nehmen an, dass die Zeitdauer bis zum Absturz eines einzelnen Routers exponentialverteilt ist. Die Abstürze der Router erfolgen unabhängig. Alle Router werden für einen reibungslosen Netzbetrieb benötigt. Geben Sie die Verteilung der Zeitdauer T bis zur ersten Störung des Netzes an und berechnen Sie $\mathbb{E}[T]$.

Lösung

Sei T_i die Zeitdauer bis zum ersten Absturz von Router i. T_i ist exponentialverteilt mit Parameter $\lambda_i = 1/t$. Da alle Router zum Betrieb benötigt werden, gilt $T = \min\{T_1, \ldots, T_n\}$. Wegen der Unabhängigkeit von $T_1, \ldots, T_n$ ist T wiederum exponentialverteilt mit Parameter $\lambda = \lambda_1 + \ldots + \lambda_n = n/t$. Somit gilt $\mathbb{E}[T] = 1/\lambda = t/n$.

Aufgabe 10

(2 Punkte)
Seien $X_1, \ldots, X_n$ unabhängig wie X verteilte Stichprobenvariablen. Zeigen Sie, dass $Y := \sum_{i=1}^{n} \lambda_i X_i$ für beliebige Werte $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ mit $\lambda_1 + \ldots + \lambda_n = 1$ einen erwartungstreuen Schätzer für $\mathbb{E}[X]$ darstellt.

Lösung

Wegen der Linearität des Erwartungswerts gilt

$$\mathbb{E}[Y] = (\lambda_1 + \ldots + \lambda_n) \cdot \mathbb{E}[X] = \mathbb{E}[X].$$

Aufgabe 11

(4 Punkte)
Seien $X_1, \dots, X_n$ unabhängig wie X verteilte Stichprobenvariablen. Dabei sei X exponentialverteilt mit unbekanntem Parameter λ. Berechnen Sie einen ML-Schätzer für λ. (Überlegen Sie sich ggf. zuvor, welches Ergebnis Sie intuitiv für plausibel halten.)

Lösung

Die Likelihood-Funktion lautet

$$L(\vec{x}, \lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i} = \lambda^n \exp\left(-\lambda \sum_{i=1}^{n} x_i\right).$$

Zur Bestimmung des Maximums berechnen wir die Ableitung

$$\frac{\partial L}{\partial \lambda} = n\lambda^{n-1}E - \left(\sum_{i=1}^{n} x_i\right) \lambda^n E$$

mit $E := \exp\left(-\lambda \sum_{i=1}^{n} x_i\right)$. Durch Nullsetzen erhalten wir

$$n = \lambda \left(\sum_{i=1}^{n} x_i\right) \iff \lambda = 1/\overline{x}.$$

Aufgabe 12

(3 Punkte)
Sie besitzen zwei Münzen, für die erste ist $\Pr[\text{Zahl}] = 0.4$ und für die zweite $\Pr[\text{Zahl}] = 0.5$. Dabei kürzen wir zur Vereinfachung 'Kopf' durch '0' ab und 'Zahl' durch '1'. Sie nehmen eine der Münzen und erzeugen eine Stichprobe $\vec{x} = (x_1, \dots, x_n)$ in $\{0, 1\}^n$, indem Sie die Münze n mal werfen. Sie wissen aber nicht, welche der beiden Münzen Sie verwendet haben, weil diese äußerlich völlig gleich aussehen.
Sie berechnen nun die Schätzvariable $\hat{p} = \hat{p}(\vec{x}) := \frac{1}{n} \sum_{i=1}^{n} x_i$. Wenn $\hat{p} \leq 0.45$, dann akzeptieren Sie die Hypothese *'Die erste Münze wurde benutzt.'*, anderenfalls verwerfen Sie die Hypothese.

a) Wie wahrscheinlich ist ein *Fehler der zweiten Art*, daß nämlich die zweite Münze verwendet wurde, Sie sich aber dennoch für die erste entscheiden? Wie wahrscheinlich ist ein *Fehler der ersten Art*, daß die erste Münze geworfen wurde, Sie aber fälschlicherweise sagen, daß es die zweite war?

   Sie brauchen die Werte nicht exakt ausrechnen. Fertigen Sie eine Skizze an, in der die Dichte von $\hat{p}$ eingezeichnet wird, einmal unter der Bedingung, daß die erste Münze verwendet wurde, wie auch unter der Bedingung, daß die zweite Münze verwendet wurde. Weiterhin dürfen die Dichten in der Skizze ähnlich wie Dichten von normalverteilten ZV aussehen, was durch den Zentralen Grenzwertsatz ja für große n gerechtfertigt ist.

   Schraffieren Sie in Ihrer Skizze geeignete *Tails*, so daß die schraffierten Flächen jeweils den gesuchten Wahrscheinlichkeiten entsprechen. [1]

[1] Diese Tails könnte man natürlich entweder (asymptotisch) exakt mit Hilfe des Zentralen Grenzwertsatzes ausrechen oder durch die Ungleichungen von Chernoff oder Chebychev abschätzen. Wir gehen davon aus, daß Sie das können.

b) Haben beide Fehler (zumindest asymptotisch) denselben Wert?

Lösung

zu a) **Fehler zweiter Art:** Wir müssen $\Pr\left[\frac{1}{n}\sum_{i=1}^{n} X_i \leq 0.45\right]$ berechnen, unter der Annahme, daß die X_i Bernoulli-Variablen sind mit Parameter $p = 0.5$. Dies ist gleich

$$\Pr[X \leq n \cdot 0.45],$$

wobei $X := \sum_{i=1}^{n} X_i$ eine binomialverteilte ZV ist mit Parametern n und 0.5.

Fehler erster Art: Wir müssen $\Pr\left[\frac{1}{n}\sum_{i=1}^{n} Y_i > 0.45\right]$ berechnen, unter der Annahme, daß die Y_i Bernoulli-Variablen sind mit Parameter $p' = 0.5$. Dies ist gleich

$$\Pr[Y > n \cdot 0.45],$$

wobei $Y := \sum_{i=1}^{n} Y_i$ eine binomialverteilte ZV ist mit Parametern n und 0.4.

Die Skizze sieht aus wie folgt: Wir zeichnen zwei Kurven, die jeweil in etwa die Form einer Gauß'schen Dichte haben (vgl. 10 DM Schein). Die eine hat ihren Schwerpunkt bei 0.4, die andere bei 0.5. Dann ziehn wir einen senkrechten Strich bei $x = 0.45$. Der Fehler zweiter Art ist die Fläche, die links vom Strich unter der zweitgenannten Kurve liegt, der Fehler erster Art die Flache, die rechts vom Strich unter der erstgenannten Kurve liegt.

Mit wachsendem n werden sich die beiden Kurven immer stärker um 0.4 bzw. 0.5 'zusammenziehen', und die markierten Tails werden sehr schnell kleiner werden.

zu b) Man beachte, daß die Fehler erster und zweiter Art vermutlich ähnlich klein sind, aber nicht exakt gleich, weil die Dichten nicht exakt 'die gleiche Form' haben. Selbst wenn wir die Tails näherungsweise mit dem Zentralen Grenzwertsatz berechnen, so werden wir nicht zwei gleiche Werte bekommen, weil sich X und Y u.a. bezüglich ihrer Varianz unterscheiden.