WOLFGANG K. SEILER
Tel. 2515

A5, ZIMMER C 201

13. Juli 2001

Klausur zur Einführung in die Wahrscheinlichkeitstheorie

• • • Schreiben Sie bitte auf jedes Blatt Ihren Namen! • • •

Fragen: *(je zwei Punkte)*
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) Auf wie viele Arten kann man auf einer zweielementigen Menge eine σ-Algebra definieren?

Lösung: Eine σ-Algebra einhält mindestens die leere Menge und die Gesamtmenge. Falls sie noch eine weitere Menge enthält, muß sie auch deren Komplement enthalten, also *alle* Teilmengen. Somit gibt es genau zwei σ-Algebren.

2) Richtig oder falsch: Mit $X\colon \Omega \to \mathbb{R}$ ist auch $Y = \sin X$ meßbar.

Lösung: *Richtig,* denn stetige Funktionen sind meßbar und die Hintereinanderausführung zweier meßbarer Funktionen ist meßbar.

3) $\Omega = \{1, \dots, 10\}$ bestehe aus den ersten zehn natürlichen Zahlen, und auf $\mathfrak{P}(\Omega)$ sei ein Maß μ definiert, das jeder Teilmenge deren Elementanzahl dividiert durch zehn zuordnet. Was ist $\int_\Omega x^2 \, d\mu$?

Lösung:

$$\int_\Omega x^2 \, d\mu = \frac{1}{10} \sum_{x=1}^{1} 0x^2 = \frac{10 \cdot 11 \cdot 21}{6 \cdot 10} = \frac{11 \cdot 7}{2} = \frac{77}{2}$$

4) Die Zufallsvariable X sei normalverteilt mit Erwartungswert null und Varianz $1/2$. Was ist $\operatorname{Cov}(X, X^2)$?

Lösung:

$$\operatorname{Cov}(X, X^2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x \cdot x^2 \cdot e^{-x^2/(2\cdot\frac{1}{2})} \, dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^3 e^{-x^2} \, dx = 0\,,$$

denn der Integrand ist eine ungerade Funktion (die für $x \to \pm\infty$ so schnell gegen null geht, daß es keine Probleme mit der Konvergenz gibt).

5) Sei $f(t) = t$ und $g(t) = \begin{cases} e^{-t} & \text{für } t \geq 0 \\ 0 & \text{sonst} \end{cases}$. Was ist $f * g$?

Lösung:

$$f * g(s) = \int\limits_{-\infty}^{\infty} f(s-t)g(t)\,dt = \int\limits_{0}^{\infty} (s-t)e^{-t}\,dt = s\int\limits_{0}^{\infty} e^{-t}\,dt - \int\limits_{0}^{\infty} te^{-t}\,dt$$

$$= se^{-t}\Big|_{0}^{\infty} + e^{-t}(t+1)\Big|_{0}^{\infty} = s-1$$

6) Sie wollen die Spannung einer Batterie mit einer Genauigkeit von $\pm 0{,}01\,\mathrm{V}$ messen, haben aber nur ein Voltmeter, mit einer Meßgenauigkeit von $\pm 0{,}02\,\mathrm{V}$. Wie gehen Sie vor?

Lösung: Nimmt man den Mittelwert von n Messungen, wird der Fehler um den Faktor $\sqrt{n}$ kleiner. Hier soll der Fehler halbiert werden, also nimmt man den Mittelwert aus vier Messungen.

7) *Richtig oder falsch:* Der stochastische Prozeß $(X_t)_{t\in T}$ sei ein Martingal. Dann ist auch $(X_t^2)_{t\in T}$ ein Martingal.

Lösung: *Falsch:* Der Prozeß $(X_t)_{t\in\mathbb{N}_0}$ mit $X_0 = 0$ und $X_{t+1} = X_t \pm 1$, je nachdem ob eine zur Zeit t geworfene Münze *Kopf* oder *Zahl* zeigt, ist ein Martingal. Für sein Quadrat ist aber $X_{t+1}^2 = X_t^2 + 2X_t + 1$ wegen des Summanden eins im Mittel ansteigend, also kein Martingal.

8) *Richtig oder falsch:* Der stochastische Prozeß $(X_t)_{t\in T}$ sei ein MARKOV-Prozeß. Dann ist auch $(X_t^2)_{t\in T}$ ein MARKOV-Prozeß.

Lösung: *Richtig,* denn wenn $\mathbb{E}\big(X_s \mid \{X_\tau | \tau \le t\}\big)$ für $s > t$ nicht von den X_τ mit $\tau < t$ abhängt, gilt dasselbe auch für X_t^2.

Aufgabe 1: (*6 Punkte*)

Eine Lostrommel enthält, jeweils in großer Zahl, Lose mit Gewinnwerten von fünf, fünfzehn und vierzig Mark sowie Nieten. Alle vier Arten von Losen sind mit gleicher Häufigkeit vorhanden.

Sie nehmen eine Handvoll Lose heraus; es sind zehn Stück.

a) Wie viele Möglichkeiten gibt es für deren Verteilung auf die vier Klassen?

Lösung: Dies ist eine ungeordnete Probe vom Umfang zehn mit Wiederholungen aus der Grundmenge, die aus den vier Gewinnklassen besteht; die Anzahl der Möglichkeiten ist also

$$\binom{4+10-1}{10} = \binom{13}{10} = \frac{13\cdot 12\cdot 11}{3!} = 13\cdot 2\cdot 11 = 286\,.$$

b) Wie groß ist der Erwartungswert für Ihren Gewinn, falls in der Lostrommel alle vier Gewinnklassen mit gleicher Häufigkeit vertreten sind?

Lösung: Der zu erwartende Gewinn pro Los ist dann

$$\frac{0+5+15+40}{4}\,\mathrm{DM} = \frac{60}{4}\,\mathrm{DM} = 15\,\mathrm{DM}\,,$$

bei zehn Losen ist also ein Gesamtgewinn von 150 Mark zu erwarten.

c) N bezeichne die Anzahl der von Ihnen gezogenen Nieten und X Ihren Gesamtgewinn. Was ist $\mathbb{E}(X|N)$?

Lösung: Bei N Nieten gibt es $10 - N$ Lose mit Gewinn; da unter diesen alle drei echten Gewinnklassen mit gleicher Wahrscheinlichkeit vorkommen, ist der erwartete Gewinn einer Nicht-Niete gleich

$$\frac{5+15+40}{3}\,\mathrm{DM} = \frac{60}{3}\,\mathrm{DM} = 20\,\mathrm{DM}\,,$$

bei $10 - N$ Gewinnlosen ist also ein Gesamtgewinn von $(10 - N)\cdot 20$ DM zu erwarten:

$$\mathbb{E}(X|N) = (10 - N)\cdot 20\,\mathrm{DM}\,.$$

Aufgabe 2: (*5 Punkte*)

Ein AIDS-Test liefert bei allen Infizierten und bei 0,01% der Nichtinfizierten ein positives Ergebnis. Von Personen, die keiner der bekannten Risikogruppen angehören, sind etwa 0,01% infiziert. Wie groß ist die Wahrscheinlichkeit dafür, daß eine solche Person, bei der ein positives Testergebnis auftritt, tatsächlich infiziert ist?

Lösung: T sei das Ereignis *Testergebnis positiv,* I das Ereignis *infiziert* und N das Ereignis *nicht infiziert.* Für Personen, die keiner Risikogruppe angehören, kennen wir die Wahrscheinlichkeiten

$$P(I) = 0{,}01\%\,,\quad P(N) = 1 - P(I) = 99{,}99\%\,,\quad P(T|I) = 1 \quad\text{und}\quad P(T|N) = 0{,}01\%\,.$$

Nach der BAYESschen Formel ist daher

$$\begin{aligned} P(I|T) &= \frac{P(T|I)\cdot P(I)}{P(T|I)\cdot P(I) + P(T|N)\cdot P(N)} = \frac{1\cdot 0{,}01\%}{1\cdot 0{,}01\% + 0{,}01\%\cdot 99{,}99\%} = \frac{1}{1+99{,}99\%} \\ &= \frac{1}{1{,}9999} \approx 0{,}500025\,. \end{aligned}$$

Die Wahrscheinlichkeit für eine Infektion liegt also nur ganz minimal über 50%.

Aufgabe 3: (*6 Punkte*)

a) Die Lebensdauer einer Energiesparlampe sei exponentialverteilt mit Parameter $\lambda = 1/4000$, die einer zweiten mit Parameter $\mu = 1/5000$. Wie groß ist die Wahrscheinlichkeit, daß die zweite Lampe länger brennt als die erste?

Lösung: Die Wahrscheinlichkeitsdichten der beiden Verteilungen sind

$$\lambda e^{-\lambda t} \quad\text{und}\quad \mu e^{-\mu t}$$

für $t \ge 0$ und null sonst; die gesuchte Wahrscheinlichkeit ist also

$$\begin{aligned} \iint\limits_{0\le s<t} \lambda e^{-\lambda s}\cdot \mu e^{-\mu t}\,ds\,dt &= \lambda\mu \int_0^\infty e^{-\lambda s}\left(\int_s^\infty e^{-\mu t}\,dt\right) ds = \lambda\mu \int_0^\infty e^{-\lambda s}\left(-\frac{1}{\mu}e^{-\mu t}\Big|_s^\infty\right) ds \\ &= \lambda \int_0^\infty e^{-\lambda s} e^{-\mu s}\,ds = \lambda \int_0^\infty e^{-(\lambda+\mu)s}\,ds = \frac{-\lambda}{\lambda+\mu} e^{-(\lambda+\mu)s}\Big|_0^\infty \\ &= \frac{\lambda}{\lambda+\mu} = \frac{\frac{1}{4000}}{\frac{1}{4000}+\frac{1}{5000}} = \frac{1}{1+4/5} = \frac{5}{9} = 0{,}\overline{5}\,. \end{aligned}$$

b) Berechnen Sie Erwartungswert und Varianz für die *Summe* der Brenndauern der beiden Lampen!

Lösung: Wenn wir die Möglichkeit ignorieren, daß beide Lampen gleichzeitig durch Blitzschlag oder eine sonstige Überspannung geschlachtet werden, können wir davon ausgehen, daß die Lebensdauern der beiden Lampen voneinander unabhängige Zufallsvariablen sind. Damit addieren sich nicht nur die Erwartungswerte, sondern nach dem Satz von BIENAYMÉ auch die Varianzen. Die Exponentialverteilung mit Parameter λ hat Erwartungswert $1/\lambda$ und Varianz

$$\int_0^\infty \left(t-\frac{1}{\lambda}\right)^2 \lambda e^{-\lambda t}\,dt = \int_0^\infty \lambda t^2 e^{-\lambda t}\,dt - \int_0^\infty 2te^{-\lambda t}\,dt + \int_0^\infty \frac{1}{\lambda}e^{-\lambda t}\,dt$$
$$= \frac{2}{\lambda^2} - \frac{2}{\lambda^2} + \frac{1}{\lambda^2} = \frac{1}{\lambda^2}\,.$$

Damit ist der Erwartungswert für die Summe der Brenndauern

$$4000 + 5000 = 9000$$

und die Varianz ist

$$4000^2 + 5000^2 = 16\,000\,000 + 25\,000\,000 = 41\,000\,000\,.$$

Aufgabe 4: *(5 Punkte)*

Die Wahrscheinlichkeitsverteilung einer zweidimensionalen Zufallsvariablen $X \otimes Y$ sei gegeben durch

$$\varphi(x,y) = \begin{cases} \frac{e^{-y-x/y}}{y} & \text{für } x,y>0 \\ 0 & \text{sonst} \end{cases}\,.$$

Was ist $\mathbb{E}(X|Y)$?

Lösung: Die Randverteilung bezüglich Y ist null für $y \le 0$, und für $y > 0$ ist

$$\varphi_Y(y) = \int_{-\infty}^\infty f(x,y)\,dx = \int_0^\infty \frac{e^{-y-x/y}}{y}\,dx = \frac{e^{-y}}{y}\int_0^\infty e^{-x/y}\,dx$$
$$= \frac{e^{-y}}{y}(-y)e^{-x/y}\Big|_0^\infty = e^{-y}\,.$$

Damit ist

$$\varphi(x|Y=y) = \frac{\varphi(x,y)}{\varphi_Y(y)} = \begin{cases} \frac{e^{-x/y}}{y} & \text{für } x,y>0 \\ 0 & \text{sonst} \end{cases}, \quad \text{also} \quad \varphi(x|Y) = \begin{cases} \frac{e^{-x/Y}}{Y} & \text{für } x,y>0 \\ 0 & \text{sonst} \end{cases},$$

und

$$\mathbb{E}(X|Y) = \int_{-\infty}^\infty x\cdot\varphi(x|Y)\,dx = \int_0^\infty x\cdot\frac{e^{-x/Y}}{Y}\,dx = -(x+Y)e^{-x/Y}\Big|_0^\infty = Y\,.$$

Aufgabe 5: *(12 Punkte)*
Beim Lotto werden aus einem Gefäß, das 49 nummerierte Kugeln enthält, nacheinander sechs Kugeln entnommen; die Zufallsvariable X_i stehe für die Zahl auf der i-ten Kugel.

a) Was ist $\mathbb{E}(X_2|X_1)$?

Lösung: Falls die erste gezogene Kugel die Nummer i hat, kann die zweite alle Nummern zwischen eins und 49 haben außer i, und jede dieser 48 Möglichkeiten ist gleichwahrscheinlich. Also ist

$$\mathbb{E}(X_2|X_1 = i) = \frac{1}{48}\sum_{\substack{j=1\\ j\neq i}} = \frac{1}{48}\left(\sum_{j=1}^{4} 9j - i\right) = \frac{1}{48}\left(\frac{49\cdot 50}{2} - i\right) = \frac{49\cdot 25 - i}{48} = \frac{1225 - i}{48}\,.$$

Somit ist

$$\mathbb{E}(X_2|X_1) = \frac{1225 - X_1}{48}\,.$$

b) Was ist $\mathbb{E}(X_1 + X_2)$?

Lösung: Nach *a)* ist

$$\mathbb{E}(X_1 + X_2|X_1 = i) = i + \frac{1225 - i}{48} = \frac{1225 + 47i}{48}\,.$$

Summation über die möglichen Werte von i ergibt

$$\begin{aligned}\mathbb{E}(X_1 + X_2) &= \sum_{i=1}^{49} P(X_1 = 1)\mathbb{E}(X_1 + X_2|X_1 = i) = \frac{1}{49}\sum_{i=1}^{49}\frac{1225 + 47i}{48}\\ &= \frac{1225}{48} + \frac{47}{48\cdot 49}\sum_{i=1}^{49} i = \frac{1225}{48} + \frac{47\cdot 49\cdot 25}{48\cdot 49} = \frac{1225 + 47\cdot 25}{48} = \frac{2400}{48} = 50\,.\end{aligned}$$

c) Verallgemeinern Sie das Ergebnis von *b)* zu einer plausiblen Formel für $\mathbb{E}(X_1 + \cdots + X_6)$ und $\mathrm{Var}(X_1 + \cdots + X_6)$ oder erraten Sie eine solche Formel, falls Sie *b)* nicht gelöst haben! Geben Sie lediglich eine *kurze* Begründung, warum Ihre Formel richtig sein sollte; ein formaler Beweis ist **nicht** gefragt.

Lösung: Da der Erwartungswert von X_1 gleich 25 ist, sollte man anhand der Rechnung aus *b)* erwarten, daß

$$\mathbb{E}(X_1 + \cdots + X_k) = 25k$$

ist für $k = 1, \ldots, 6$. Dies ist auch aus Symmetriegründen plausibel, denn zwar hängt die Verteilung eines jeden X_k ab von den bereits gezogenen Kugeln, aber es ist kein Grund zu sehen, warum dies im Mittel den Erwartungswert erhöhen oder erniedrigen sollte. Damit sollte $X_1 + \cdots + X_6)$ Erwartungswert $25\cdot 6 = 150$ haben.
Aus demselben Grund erwartet man auch, daß die Varianz der Summe der ersten X Ziehungen gleich dem k-fachen der Varianz von X_1 ist, also dem k-fachen von

$$\frac{1}{49}\sum_{i=1}^{49}(i - 25)^2 = \frac{1}{49}\sum_{j=-24}^{24} j^2 = \frac{24\cdot 25\cdot 49}{49\cdot 3} = 8\cdot 25 = 200\,.$$

Speziell für $k = 6$ sollte die Varianz also gleich 1200 sein.

d) Berechnen Sie aufgrund der Formeln aus *c)* Erwartungswert und Varianz der Summe aller in einem Jahr (mit 52 Ziehungen) gezogenen Lottozahlen (ohne Zusatzzahl)!

Lösung: Da die Ziehungen der einzelnen Wochen voneinander unabhängig sind, addieren sich nach dem Satz von BIENAYMÉ die Varianzen, und die Erwartungswerte tun dies natürlich erst recht. Somit ist der Erwartungswert für die Summe $52 \cdot 150 = 7\,800$ und die Varianz $52 \cdot 1200 = 62\,400$.

e) Was können Sie über die Verteilung dieser Summe sagen?

Lösung: Nach dem zentralen Grenzwertsatz sollte sie ziemlich nahe bei einer Normalverteilung mit Erwartungswert 7 800 und Varianz 62 400 liegen.

f) Geben Sie ein Intervall an, in dem diese Summe mit mindestens 95%-iger Wahrscheinlichkeit liegt!

Lösung: Wegen

$$\frac{1}{\sqrt{2\pi}} \int_{-2}^{2} e^{-t^2/2}\,dt \approx 0{,}9544997360$$

nimmt eine normalverteilte Zufallsvariable mit einer Wahrscheinlichkeit von etwas über 95% Werte an, die sich um höchstens zwei Standardabweichungen vom Erwartungswert unterscheiden. Hier ist die Standardabweichung

$$\sigma = \sqrt{62\,400} \approx 249{,}8 \quad \text{und} \quad 2\sigma \approx 499{,}6\,,$$

die Summe liegt also mit über 95%-iger Wahrscheinlichkeit zwischen 7 300,4 und 8 299,6.

Einige nützliche Formeln, die ohne Beweis verwendet werden können:

$$\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}, \qquad \sum_{i=-n}^{n} i^2 = \frac{n(n+1)(2n+1)}{3}$$

$$\int \lambda t e^{-\lambda t}\,dt = -e^{-\lambda t}\left(t + \frac{1}{\lambda}\right) + C, \qquad \int \lambda t^2 e^{-\lambda t}\,dt = -e^{-\lambda t}\left(t^2 + \frac{2t}{\lambda} + \frac{2}{\lambda^2}\right) + C$$

$$\frac{1}{\sqrt{2\pi}} \int_{-1}^{1} e^{-t^2/2}\,dt \approx 0{,}6826894920 \approx \frac{2}{3}, \qquad \frac{1}{\sqrt{2\pi}} \int_{-2}^{2} e^{-t^2/2}\,dt \approx 0{,}9544997360$$