Wolfgang K. Seiler
Tel. 2515

A5, C201

20. Juni 2001

9. Übungsblatt zur Einführung in die Wahrscheinlichkeitstheorie

Fragen: *(je ein Punkt)*
Die Antworten auf die nachfolgenden Fragen sollten nicht länger als etwa zwei Zeilen sein und lediglich eine kurze *Begründung enthalten. Antworten ohne Begründung werden nicht gewertet.*

1) Richtig oder falsch: X, Y seien beliebige, nicht notwendigerweise unabhängige Zufallsvariablen. Dann ist $V(X+Y) \geq V(X) + V(Y)$.

Lösung: *Falsch:* Sei beim Würfeln etwa X gleich der geworfenen Augenzahl und $Y = 7-X$. Dann hat $X + Y = 7$ Varianz null, aber $V(X) = V(Y) = 35/2$.

2) Richtig oder falsch: X, Y seien beliebige, nicht notwendigerweise unabhängige Zufallsvariablen. Dann ist $V(X+Y) \leq V(X) + V(Y)$.

Lösung: *Falsch:* Für $X = Y$ und $V(X) > 0$ ist

$$V(X+Y) = V(2X) = \mathbb{E}\Big(\big(2X - \mathbb{E}(2X)\big)^2\Big) = 4V(X) > 2V(X) = V(X) + V(Y)\,.$$

NB: Die korrekte Formel ist, wie wir beim Beweis des Satzes von Bienaymé gesehen haben,

$$V(X+Y) = V(X) + V(Y) + 2\,\mathrm{Cov}(X,Y)\,,$$

und $\mathrm{Cov}(X,Y)$ kann positiv, negativ oder null sein.

3) Sei $f(x) = 1$ für $x \in [0, 2]$ und $f(x) = 0$ sonst. Was ist $(f * f)(1)$?

Lösung: Nach Definition ist $f * f(1) = \int\limits_{-\infty}^{\infty} f(1-y)f(y)\,dy$. Der Integrand ist eins für $y \in [0, 2]$ und $1-y \in [0, 2]$, also für $y \in [0, 1]$, und null sonst. Somit ist $f*f(1) = \int\limits_0^1 dy = 1$.

4) Richtig oder falsch: Die Summe zweier unabhängiger gleichverteilter kontinuierlicher Zufallsvariablen ist wieder gleichverteilt.

Lösung: *Falsch:* Sind X und Y etwa beide gleichverteilt auf $[0, 1]$, so nimmt $X+Y$ Werte aus $[0, 2]$ an mit der Dichtefunktion $f * f$ aus der vorigen Frage. Diese ist null an den Randpunkten, da es dort nur eine Art der Darstellung gibt ($0 = 0 + 0$ und $2 = 1 + 1$), aber nach Frage drei eins in der Intervallmitte.

5) Richtig oder falsch: Die Summe zweier unabhängiger gleichverteilter diskreter Zufallsvariablen ist wieder gleichverteilt.

Lösung: *Falsch:* Beim Werfen zweier Würfel liegt die Summe zwischen zwei und zwölf, wobei die Randwerte nur auf eine Art und Weise zustandekommen können, der mittlere Wert sieben aber auf sechs.

Aufgabe 1: *(5 Punkte)*

Die Betriebsdauer eines bestimmten Notebooks mit einem vollem Akku sei normalverteilt mit Mittelwert vier Stunden und Standardabweichung ($=\sqrt{\text{Varianz}}$) eine Stunde.

a) T sei die mögliche Betriebsdauer des Notebooks mit vollem Akku plus zwei vollen Reserveakkus. Welche Verteilung hat T?

Lösung: Die Abweichungen der Betriebsdauern mit den verschiedenen Akkus vom Mittelwert sollten unabhängig voneinander sein, wir können T also beschreiben als Summe dreier voneinander unabhängiger Zufallsvariablen mit Erwartungswert drei Stunden und Standardabweichung eine Stunde. Da die Summe zweier und damit beliebig vieler voneinander unabhängiger normalverteilter Zufallsvariablen wieder normalverteilt ist und sich Erwartungswerte und Varianzen bei der Summierung einfach addieren, ist T normalverteilt mit Erwartungswert zwölf Stunden und Varianz drei „Quadratstunden", also Standardabweichung $\sigma = \sqrt{3}$ Stunden.

b) Sie wollen mit 99%-iger Wahrscheinlichkeit garantieren, daß das Notebook für eine Betriebsdauer von 12 Stunden zur Verfügung steht. Wie viele Reserveakkus brauchen Sie?

Lösung: Zwei Reserveakkus reichen sicherlich nicht, denn dann ist die Gesamtbetriebsdauer normalverteilt mit Erwartungswert zwölf und der Wert zwölf wird wird wegen der Symmetrie der Normalverteilung in Bezug auf ihren Mittelwert nur mit einer einer Wahrscheinlichkeit von 50% erreicht oder überschritten.

Für *drei* Reserveakkus ist die Gesamtbetriebsdauer aus den bei der Lösung von *a)* dargelegten Gründen normalverteilt mit Erwartungswert sechzehn und Standardabweichung $\sqrt{4} = 2$. Eine Betriebsdauer von mindestens zwölf Stunden wird also genau dann erreicht, wenn die Zufallsvariable nicht um mehr als zwei Standardabweichungen nach unten von ihrem Erwartungswert abweicht.

Bezeichnet $F(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-t^2/2}\,dt$ die Verteilungsfunktion der Standardnormalverteilung, so ist die Wahrscheinlichkeit des komplementären Ereignisses gleich $F(-2) = 1-F(2)$, die gesuchte Wahrscheinlichkeit ist nach der Tabelle auf der Rückseite des fünften Übungsblatts

$$F(2) \approx 0{,}977250\,,$$

was leider noch deutlich unter den gewünschten 99% liegt.

Mit *vier* Reserveakkus ist die Gesamtbetriebsdauer normalverteilt mit Erwartungswert zwanzig und Standardabweichung $\sigma = \sqrt{5}$. Um kleiner als zwölf Stunden zu sein, müßte sie also um acht Stunden oder $8/\sqrt{5} \approx 3{,}5777$ Standardabweichungen nach unten vom Erwartungswert abweichen, was mit Wahrscheinlichkeit $F(8/\sqrt{5})$ passiert. Leider geht die Tabelle nur knapp über drei Standardabweichungen, aber da schon

$$F(3) \approx 0{,}998650$$

über 99% liegt und F eine monoton wachsende Funktion ist, liegt auch $F(8/\sqrt{5})$ über 99%. Tatsächlich ist übrigens

$$F(8/\sqrt{5}) \approx 0{,}9998266904\,,$$

die Wahrscheinlichkeit, daß die zwölf Stunden erreicht werden, liegt also sogar bei über 99,98%.

c) Wäre es realistischer gewesen, die Betriebsdauer mit einem Akku durch eine Exponentialverteilung zu modellieren?

Lösung: Auf keinen Fall! Wie wir bei Aufgabe 1*b* des fünften Übungsblatts gesehen haben, ist bei einer exponentialverteilten Zufallsvariablen X die bedingte Wahrscheinlichkeit für

$X \geq t+\tau$ unter der Voraussetzung $X \geq t$ unabhängig von t. Bei einem Akku wird man aber erwarten, daß die Wahrscheinlichkeit des Leerwerdens in den nächsten zehn Minuten für einen fast vollen Akku (t nahe null) deutlich kleiner ist als für einen Akku, der bereits mehrere Stunden Strom lieferte.

Aufgabe 2: *(5 Punkte)*

a) Die beiden voneinander unabhängigen Zufallsvariablen X und Y seien exponentialverteilt mit Parametern λ und μ. Berechnen Sie die Verteilungsdichte von $X+Y$!

Lösung: X und Y haben Verteilungsdichten $\varphi_X(t)$ und $\varphi_Y(t)$, die für nichtnegative t gleich $\lambda e^{-\lambda t}$ bzw. $\mu e^{-\mu t}$ sind und für negative Werte von t verschwinden. Die Verteilungsdichte von $X+Y$ ist die Faltung

$$\varphi_X * \varphi_Y(x) = \int_{-\infty}^{\infty} \varphi_X(x-y)\varphi_Y(y)\,dy\,,$$

deren Integrand für negative y sowie für $y > x$ verschwindet. Insbesondere verschwindet das Integral also für $x < 0$, und für $x \geq 0$ ist

$$\begin{aligned}\varphi_X * \varphi_Y(x) &= \int_0^x \lambda e^{-\lambda(x-y)} \cdot \mu e^{-\mu y}\,dy = \lambda\mu e^{-\lambda x}\int_0^x e^{(\lambda-\mu)y}\,dy \\ &= \begin{cases} \frac{\lambda\mu e^{-\lambda x}}{\lambda-\mu}\left(e^{(\lambda-\mu)x-1}\right) = \frac{\lambda\mu}{\lambda-\mu}\left(e^{-\mu x}-e^{-\lambda x}\right) & \text{für } \lambda \neq \mu \\ \lambda^2 x e^{-\lambda x} & \text{für } \lambda = \mu \end{cases}\end{aligned}$$

b) Die Zufallsvariablen $Z_1, \ldots, Z_n$ seien unabhängig; jede einzelne sei exponentialverteilt mit Parameter λ. Zeigen Sie: Die Wahrscheinlichkeitsdichte für die Summe Z der Z_i ist

$$\varphi(t) = \lambda\frac{(\lambda t)^{n-1}}{(n-1)!}e^{-\lambda t}\,.$$

Lösung: Die Formel für die Wahrscheinlichkeitsdichte kann beispielsweise durch vollständige Induktion bewiesen werden: Der Fall $n=1$ ist klar, und auch $n=2$ wurde bereits in *a)* miterledigt.

Die Behauptung sei also richtig für eine gewisse Anzahl $n \geq 1$ von Zufallsvariablen, deren Summe Z sei; für eine weitere Zufallsvariable Z_{n+1} sei $Z' = Z + Z_{n+1}$. Die Wahrscheinlichkeitsdichte $\varphi_{Z'}$ von Z' ist dann die Faltung

$$\varphi_{Z'}(x) = \left(\varphi_Z * \varphi_{Z_{n+1}}\right)(x) = \int_{-\infty}^{\infty} \varphi_Z(x-y)\varphi_{Z_{n+1}}(y)\,dy\,,$$

und wie in *a)* verschwindet das Integral für $x < 0$ und der Integrand für alle y, die nicht zwischen null und x liegen. Somit ist

$$\varphi_{Z'}(x) = \int_0^x \lambda\frac{\big(\lambda(x-y)\big)^{n-1}}{(n-1)!}e^{-\lambda(x-y)} \cdot \lambda e^{-\lambda y}\,dy$$

Dies schreit natürlich geradezu nach entweder der Substitution $u = x-y$ oder aber der Anwendung des Kommutativgesetzes für Faltungen; der Aufwand dürfte in beiden Fällen vergleichbar sein. Beim ersten Weg ist $du = -dy$; andererseits läuft u von x nach 0, wenn y von 0 nach x läuft, was ebenfalls einen Vorzeichenwechsel bedingt; somit ist

$$\varphi_{Z'}(x) = \int_0^x \lambda \frac{(\lambda u)^{n-1}}{(n-1)!} e^{-\lambda u} \cdot \lambda e^{-\lambda(x-u)}\, du = \lambda e^{-\lambda x} \int_0^x \lambda \frac{(\lambda u)^{n-1}}{(n-1)!}\, du = \lambda e^{-\lambda x} \frac{(\lambda x)^n}{n!},$$

was genau das gewünschte Ergebnis ist. Die berechnete Verteilung heißt übrigens (spezielle) ERLANG-Verteilung der Stufe n.

c) Welchen Erwartungswert und welche Varianz hat Z?

Lösung: Wie vom siebten Übungsblatt her bekannt ist, hat jedes Z_i Erwartungsert $1/\lambda$ und Varianz $1/\lambda^2$; somit ist wegen der Linearität der Integration/Erwartungswertbildung und des Satzes von BIENAYMÉ

$$\mathbb{E}(Z) = \frac{n}{\lambda} \quad \text{und} \quad V(Z) = \frac{n}{\lambda^2}.$$

Aufgabe 3: (*5 Punkte*)

a) Die Summe zweier POISSON-verteilter Zufallsvariablen ist wieder POISSON-verteilt.

Lösung: Die beiden Zufallsvariablen X, Y seien POISSON-verteilt mit Parametern λ und μ; die Wahrscheinlichkeiten der Ereignisse $X = k$ und $Y = \ell$ sind also

$$e^{-\lambda} \frac{\lambda^k}{k!} \quad \text{und} \quad e^{-\mu} \frac{\mu^\ell}{\ell!}.$$

Das Ereigniss $X + Y = k$ ist zusammengesetzt aus den Elementarereignissen

$$X = 0, Y = k, \quad X = 1, Y = k-1, \quad \ldots, \quad X = k, Y = 0$$

und hat somit die Wahrscheinlichkeit

$$\begin{aligned}\sum_{\ell=0}^{k} e^{-\lambda} \frac{\lambda^\ell}{\ell!} \cdot e^{-\mu} \frac{\mu^{k-\ell}}{(k-\ell)!} &= e^{-\lambda-\mu} \sum_{\ell=0}^{k} \frac{\lambda^\ell \mu^{k-\ell}}{\ell!(k-\ell)!} = e^{-\lambda-\mu} \sum_{\ell=0}^{k} \frac{1}{k!} \binom{\ell}{k} \lambda^\ell \mu^{k-\ell} \\ &= \frac{e^{-(\lambda+\mu)}}{k!} \sum_{\ell=0}^{k} \binom{\ell}{k} \lambda^\ell \mu^{k-\ell} = e^{-(\lambda+\mu)} \frac{(\lambda+\mu)^k}{k!},\end{aligned}$$

die (diskrete) Faltung der beiden obigen Wahrscheinlichkeitsverteilngen. Daher ist $X + Y$ POISSON-verteilt mit Parameter $\lambda + \mu$.

b) Wie von Blatt sechs bekannt, wurden für zehn preußische Kavallerieregimentern über zwanzig Jahre hinweg die jährlichen Todesfälle durch Hufschlag gezählt; die Zahlen pro Jahr und Regiment sind mit guter Genauigkeit POISSON-verteilt mit Parameter $\lambda = 0{,}61$. Welche Verteilung beschreibt die Gesamtzahl der in einem Jahr durch Hufschlag getöteten Offiziere aus allen zehn Regimentern?

Lösung: Da die Summe zweier POISSON-verteilter Zufallsvariablen wieder POISSON-verteilt ist, folgt induktiv, daß auch die Summe von n POISSON-verteilten Zufallsvariablen ist wieder POISSON-verteilt ist. Der Parameter λ der POISSON-Verteilung ist gleich dem Erwartungswert, und Erwartungswerte addieren sich; also ist der Parameter der Summenverteilung gleich der Summe der Parameter der Summanden. Hier geht es um zehn Verteilungen mit identischem Parameter $\lambda = 0{,}61$; die Summe ist also POISSON-verteilt mit Parameter 6,1.