

1. Übungsblatt

1. Beim TÜV werden n Fahrzeuge überprüft. Für $i = 1, \dots, n$ bezeichne A_i das Ereignis "das i -te Fahrzeug erhält die Prüfplakette". Beschreiben Sie die folgenden Ereignisse durch mengentheoretische Verknüpfungen der Ereignisse A_i :
 - (a) mindestens eines der n Fahrzeuge erhält keine Plakette;
 - (b) kein Fahrzeug erhält eine Plakette;
 - (c) genau ein Fahrzeug erhält keine Plakette;
 - (d) höchstens ein Fahrzeug erhält eine Plakette.
2. Es sei \mathcal{F} eine σ -Algebra über Ω . Zeige, dass für jedes Wahrscheinlichkeitsmaß P auf \mathcal{F} gilt:
 - (a) $P(\emptyset) = 0$;
 - (b) $A, B \in \mathcal{F}, A \subseteq B \Rightarrow P(A) \leq P(B)$;
 - (c) $\forall A, B \in \mathcal{F} : P(A \cup B) = P(A) + P(B) - P(A \cap B)$;
 - (d) $\forall A_n \in \mathcal{F}, n \geq 1 : P(\bigcup_{n \geq 1} A_n) \leq \sum_{n \geq 1} P(A_n)$ (*Subadditivität*).

Beweise ferner, dass jede normierte, additive Mengenfunktion $Q : \mathcal{F} \rightarrow [0, 1]$ (d.h. $Q(\Omega) = 1, Q(A \cup B) = Q(A) + Q(B)$ für alle disjunkten $A, B \in \mathcal{F}$), die σ -stetig ist, auch σ -additiv und damit ein Wahrscheinlichkeitsmaß ist.

3. Löse die folgenden Textaufgaben jeweils mit vollständiger Angabe und Begründung der wahrscheinlichkeitstheoretischen Modellierung:
 - (a) Wie viele Rosinen müssen in 500g Teig vorhanden sein, damit ein 50g-Brötchen mit mindestens 99% Wahrscheinlichkeit eine Rosine enthält?
 - (b) Ein gewisser Chevalier de Méré wunderte sich, dass er beim Werfen mit drei Würfeln die Augensumme 11 häufiger beobachtet hatte als die Augensumme 12, obwohl doch 11 durch die Kombinationen $6-4-1, 6-3-2, 5-5-1, 5-4-2, 5-3-3, 4-4-3$ und die Augensumme 12 durch ebensoviele (welche?) Kombinationen erzeugt würde. Kann diese Beobachtung als „vom Zufall bedingt“ angesehen werden oder ist die Argumentation falsch?

4. Für ganze Zahlen $N \geq 1$, $0 \leq W \leq N$, $0 \leq n \leq N$, $0 \leq w \leq W$ gebe $p_{N,W,n}(w)$ die Wahrscheinlichkeit an, dass bei n -fachem Ziehen (ohne Zurücklegen) aus einer Urne mit W weißen und $N - W$ schwarzen Kugeln genau w weiße Kugeln gezogen werden.

(a) Begründe mit kombinatorischen Argumenten die Formel

$$p_{N,W,n}(w) = \frac{\binom{N-W}{n-w} \binom{W}{w}}{\binom{N}{n}}.$$

(b) Weise anhand der Formel nach, dass $p_{N,W,n}$ eine Zähldichte auf $\Omega = \{0, 1, \dots, W\}$ ist (diese definiert die *hypergeometrische Verteilung*).

(c) Berechne mit Hilfe dieser Formel die Wahrscheinlichkeit für k Richtige im Lotto 6 aus 49 ($0 \leq k \leq 6$).

Abgabe in der Vorlesung am Donnerstag, dem 23.4.09



2. Übungsblatt

1. Zu einer Tanzstunde kommen n Paare. Um für Abwechslung zu sorgen, wird jeder Dame rein zufällig einer der Herren zugelost. Wie groß ist die Wahrscheinlichkeit, dass mindestens ein ursprüngliches Paar miteinander tanzen wird? Bestimmen Sie den Grenzwert dieser Wahrscheinlichkeit für $n \rightarrow \infty$.
Anleitung: Sei A_k das Ereignis „Dame k wird ursprünglicher Partner zugelost“. Beweise die *Einschluss-Ausschluss-Formel*:

$$\mathbb{P}\left(\bigcup_{k=1}^n A_k\right) = \sum_{l=1}^n \left((-1)^{l-1} \sum_{\{k_1, \dots, k_l\} \subseteq \{1, \dots, n\}} \mathbb{P}(A_{k_1} \cap \dots \cap A_{k_l})\right).$$

Bestimme die rechte Seite z.B. mittels der Ergebnisse für Urnenmodelle.

2. Es sei $\mathcal{Z} = \{Z_i \mid i \in I\}$ mit einer Indexmenge $I \subseteq \mathbb{N}$ eine abzählbare Zerlegung von Ω in disjunkte Teilmengen.
- (a) Gib die kleinste σ -Algebra $\sigma(\mathcal{Z})$ über Ω an, die das Mengensystem \mathcal{Z} umfasst.
- (b) Bestimme mittels (a) für $\Omega = [0, 1)$

$$\mathcal{F}_n := \sigma\left(\left\{[k2^{-n}, (k+1)2^{-n}) \mid k \in \{0, 1, \dots, 2^n - 1\}\right\}\right).$$

- (c) Zeige, dass $\mathcal{F} := \bigcup_{n \geq 1} \mathcal{F}_n$ eine Algebra über $[0, 1)$ bildet. Ist \mathcal{F} auch eine σ -Algebra?
3. Es sei $F : \mathbb{R} \rightarrow [0, 1]$ gegeben durch

$$F(x) = \sum_{n=1}^{\infty} 2^{-n} \mathbf{1}_{[1/n, \infty)}(x).$$

Zeige, dass es sich um die Verteilungsfunktion eines Wahrscheinlichkeitsmaßes \mathbb{P} auf $\mathfrak{B}_{\mathbb{R}}$ handelt, und berechne folgende Wahrscheinlichkeiten

$$\mathbb{P}([1, \infty)), \mathbb{P}([1/10, \infty)), \mathbb{P}(\{0\}), \mathbb{P}((-5, 1/2)), \mathbb{P}(\mathbb{Q}).$$

4. Es sei \mathbb{P} ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}^2, \mathfrak{B}_{\mathbb{R}^2})$. Zeige folgende Eigenschaften der zugehörigen Verteilungsfunktion $F(x, y) := \mathbb{P}((-\infty, x] \times (-\infty, y])$:

(a) Durch die Angabe von F ist \mathbb{P} eindeutig bestimmt.

(b) Es gilt für alle $a_1 \leq b_1, a_2 \leq b_2$ (Skizze!)

$$F(b_1, b_2) - F(a_1, b_2) - F(b_1, a_2) + F(a_1, a_2) = \mathbb{P}((a_1, b_1] \times (a_2, b_2]) \geq 0.$$

(c) Für $x^{(k)} \in \mathbb{R}^2$ mit $x_i^{(k)} \downarrow x_i, i = 1, 2$, für $k \rightarrow \infty$ folgt $F(x^{(k)}) \downarrow F(x)$.

(d) $\lim_{k \rightarrow \infty} F(k, k) = 1, \lim_{k \rightarrow -\infty} F(k, k) = 0$.

Bemerkung: Wie im eindimensionalen Fall kann man umgekehrt zeigen, dass jede Funktion F mit den Eigenschaften (b)-(d) ein zugehöriges Wahrscheinlichkeitsmaß \mathbb{P} auf $\mathfrak{B}_{\mathbb{R}^2}$ generiert.

Abgabe in der Vorlesung am Donnerstag, dem 30.4.09



3. Übungsblatt

1. Es sei φ_{μ,σ^2} die Dichte der Normalverteilung $\mathcal{N}(\mu, \sigma^2)$, also

$$\varphi_{\mu,\sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad x \in \mathbb{R}.$$

- (a) Bestimme die Extrema und Wendepunkte von φ_{μ,σ^2} und skizziere den Funktionsgraphen.
(b) Nun sei $\mu = 0$, $\sigma^2 = 1$. Zeige für alle $\eta > 0$ die Abschätzung

$$\frac{\eta}{\sqrt{2\pi(1+\eta^2)}} e^{-\eta^2/2} \leq \int_{\eta}^{\infty} \varphi_{0,1}(x) dx \leq \frac{1}{\sqrt{2\pi\eta^2}} e^{-\eta^2/2}.$$

Anleitung: für die erste Ungleichung zeige $\int_{\eta}^{\infty} (1+x^{-2})e^{-x^2/2} dx = \eta^{-1}e^{-\eta^2/2}$ und schließe $\eta^{-1}e^{-\eta^2/2} \leq (1+\eta^{-2}) \int_{\eta}^{\infty} e^{-x^2/2} dx$, für die zweite Ungleichung benutze $e^{-x^2/2} \leq \frac{x}{\eta} e^{-x^2/2}$ falls $x \geq \eta$.

- (c) Bestimme unter Verwendung von Teil (b) approximativ die Werte $\int_m^{\infty} \varphi_{0,1}(x) dx$ für $m = 0, 1, 2, 3, 4, 5$.
freiwillig: Welche Werte liefert eine numerische Integration (beliebige Mathematik-Software)?

2. In einem Kreis vom Radius r werde „rein zufällig“ eine Sehne ausgewählt. Mit welcher Wahrscheinlichkeit ist die Länge dieser Sehne größer als r ? Verwende folgende Zufallsbeschreibungen:

- (a) Die Sehne ist durch ihren Mittelpunkt eindeutig bestimmt. Die Lage des Mittelpunkts ist gleichmäßig in der Kreisscheibe verteilt.
(b) Die Sehne ist durch ihre Endpunkte eindeutig bestimmt und aus Symmetriegründen wählen wir den einen Endpunkt fest. Der andere möge gleichmäßig auf dem Kreisrand verteilt sein.
(c) Die Sehne ist durch ihren Abstand vom Kreismittelpunkt und die entsprechende Richtung eindeutig festgelegt. Aus Symmetriegründen kann die Richtung fest gewählt werden, der Abstand sei gleichmäßig auf $[0, r]$ verteilt.

3. Es bezeichne P_λ die Exponentialverteilung zum Parameter $\lambda > 0$.

- (a) Zeige, dass die Exponentialverteilung in folgendem Sinne *gedächtnislos* ist:

$$\forall t, x > 0 : P_\lambda((x+t, \infty) | (t, \infty)) = P_\lambda((x, \infty)).$$

Erkläre diese Eigenschaft am Beispiel einer zufälligen Wartezeit.

- (b) Beweise umgekehrt, dass jede solche gedächtnislose Verteilung auf $(\mathbb{R}^+, \mathfrak{B}_{\mathbb{R}^+})$ eine Exponentialverteilung ist.
- (c) Bestimme die Verteilungen P auf $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$, die im folgenden diskreten Sinne gedächtnislos sind:

$$\forall m, n \in \mathbb{N} : P(\{k \in \mathbb{N} | k \geq m+n\} | \{k \in \mathbb{N} | k \geq m\}) = P(\{k \in \mathbb{N} | k \geq n\}).$$

4. Die amerikanische Journalistin Marilyn vos Savant (mit angeblich dem höchsten IQ der Welt) bekam 1990 für ihre Denksportkolumne im „Parade Magazine“ von einem Leser folgende Aufgabe:

Suppose you're on a game show, and you're given the choice of three doors. Behind one door is a car, behind the others, goats. You pick a door, say number one, and the host, who knows what's behind the doors, opens another door, say number 3, which has a goat. He says to you "Do you want to pick door number 2?" Is it to your advantage to switch your choice of doors?

Ihre Antwort lautete: *"Yes, you should switch. The first door has a 1/3 chance of winning, but the second door has a 2/3 chance."* Hat sie recht?

(Bitte mathematisch exakt begründen!)

5. (*Zusatzaufgabe*) Betrachte den Ergebnisraum $\Omega = \{0, 1\}^{\mathbb{N}}$ des unendlich oft wiederholten Münzwurfs. Es sei $\Pi_n : \Omega \rightarrow \{0, 1\}^n$ die durch $\Pi_n(\omega) = (\omega_1, \dots, \omega_n)$ gegebene Koordinatenprojektion. Zeige, dass das System der *Zylindermengen*

$$\mathfrak{A} := \{\Pi_n^{-1}(A_n) | n \in \mathbb{N}, A_n \subseteq \{0, 1\}^n\}$$

eine Algebra über Ω bildet. Setze $P(\Pi_n^{-1}(A_n)) := |A_n|/2^n$ und zeige, dass P ein Prämaß auf \mathfrak{A} definiert. Konstruiere damit einen Wahrscheinlichkeitsraum, der den unendlich oft wiederholten Münzwurf modelliert.



4. Übungsblatt

1. Beweise: Sind $(A_i)_{i \in I}$ unabhängige Ereignisse, so ist auch die Familie der erzeugten σ -Algebren $\mathcal{F}_i := \{\emptyset, \Omega, A_i, A_i^c\}$, $i \in I$, unabhängig.
2. Es sei X ein Zufallsvektor auf (Ω, \mathcal{F}, P) mit Dichte $f_X : \mathbb{R}^d \rightarrow [0, \infty)$.

- (a) Beweise im Fall $d = 1$, dass $Y = g(X)$ für eine streng monotone Funktion $g : I \rightarrow J$ mit offenen Intervallen $I, J \subseteq \mathbb{R}$, deren Inverse g^{-1} stetig differenzierbar ist, eine Dichte besitzt, nämlich

$$f_Y(y) = f_X(g^{-1}(y)) |(g^{-1})'(y)| \mathbf{1}_{g(X(\Omega))}(y), \quad y \in \mathbb{R}.$$

Tipp: Betrachte die Verteilungsfunktion von Y .

- (b) Zeige im Fall $d = 1$, dass auch X^2 eine Dichte besitzt und bestimme diese explizit für $N(0, 1)$ -verteiltes X .
 - (c) Verallgemeinere Teil (a) auf den allgemeinen d -dimensionalen Fall.
3. Es sei $(X_i)_{i \in I}$ eine Familie unabhängiger (reellwertiger) Zufallsvariablen. Zeige:
 - (a) Sind $g_i : \mathbb{R} \rightarrow \mathbb{R}$, $i \in I$, beliebige Borel-messbare Funktionen, so ist auch $(g_i(X_i))_{i \in I}$ eine Familie unabhängiger Zufallsvariablen.
 - (b) Im Fall $I = \{1, 2, 3\}$ und für Borel-messbare Funktionen $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $g : \mathbb{R} \rightarrow \mathbb{R}$ sind die Zufallsvariablen $Y = f(X_1, X_2)$ und $Z = g(X_3)$ unabhängig.
 4. Simulation von Zufallsvariablen:

- (a) Es sei U eine $U([0, 1])$ -verteilte Zufallsvariable und F eine Verteilungsfunktion (auf \mathbb{R}). Weise nach, dass durch

$$X := \inf\{x \in \mathbb{R} \mid F(x) \geq U\}$$

eine Zufallsvariable konstruiert wird, deren Verteilungsfunktion F ist.

- (b) Es seien U, V unabhängige $U([0, 1])$ -verteilte Zufallsvariablen. Setze $R = \sqrt{-2 \log(U)}$, $X = R \cos(2\pi V)$ und $Y = R \sin(2\pi V)$. Beweise, dass X und Y unabhängige, standard-normalverteilte Zufallsvariablen sind.

Tipp: Berechne die Dichte von R und betrachte dann die Polarkoordinatentransformation $(R, V) \mapsto (X, Y)$, verwende Aufgaben 2 und 3.

- (c) *freiwillig:* Simuliere 100.000 unabhängige $N(0, 1)$ -verteilte Zufallsvariablen gemäß Methode (a) und (b). Gib jeweils die Rechenzeit an und stelle die Werte in einem Histogramm dar (mit Programmcode).



5. Übungsblatt

- Es seien X_1, \dots, X_n unabhängige Zufallsvariablen mit Verteilungsfunktionen F_1, \dots, F_n .
 - Zeige, dass die Verteilungsfunktionen von $M = \max(X_1, \dots, X_n)$ und $m = \min(X_1, \dots, X_n)$ gegeben sind durch $F_M(x) = \prod_{i=1}^n F_i(x)$ bzw. $F_m(x) = 1 - \prod_{i=1}^n (1 - F_i(x))$.
 - Bestimme für den Fall, dass jedes X_i gleichmäßig auf $[0, 1]$ verteilt ist, die Dichte von M und m . Sind M und m unabhängig?
 - Die Zeit bis zum Zerfall eines radioaktiven Atoms wird durch die Exponentialverteilung $\text{Exp}(\lambda)$ beschrieben. Wie ist die Zeit bis zum ersten Atomzerfall in einer Probe aus N solchen Atomen verteilt?
- Es sei $X = (X_1, \dots, X_r)$ eine Zufallsvariable auf (Ω, \mathcal{F}, P) mit Werten in $(S^r, \mathcal{P}(S^r))$, wobei S eine abzählbare Menge ist.
 - Beweise formal, dass die i -te Randverteilung P^{X_i} von X durch folgende Zähldichte beschrieben wird:

$$p^{X_i}(x_i) = \sum_{x_j \in S; j \in \{1, \dots, r\} \setminus \{i\}} P(X_1 = x_1, \dots, X_r = x_r), \quad x_i \in S.$$

- Betrachte ein Spiel mit drei Ausgängen (Gewinn, Verlust, Remis), das n -mal wiederholt wird, und bezeichne mit $X = (X_1, X_2, X_3)$ die zufällige Anzahl der jeweiligen Spielausgänge. Erkläre die Modellierung von X als *Multinomial*-verteilte Zufallsvariable mit Parametern $n \geq 1$ und $p_1, p_2, p_3 \in [0, 1]$ mit $p_1 + p_2 + p_3 = 1$, d.h. für $k_1, k_2, k_3 \in \mathbb{N}_0$

$$P(X_1 = k_1, X_2 = k_2, X_3 = k_3) = \frac{n!}{k_1! k_2! k_3!} p_1^{k_1} p_2^{k_2} p_3^{k_3} \mathbf{1}_{\{n\}}(k_1 + k_2 + k_3).$$

Weise nach, dass für $i = 1, 2, 3$ die i -te Randverteilung P^{X_i} gerade die Binomialverteilung $\text{Bin}(n, p_i)$ ist. Sind X_1, X_2 und X_3 unabhängig?

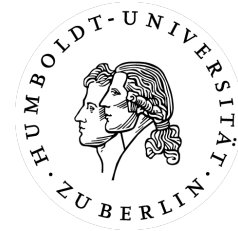
3. Ein System besteht aus n gleichartigen, voneinander unabhängigen Komponenten K_1, \dots, K_n . Es funktioniert nur, solange alle Komponenten funktionieren. Für jede einzelne Komponente sei die Zeit bis zum Ausfall geometrisch verteilt mit demselben Parameter $p \in (0, 1)$.
- (a) Bestimme für jedes feste $i = 1, \dots, n$ und $m \in \mathbb{N}_0$ die Wahrscheinlichkeit des Ereignisses $A_{i,m}$, dass K_i zur Zeit m noch funktionsfähig ist.
 - (b) Konstruiere einen diskreten Wahrscheinlichkeitsraum und formalisiere die Ereignisse $A_{i,m}$, so dass $A_{1,m}, \dots, A_{n,m}$ stochastisch unabhängig sind.
 - (c) Leite formal ab, dass die Zeit bis zum Ausfall des Gesamtsystems geometrisch verteilt ist mit Parameter $1 - (1 - p)^n$.
4. Es seien $X_k, k \geq 1$, unabhängige $[0, \infty)$ -wertige Zufallsvariablen. Entscheide (mit formaler Begründung), welche der folgenden Ereignisse in der zu $(X_k)_{k \geq 1}$ gehörigen asymptotischen σ -Algebra liegen: $A_1 := \{\sum_{k \geq 1} X_k < \infty\}$, $A_2 := \{\sum_{k \geq 1} X_k < 1\}$, $A_3 := \{\inf_{k \geq 1} X_k < 1\}$, $A_4 := \{\liminf_{k \rightarrow \infty} X_k < 1\}$.

Abgabe in der Vorlesung am Dienstag(!), dem 26.5.09



6. Übungsblatt

- Es seien X eine $\text{Poiss}(\lambda)$ -verteilte und Y eine $\text{Geo}(p)$ -verteilte Zufallsvariable mit Parameter $\lambda > 0$ bzw. $p \in (0, 1)$.
 - Bestimme $\mathbb{E}[X]$, $\mathbb{E}[X^2]$ sowie $\mathbb{E}[Y]$.
 - Zeige, dass X und Y in \mathcal{L}^q liegen für jedes $q \geq 1$.
- Bestimme für eine Zufallsvariable X Erwartungswert $\mathbb{E}[X]$ und Varianz $\text{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$ (sofern diese existieren) im Fall folgender Verteilungen:
 - $N(\mu, \sigma^2)$; (b) $\text{Exp}(\lambda)$; (c) Cauchyverteilung (d.h. $f^X(x) = \frac{1}{\pi} \frac{1}{1+x^2}$);
 - $\chi^2(1)$ -Verteilung (d.h. $f^X(x) = (2\pi x)^{-1/2} e^{-x/2} \mathbf{1}_{\mathbb{R}^+}(x)$).
- Es sei X eine Zufallsvariable in \mathcal{L}^2 . Zeige, dass die Funktion $v : \mathbb{R} \rightarrow [0, \infty)$ mit $v(x) := \mathbb{E}[(X - x)^2]$ ihr Minimum bei $x = \mathbb{E}[X]$ annimmt. Charakterisiere ferner die Minimalstellen $m \in \mathbb{R}$ von $d : \mathbb{R} \rightarrow [0, \infty)$ mit $d(x) := \mathbb{E}[|X - x|]$ im Fall, dass $X \in \mathcal{L}^1$ eine Dichte f^X besitzt.
Freiwillig: charakterisiere m für allgemeine Zufallsvariablen $X \in \mathcal{L}^1$.
- Bei der Fußballweltmeisterschaft treten 32 Mannschaften mit einem Kader von jeweils 20 Spielern gegeneinander an. Es gibt von jedem Spieler ein Sammelbildchen. Am Kiosk wird ein sichtgeschützt verpacktes Bildchen für einen Cent verkauft. Wieviel wird ein Sammler im Mittel am Kiosk ausgeben, bis er von jedem Spieler (mindestens) ein Bildchen besitzt?
Tipp: Bezeichne mit N_i die Anzahl der erworbenen Bildchen, bis man von i Spielern ein Bildchen besitzt, und bestimme die Verteilung von $D_i = N_i - N_{i-1}$.
Freiwillige Zusatzaufgabe: Was ergibt sich, falls sich zwei Sammler zusammenschließen und Bildchen tauschen?



7. Übungsblatt

1. Zu Daten $(x_i, y_i)_{1 \leq i \leq n}$ in \mathbb{R}^2 wird die *Regressionsgerade* $y = \hat{a}x + \hat{b}$ definiert mittels der *Methode der kleinsten Quadrate*:

$$(\hat{a}, \hat{b}) := \operatorname{argmin}_{(a,b) \in \mathbb{R}^2} \left\{ \sum_{i=1}^n (y_i - (ax_i + b))^2 \right\}.$$

- (a) Bestimme \hat{a} und \hat{b} als Funktion der empirischen Mittel $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$, $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$, der empirischen Varianzen $\bar{\sigma}_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$, $\bar{\sigma}_y^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2$ und der empirischen Korrelation $\bar{\rho}_{xy} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) / (\bar{\sigma}_x \bar{\sigma}_y)$ (falls $\bar{\sigma}_x, \bar{\sigma}_y > 0$).
- (b) Um die Abhängigkeit der durch Melanome (Hautkrebs) verursachten Todesfälle von der Sonneneinstrahlung zu bestimmen, wurde in den Bundesstaaten der USA die Mortalität (Todesfälle pro 10^7 Einwohner) und der Breitengrad erfasst. Bestimme aus den Daten

Staat	Delaware	Iowa	Michigan	New Hampshire	Oklahoma	Texas	Wyoming
Mort.	200	128	117	129	182	229	134
Breite	39	42	44	44	35	31	43

die zugehörige Regressionsgerade und zeichne diese zusammen mit den Daten in ein Koordinatensystem (Computereinsatz gestattet). Welche Mortalität ist in Ohio (Breitengrad 40) in etwa zu erwarten?

2. Es sei $X \in \mathcal{L}^4$ eine Zufallsvariable mit symmetrischer Verteilung (d.h. $P(X \geq x) = P(X \leq -x)$ für alle $x \in \mathbb{R}$). Setze $Y = X^2$. Weise nach, dass X und Y unkorreliert, aber im Allgemeinen nicht unabhängig sind. Bestimme die beste lineare Vorhersage von Y durch X (bzgl. mittlerer quadratischer Abweichung). Welche ist die beste nichtlineare Vorhersage?

3. Es sei P ein Wahrscheinlichkeitsmaß auf $(\mathbb{N}_0, \mathcal{P}(\mathbb{N}_0))$. Dann heißt

$$\varphi_P(s) := \sum_{k=0}^{\infty} P(\{k\})s^k, \quad s \in [-1, 1]$$

erzeugende Funktion von P . Als Potenzreihe ist φ_P wohldefiniert und in $(-1, 1)$ beliebig oft differenzierbar. Zeige:

- (a) P ist durch φ_P eindeutig bestimmt.
- (b) Der Erwartungswert von P ist gleich $\varphi'_P(1) := \lim_{s \uparrow 1} \varphi'_P(s) \in [0, +\infty]$ und daher endlich für $\varphi'_P(1) < \infty$. Formuliere ein analoges Resultat für die Varianz.
- (c) Für die Faltung gilt $\varphi_{P_1 * P_2}(s) = \varphi_{P_1}(s)\varphi_{P_2}(s)$, $s \in [-1, 1]$.

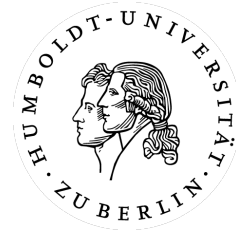
Bestimme mittels erzeugender Funktion den Erwartungswert und die Varianz der geometrischen Verteilung $\text{Geo}(p)$ sowie für $m \in \mathbb{N}$ die Zähldichte von $\text{Geo}(p)^{*m}$ (m -fache Faltung; Interpretation: Wartezeit bis zum m -ten Erfolg).

4. (Galton-Watson-Prozess) In einem Land wird der Name in weiblicher Linie vererbt. In Generation n bezeichne Z_n die Anzahl der Frauen mit einem bestimmten Namen, mit $X_{n,i}$, $i = 1, \dots, Z_n$, sei die Anzahl der weiblichen Nachkommen von Frau i in Generation n bezeichnet. Dann gilt also

$$Z_{n+1} = \sum_{i=1}^{Z_n} X_{n,i}, \quad n \geq 0,$$

mit $\sum_{i=1}^0(\dots) := 0$. Wir nehmen an: $Z_0 = 1$ und die $X_{n,i}$ sind unabhängige \mathbb{N}_0 -wertige Zufallsvariablen mit derselben Verteilung P . Zeige:

- (a) Für die erzeugende Funktion φ_n der Verteilung von Z_n gilt die Rekursionsformel $\varphi_{n+1}(s) = \varphi_n(\varphi_P(s))$, $n \geq 0$.
- (b) Die Wahrscheinlichkeit, dass der Name nach der n -ten Generation ausgestorben ist, berechnet sich durch $P(Z_n = 0) = \varphi_P^{\circ n}(0)$ mit der n -fachen Hintereinanderausführung $\varphi_P^{\circ n}$ von φ_P . Die Wahrscheinlichkeit, dass der Name jemals ausstirbt, ist dann gegeben durch $\lim_{n \rightarrow \infty} \varphi_P^{\circ n}(0)$.
- (c) Betrachtet man φ_P auf $[0, 1]$, so ist φ_P monoton wachsend und konvex mit $\varphi_P(1) = 1$. Der Wert $\lim_{n \rightarrow \infty} \varphi_P^{\circ n}(0)$ ist der kleinste Fixpunkt von $\varphi_P : [0, 1] \rightarrow [0, 1]$.
- (d) Im Fall $\mathbb{E}[X_{n,i}] \leq 1$ und $P(X_{n,i} = 1) < 1$ stirbt der Name mit Wahrscheinlichkeit eins irgendwann aus, im Fall $\mathbb{E}[X_{n,i}] > 1$ stirbt er mit positiver Wahrscheinlichkeit nie aus.



8. Übungsblatt

- Es seien U_1, U_2, U_3, U_4 unabhängige gleichmäßig auf $[-1, 1]$ verteilte Zufallsvariablen.
 - Bestimme die Dichten von $U_1 + U_2$, $U_1 + U_2 + U_3$ und $U_1 + U_2 + U_3 + U_4$.
 - Zeichne die Dichte von $S_n^* := \frac{\sqrt{3}}{\sqrt{n}}(U_1 + \dots + U_n)$ für $n = 1, 2, 3, 4$ sowie die Dichte der Standardnormalverteilung in ein Koordinatensystem (Computereinsatz gestattet).
- Es seien X_1, \dots, X_n unabhängige $N(\mu, \sigma^2)$ -verteilte Zufallsvariablen mit $\mu \in \mathbb{R}$ und $\sigma > 0$. Zeige für $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$, $\bar{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$:
 - Die Zufallsvariablen \bar{X} und $\bar{\sigma}^2$ sind unabhängig.
Tipp: Beweise zunächst, dass \bar{X} und $(X_k - \bar{X})_{1 \leq k \leq n}$ gemeinsam normalverteilt und unabhängig sind.
 - \bar{X} ist $N(\mu, \sigma^2/n)$ -verteilt und $\frac{n}{\sigma^2} \bar{\sigma}^2$ ist $\chi^2(n-1)$ -verteilt (zur Erinnerung $\chi^2(p) = \chi^2(1)^{*p}$). Was ist der Erwartungswert von $\bar{\sigma}^2$?
- Auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) sei ein zweidimensionaler Zufallsvektor (X, Y) mit stetiger Dichte $f^{(X,Y)}$ gegeben. Es gelte $f^Y(y) = \int_{-\infty}^{\infty} f^{(X,Y)}(\xi, y) d\xi > 0$ für alle $y \in \mathbb{R}$.
 - Bestimme die bedingte Wahrscheinlichkeit $P(X \leq x | Y \in [y-h, y+h])$ für $x, y \in \mathbb{R}$, $h > 0$. Zeige, dass der Grenzwert der bedingten Wahrscheinlichkeit für $h \rightarrow 0$ existiert und gleich

$$\frac{\int_{-\infty}^x f^{(X,Y)}(\xi, y) d\xi}{f^Y(y)} =: F^{X|Y=y}(x)$$

ist. Man definiert dann die *bedingte Dichte* von X gegeben $Y = y$ als

$$f^{X|Y=y}(x) := \frac{\partial}{\partial x} F^{X|Y=y}(x) = \frac{f^{(X,Y)}(x, y)}{f^Y(y)}.$$

- Bestimme $f^{X|Y=y}$, falls (X, Y) $N(\mu, \Sigma)$ -verteilt ist mit $\mu \in \mathbb{R}^2$ und $\Sigma \in \mathbb{R}^{2 \times 2}$ strikt positiv-definit. Untersuche das Verhalten von $f^{X|Y=y}$, falls die Korrelation $\rho(X, Y)$ gegen 0, +1 bzw. -1 konvergiert. Interpretation?

4. In einem Labor wurde eine Probe gefunden, die entweder aus einem schwach oder einem stark radioaktiven Material stammt. Um dies herauszufinden, sollen n -mal die Zeiten T_1, \dots, T_n bis zum nächsten Atomzerfall gemessen werden.

(a) Begründe, warum T_1, \dots, T_n näherungsweise als exponentialverteilt und unabhängig modelliert werden können. Gib ein entsprechendes statistisches Modell an, um für den Intensitätsparameter λ der Exponentialverteilung die Hypothese $H_0 : \lambda = \lambda_0$ gegen $H_1 : \lambda = \lambda_1$ für $\lambda_0 > \lambda_1$ zu testen.

(b) Begründe, weshalb ein gleichmäßig bester Test von der Gestalt

$$\varphi(t_1, \dots, t_n) = \mathbf{1}_{[c, \infty)}(t_1 + \dots + t_n)$$

gewählt werden kann mit einem geeigneten *kritischen Wert* $c \geq 0$.

(c) Bestimme approximativ den kritischen Wert in einem solchen Test für $\lambda_0 = 10$, $\lambda_1 = 2$ und $n = 5$ zum Niveau $\alpha = 0,01$. Wie groß ist die Wahrscheinlichkeit für einen Fehler zweiter Art?

Erinnerung: $\text{Exp}(\lambda) = \Gamma(\lambda, 1)$, $\Gamma(\lambda, p_1) * \Gamma(\lambda, p_2) = \Gamma(\lambda, p_1 + p_2)$

Abgabe in der Vorlesung am Donnerstag, dem 18.6.09



9. Übungsblatt

1. Betrachte für $p_0 \in (0, 1)$ und für die $\text{Bin}(n, p)$ -Verteilungen P_p mit Parametern $p \in [0, 1]$ das Testproblem $H_0 : p \leq p_0$ gegen $H_1 : p > p_0$ zum Niveau α .
 - (a) Zeige, dass der Likelihoodquotient $R_{p', p}(x)$ von $P_{p'}$ bezüglich P_p für $p < p'$ monoton in x wächst.
 - (b) Gib die Form des Neyman-Pearson-Tests vom Niveau α für $H_0 : p = p_0$ gegen $H_1 : p = p_1$ für beliebiges $p_1 > p_0$ an. Begründe mittels (a), dass dieser Test gleichmäßig bester Test zum Niveau α für die einfache Hypothese $H_0 : p = p_0$ gegen die zusammengesetzte Alternative $H_1 : p > p_0$ ist.
 - (c) Dieser Test besitzt Niveau α sogar auf der zusammengesetzten Hypothese $H_0 : p \leq p_0$. Schließe, dass dieser Test sogar gleichmäßig bester Test von $H_0 : p \leq p_0$ gegen $H_1 : p > p_0$ zum Niveau α ist.
2. Verwende Aufgabe 1 zur Lösung folgenden Problems: Zur Behandlung einer Krankheit gibt es ein bewährtes Medikament, das in 80% der Fälle zu einer Heilung führt. Der Hersteller eines neuen Medikaments verspricht größere Heilungschancen, da nach Verabreichung an 50 erkrankten Personen 44 geheilt wurden. Ist diese Aussage gerechtfertigt durch einen gleichmäßig besten Test zum Irrtumsniveau 5% der Hypothese, dass das neue Medikament maximal gleiche Heilungschancen bietet, gegen die Alternative, dass es bessere bietet? Skizziere die Gütefunktion dieses Tests durch Interpolation einiger Werte.

3. Es sei $(\mathcal{X}, \mathcal{F}, (P_0, P_1))$ ein *binäres* statistisches Modell für das Testen von $H_0 : \vartheta = 0$ gegen $H_1 : \vartheta = 1$. Es soll ein Test φ konstruiert werden, der die Summe S_φ der Fehlerwahrscheinlichkeiten erster und zweiter Art minimiert. Zeige:

- (a) Die Summe S_φ ist gegeben durch $1 + \mathbb{E}_0[(\varphi - \varphi R)\mathbf{1}_{[0, \infty)}(R)] - \mathbb{E}_1[\varphi \mathbf{1}_{\{+\infty\}}(R)]$, wobei R den Likelihoodquotienten von P_1 bezüglich P_0 bezeichnet (der existiere).
- (b) Ein Test der Form $\varphi(x) = \mathbf{1}_{(1, \infty]}(R(x))$ minimiert diese Summe S_φ unter allen Tests, und es gilt

$$\begin{aligned} \min_{\varphi} S_{\varphi} &= 1 - (P_1(R \in (1, \infty]) - P_0(R \in (1, \infty])) \\ &= 1 - \sup_{A \in \mathcal{F}} |P_0(A) - P_1(A)| \\ &=: 1 - \|P_0 - P_1\|_{TV}, \end{aligned}$$

wobei $\|P_0 - P_1\|_{TV}$ als *Totalvariationsabstand* zwischen P_0 und P_1 bezeichnet wird.

- (c) Weise nach, dass der Totalvariationsabstand stets im Intervall $[0, 1]$ liegt. Beschreibe die Fälle, wo der Totalvarianzabstand gleich null bzw. gleich eins ist, und interpretiere sie im Rahmen des Testproblems.

4. Ein Algorithmus zur Erzeugung von Pseudozufallsziffern soll getestet werden. Bei der Erzeugung von 10.000 Ziffern ergaben sich folgende Häufigkeiten:

Ziffer	0	1	2	3	4	5	6	7	8	9
Häufigkeit	1007	987	928	986	1010	1029	987	1006	1034	1026

Führe zum Niveau $\alpha = 0,1$ einen χ^2 -Anpassungstest auf Gleichverteilung durch.



10. Übungsblatt

1. Es sei X eine $N(\mu, \sigma^2)$ -verteilte Zufallsvariable. Benutze die Tschebyschew-Ungleichung, um $P(|X - \mu| \geq k\sigma)$ für $k = 1, 2, 3$ abzuschätzen. Vergleiche mit den genaueren Werten, die sich ergeben aus

$$\Phi(1) \approx 0,8413, \quad \Phi(2) \approx 0,9772, \quad \Phi(3) \approx 0,9987 \quad \text{mit} \quad \Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx.$$

Gibt es Zufallsvariablen, für die Gleichheit in der Tschebyschew-Ungleichung gilt?

2. Es seien $(X_n)_{n \geq 1}$ Zufallsvariablen, die P -fast sicher gegen eine Zufallsvariable X konvergieren. Folgere schrittweise:

- (a) $P(\exists \varepsilon > 0 \forall n \geq 1 : \sup_{m \geq n} |X_m - X| > \varepsilon) = 0.$
- (b) $\forall \varepsilon > 0 : P(\bigcap_{n=1}^{\infty} \{\sup_{m \geq n} |X_m - X| > \varepsilon\}) = 0.$
- (c) $\forall \varepsilon > 0 : \lim_{n \rightarrow \infty} P(\sup_{m \geq n} |X_m - X| > \varepsilon) = 0.$
- (d) X_n konvergiert gegen X P -stochastisch.

3. Ein Spieler startet mit dem Anfangskapital $K_0 = 1$. Bei jeder Runde $i = 1, \dots, n$ setzt er sein gesamtes Kapital ein, es wird eine faire Münze geworfen, und bei 'Kopf' erhält er den anderthalbfachen Einsatz zurück, bei 'Zahl' nur den halben.

- (a) Stelle das Kapital nach der n -ten Runde als $K_n = \prod_{i=1}^n R_i$ mit geeigneten unabhängigen Zufallsvariablen R_i dar.
- (b) Weise nach, dass das Spiel fair ist in dem Sinne, dass $\mathbb{E}[K_n] = 1$ gilt.
- (c) Zeige, dass trotzdem $\lim_{n \rightarrow \infty} K_n = 0$ (fast sicher) gilt.
Tipp: Wende das starke Gesetz der großen Zahlen auf $\log(K_n)$ an.

4. Es seien $(X_n)_{n \geq 1}$ und X reellwertige Zufallsvariablen auf (Ω, \mathcal{F}, P) .

(a) Beweise, dass aus $X_n \xrightarrow{P} X$ die Existenz einer Teilfolge $(X_{n(k)})_{k \geq 1}$ folgt mit $X_{n(k)} \rightarrow X$ P -fast sicher für $k \rightarrow \infty$.

Tipp: Benutze ein Borel-Cantelli-Argument

(b) Für den Fall, dass X_n nicht P -stochastisch gegen X konvergiert, zeige, dass es $\varepsilon > 0$ und eine Teilfolge $(X_{n(k)})_{k \geq 1}$ gibt mit

$$\forall k \geq 1 : P(|X_{n(k)} - X| > \varepsilon) \geq \varepsilon.$$

Schließe weiter, dass diese Teilfolge keine Teilteilfolge $(X_{n(k(l))})_{l \geq 1}$ besitzt, die P -f.s. gegen X konvergiert.

(c) Folgere aus (a) und (b) die Äquivalenz: Die Folge (X_n) konvergiert P -stochastisch gegen X genau dann, wenn jede Teilfolge $(X_{n(k)})$ eine Teilteilfolge $(X_{n(k(l))})$ besitzt mit $X_{n(k(l))} \rightarrow X$ P -f.s. für $l \rightarrow \infty$.

Freiwillig: Folgere, dass fast sichere Konvergenz nicht metrisierbar ist auf Wahrscheinlichkeitsräumen, wo stochastische und fast sichere Konvergenz nicht identisch sind. Gib Wahrscheinlichkeitsräume an, wo fast sichere und stochastische Konvergenz identisch sind.

Abgabe in der Vorlesung am Donnerstag, dem 2.7.09

Klausur: am Donnerstag 16.7.09, 13-15 Uhr, Räume RUD 26, 0'310 & 0'311

Stochastik I

Probeklausur

1. Entscheide, ob die folgenden Aussagen wahr oder falsch sind: (5P)
- (a) Jedes Prämaß auf einer Algebra \mathcal{A} kann auf eindeutige Weise zu einem Maß auf der von \mathcal{A} erzeugten σ -Algebra fortgesetzt werden.
 - (b) Jede Wahrscheinlichkeitsverteilung P auf $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})$ besitzt eine Dichte.
 - (c) Ereignisse A_1, \dots, A_n heißen stochastisch unabhängig unter dem Maß P , falls $P[A_i \cap A_j] = P[A_i]P[A_j]$ für alle $i, j \in \{1, \dots, n\}$, $i \neq j$, gilt.
 - (d) Zwei Zufallsvariablen X, Y mit Dichten f^X, f^Y sind unabhängig genau dann, wenn die gemeinsame Dichte $f^{X,Y}$ existiert und die Form $f^{X,Y}(x, y) = f^X(x)f^Y(y)$, $x, y \in \mathbb{R}$, besitzt.
 - (e) Besitzen zwei Zufallsvariablen X, Y Dichten, so besitzt auch $X + Y$ eine Dichte.
 - (f) Die Summe von zwei unabhängigen exponentialverteilten Zufallsvariablen ist wiederum exponentialverteilt.
 - (g) Eine Folge von Zufallsvariablen $(X_n)_{n \in \mathbb{N}}$ konvergiert gegen die Zufallsvariable X in Verteilung genau dann, wenn für die zugehörigen Verteilungsfunktionen $\lim_{n \uparrow \infty} F^{X_n}(y) = F^X(y)$ für alle $y \in \mathbb{R}$ gilt.
 - (h) Ist $\varphi : \mathcal{X} \rightarrow [0, 1]$ ein Test, so führt $\varphi(x) = 1$ zur Annahme der Hypothese H_0 .
 - (i) Sind die Zufallsvariablen $(X_n)_{n \in \mathbb{N}}$ unabhängig und $N(\mu, \sigma^2)$ -verteilt, so gilt $\lim_{n \uparrow \infty} \frac{1}{n} \sum_{i=1}^n X_i^2 = \sigma^2 + \mu^2$ fast sicher.
 - (j) Sind die Zufallsvariablen $(X_n)_{n \in \mathbb{N}}$ unabhängig und $N(0, 1)$ -verteilt, so konvergiert wegen $\mathbb{E}[X_1^4] = 3$ die Folge $\frac{1}{\sqrt{n}} \sum_{i=1}^n (X_i^2 - 1)$ in Verteilung gegen die $N(0, 1)$ -Verteilung.
2. (a) Formuliere und beweise beide Teile des Lemmas von Borel-Cantelli. (3P)
- (b) Es sei $(X_n)_{n \in \mathbb{N}}$ eine Folge von unabhängigen und identisch zum Parameter 1 exponentialverteilten Zufallsvariablen. Zeige: $P[\overline{\lim}_{n \uparrow \infty} \frac{X_n}{\ln n} = 1] = 1$. (2P)
- Hinweis:* Betrachte für $\epsilon \in (-1, 1)$ die Ereignisse $A_n := \{\frac{X_n}{\ln n} \geq 1 + \epsilon\}$, $n \in \mathbb{N}$.

3. Eine faire Münze wird solange in unabhängiger Folge geworfen, bis zum ersten Mal sowohl Wappen als auch Zahl erschienen sind.

(a) Zeige, dass die Wahrscheinlichkeitsverteilung der zufälligen Anzahl T der Würfe durch die Zähldichte $p_k = (\frac{1}{2})^{k-1}$, $k \geq 2$, gegeben ist. (2P)

(b) Berechne die erzeugende Funktion dieser Verteilung und mit deren Hilfe den Erwartungswert und die Varianz von T . (3P)

4. Es seien X und Y zwei Zufallsvariablen mit der gemeinsamen Dichte

$$f^{X,Y}(x,y) = \begin{cases} 2e^{-x-y} & \text{falls } 0 < x < y, \\ 0 & \text{sonst.} \end{cases}$$

(a) Bestimme die Randverteilungsdichten von X und Y . Sind X und Y unkorreliert? (3P)

(b) Bestimme die Dichten der Zufallsvariablen $U := e^X$ und $V := aX$, $a \in (0, \infty)$. (2P)

5. Ein Versicherungsunternehmen hat n gleichartige Verträge mit einjähriger Laufzeit abgeschlossen und muss für den i -ten Vertrag den zufälligen Schaden X_i begleichen. Es wird angenommen, dass die X_i unabhängig und identisch verteilt sind mit Erwartungswert m und Varianz $\sigma^2 \in (0, \infty)$. Als Prämie verlangt die Versicherung (gemäß dem sogenannten Varianzprinzip) jeweils $\pi = m + \lambda\sigma^2$ für ein $\lambda > 0$.

(a) Es bezeichne R die Kapitalreserve der Versicherung, und S_n sei die Summe der Einzelleistungen X_i . Bestimme näherungsweise die Ruinwahrscheinlichkeit $P(S_n > R + n\pi)$. (2P)

(b) Wie groß ist nach (a) die Ruinwahrscheinlichkeit für $R = 1440$, $\sigma = 40$, $\lambda = 0,001$ und $n = 900$? Vergleiche die ermittelte Wahrscheinlichkeit mit der Schranke aus der Tschebyschev-Ungleichung. (2P)

(c) Wie groß muss nach (a) die Anzahl der Verträge mindestens sein, damit für $R = 0$, $\sigma = 40$ und $\lambda = 0,001$ die näherungsweise Ruinwahrscheinlichkeit kleiner als 0,01 ausfällt? (1P)