WS 2008/2009 *Name:*
Matrikel-Nr.:

PROBEKLAUSUR
STOCHASTIK I – WAHRSCHEINLICHKEITSTHEORIE

HILFSMITTEL: A4 BLATT MIT NOTIZEN, TR

Die Prüfung besteht aus zwei Teilen, die insgesamt 300 Punkte ergeben:
Teil 1: Multiple Choice (jede Aufgabe 5 Punkte)
Teil 2: 5 von 8 Aufgaben mit jeweils 50 Punkten

Bitte ankreuzen, welche Aufgaben bewertet werden sollen (5 von 8)!

Aufgabe:	1	2	3	4	5	6	7	8	MC	Gesamt
Bewerten:										
Punkte:										

Schreiben Sie bitte auf jedes Blatt oben rechts Ihren Namen!

Viel Erfolg!

Teil 1. Multiple Choice

Zu jeder Frage ist genau eine richtige Antwortmöglichkeit vorgegeben. Tragen Sie Ihre Lösungen in die Kästchen auf der **nächsten** Seite ein. Die Rückseite der Blätter können Sie für Berechnungen sowie zu Anmerkungen und Erläuterungen Ihrer Lösung verwenden.

(1) Eine Zufallsvariable X besitze die Dichtefunktion $f(x) = ke^{-\frac{1}{2}\frac{(x-a)^2}{(\sigma^2+\tau)}}$. Was ist k,
a) $\frac{1}{\sigma\sqrt{2\pi}}$ b) $\frac{1}{\sqrt{\sigma+\tau}\sqrt{2\pi}}$ c) $\frac{1}{\sqrt{\sigma^2+\tau}\sqrt{2\pi}}$ d) $\frac{1}{\sqrt{\sigma^2+\tau^2}\sqrt{2\pi}}$?

(2) Welches der folgenden Ereignispaare könnte man am Besten als unabhängig bezeichnen?
a) Es wird am Samstag in Augsburg regnen. Es wird am Sonntag in Augsburg regnen.
b) Es wird am Samstag in Augsburg regnen. Es wird am Sonntag in Beijing regnen.
c) Es wird am Samstag in Augsburg regnen. Es wird am Sonntag in München regnen.
d) Es wird am Samstag in Augsburg regnen. Es wird am entsprechenden Samstag im nächsten Jahr in Augsburg regnen.

(3) Die Pareto Verteilung wird oft für das Modellieren von Einkommen eingesetzt. Die Dichte hat die Form

$$\begin{aligned} f(x) &= \frac{a}{k}\left(\frac{k}{x}\right)^{a+1} && x \geq k > 0 \\ f(x) &= 0 && x < k \end{aligned}$$

Wie sollte man die Dichte interpretieren?
a) Es gibt keine Leute mit einem Einkommen kleiner k.
b) Es gibt zuviele Leute mit kleinen Einkommen, so dass man diesen Teil der Verteilung nicht modellieren kann.
c) Das Modell ist eigentlich nur für $x \geq k$ gültig.
d) $P(\text{Jemand ein Einkommen} < k \text{ hat }) = 0.05$.

(4) Die Wahrscheinlichkeitserzeugendefunktion einer Zufallsvariablen X hat die Form

$$g_X(z) = 0,1 + 0,2z + 0,2z^2 + 0,3z^3 + 0,2z^6$$

Was ist die Varianz von X?
a) $8,20$ b) $3,61$ c) $1,90$ d) $2,70$

(5) Im Rahmen eines medizinischen Forschungsprojekts muss eine Ärztin 20 Menschen in zwei Gruppen verteilen. Nach Vorschrift verteilt sie die Leute zufällig. Unter den Menschen waren 10 Frauen und 10 Männer. Was ist die Wahrscheinlichkeit, dass eine der zwei Gruppen mehr als 7 Frauen hat?

a) $\sum_{i=8}^{10} \binom{10}{i} \left(\frac{1}{2}\right)^{10}$ b) $\sum_{i=8}^{10} \frac{\binom{10}{i}\binom{10}{10-i}}{\binom{20}{10}}$ c) $2\sum_{i=8}^{10} \frac{\binom{10}{i}\binom{10}{10-i}}{\binom{20}{10}}$ d) $\frac{\binom{10}{8}\binom{10}{2}}{\binom{20}{10}}$

(6) $X \sim N(\mu, \sigma^2)$. Der Parameter μ wird als Zufallsvariable mit apriori Verteilung $U(10, 20)$ modelliert. Wie sieht die aposteriori Verteilung für μ aus nach einer Beobachtung $X = 14$?

a) $g(\mu) = \dfrac{\frac{1}{\sigma^2\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(14-\mu)^2}{\sigma^2}} \frac{1}{10}}{\int_{10}^{20} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(14-\mu)^2}{\sigma^2}} \frac{1}{10} \mathrm{d}\mu}$ b) $N(14, \sigma^2)$

c) $g(\mu) = \dfrac{\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(14-\mu)^2}{\sigma^2}}}{\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(14-\mu)^2}{\sigma^2}} \mathrm{d}\mu}$ d) $g(\mu) = \dfrac{e^{-\frac{1}{2}\frac{(14-\mu)^2}{\sigma^2}}}{\int_{10}^{20} e^{-\frac{1}{2}\frac{(14-\mu)^2}{\sigma^2}} \mathrm{d}\mu}$

(7) Für den Median, m, einer Zufallsvariable gilt $F(m) = 0,5$. Der Median von $X \sim E(\lambda)$ ist

a) $\frac{1}{\lambda \log 2}$ b) $\frac{1}{\lambda}$ c) $\frac{1}{\lambda^2}$ d) $\frac{\log 2}{\lambda}$?

(8) $X \sim B(600; 0.4)$ und die Dichte wird mit einer Normalverteilung approximiert. Wie würden Sie die Wahrscheinlichkeit $P(200 \leq X < 240)$ schätzen?

a) $0.5 - \Phi\left(-3\frac{1}{3}\right)$ *b)* $\Phi\left(\frac{1}{24}\right) - \Phi\left(-3\frac{3}{8}\right)$

c) $0.5 + \Phi\left(-3\frac{3}{8}\right)$ *d)* $\Phi\left(-\frac{1}{24}\right) - \Phi\left(-3\frac{3}{8}\right)$

(9) Was kennzeichnet ein Bernoulli-Experiment?
(a) Es ist ein Zufallsexperiment, das beliebig viele Ausgänge haben kann.
(b) Es ist ein Zufallsexperiment mit genau zwei möglichen Ausgängen.
(c) Die möglichen Ausgänge eines einzelnen Bernoulli-Experiments haben immer die gleiche Wahrscheinlichkeit.
(d) Es ist ein Zufallsexperiment, dessen Ausgang durch eine stetige Zufallsvariable beschrieben werden kann.

(10) Wie wahrscheinlich ist es, dass bei viermaligem Würfeln eine 4, 5 oder 6 dabei ist?
(a) $\frac{11}{24}$ (b) $\frac{15}{16}$. (c) $\frac{35}{36}$ (d) $\frac{15}{36}$

Antworten:

1	2	3	4	5	6	7	8	9	10

Verteilungstabelle der Standardnormalverteilung

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0.000	0.5000	0.750	0.7734	1.500	0.9332	2.250	0.9878
0.025	0.5100	0.775	0.7808	1.525	0.9364	2.275	0.9885
0.050	0.5199	0.800	0.7881	1.550	0.9394	2.300	0.9893
0.075	0.5299	0.825	0.7953	1.575	0.9424	2.325	0.9900
0.100	0.5398	0.850	0.8023	1.600	0.9452	2.350	0.9906
0.125	0.5497	0.875	0.8092	1.625	0.9479	2.375	0.9912
0.150	0.5596	0.900	0.8159	1.650	0.9505	2.400	0.9918
0.175	0.5695	0.925	0.8225	1.675	0.9530	2.425	0.9923
0.200	0.5793	0.950	0.8289	1.700	0.9554	2.450	0.9929
0.225	0.5890	0.975	0.8352	1.725	0.9577	2.475	0.9933
0.250	0.5987	1.000	0.8413	1.750	0.9599	2.500	0.9938
0.275	0.6083	1.025	0.8473	1.775	0.9621	2.525	0.9942
0.300	0.6179	1.050	0.8531	1.800	0.9641	2.550	0.9946
0.325	0.6274	1.075	0.8588	1.825	0.9660	2.575	0.9950
0.350	0.6368	1.100	0.8643	1.850	0.9678	2.600	0.9953
0.375	0.6462	1.125	0.8697	1.875	0.9696	2.625	0.9957
0.400	0.6554	1.150	0.8749	1.900	0.9713	2.650	0.9960
0.425	0.6646	1.175	0.8800	1.925	0.9729	2.675	0.9963
0.450	0.6736	1.200	0.8849	1.950	0.9744	2.700	0.9965
0.475	0.6826	1.225	0.8897	1.975	0.9759	2.725	0.9968
0.500	0.6915	1.250	0.8944	2.000	0.9772	2.750	0.9970
0.525	0.7002	1.275	0.8988	2.025	0.9786	2.775	0.9972
0.550	0.7088	1.300	0.9032	2.050	0.9798	2.800	0.9974
0.575	0.7174	1.325	0.9074	2.075	0.9810	2.825	0.9976
0.600	0.7257	1.350	0.9115	2.100	0.9821	2.850	0.9978
0.625	0.7340	1.375	0.9154	2.125	0.9832	2.875	0.9980
0.650	0.7422	1.400	0.9192	2.150	0.9842	2.900	0.9981
0.675	0.7502	1.425	0.9229	2.175	0.9852	2.925	0.9983
0.700	0.7580	1.450	0.9265	2.200	0.9861	2.950	0.9984
0.725	0.7658	1.475	0.9299	2.225	0.9870	2.975	0.9985

Teil 2.

Bearbeiten Sie 5 der 8 Aufgaben!

1. Reisezeiten

Es gibt zwei Routen von A bis C, direkt oder über B. Die Fahrtzeiten werden als Zufallsvariablen betrachtet, wobei

$$T_{AB} \sim U(10, 20), \qquad T_{BC} \sim U(5, 15), \qquad T_{AC} \sim U(20, 30)$$

(a) Was ist die Wahrscheinlichkeit, dass die Direktroute AC schneller ist?

Statt Gleichverteilungen könnten Normalverteilungen mit denselben Erwartungswerten und Varianzen verwendet werden. Dagegen wird argumentiert, dass sehr kleine oder sogar negative Reisezeiten möglich wären.

(b) Auf welcher Strecke gäbe es die höchste Wahrscheinlichkeit einer negativen Reisezeit und wie groß wäre sie?

2. Ausreisser

Werte, die weit weg von anderen liegen, werden als Ausreißer oder Extremwerte bezeichnet. Es gibt verschiedene Faustregel, wie, z.B., dass Werte außerhalb $\bar{x} \pm 3s$, als Ausreißer betrachtet werden sollten (wo $\bar{x}$ der Stichprobenmittelwert und s die Stichprobenstandardabweichung sind).

(a) Wenn X normalverteilt ist, was ist die Wahrscheinlichkeit, dass ein Wert außerhalb $\mu \pm 3\sigma$ liegt?

(b) Was ist die Wahrscheinlichkeit, dass in einer Stichprobe der Größe N, r oder mehr Werte außerhalb dieser Grenzen liegen?

(c) Sei $N = 100$; was ist die genaue Wahrscheinlichkeit, dass keine Werte außerhalb dieser Grenzen liegen? Wie gut sind die Poisson- und Normalapproximationen hier?

(d) In der Praxis muss man mit $\bar{x}$ und s, nicht mit μ und σ, arbeiten. Wie groß müsste N sein, dass mit einer Wahrscheinlichkeit von 99.5%, $\bar{x}$ nicht weiter als 0.01σ von μ fällt?

3. Benfords Gesetz

Nach dem Gesetz von Benford ist die Verteilung der Werte der ersten Ziffer einer Zahl

$$P(X_1 = n) = \log_{10}(1 + \frac{1}{n}) \quad n = 1, \dots, 9$$

Sei G eine große Gruppe von achtstelligen Zahlen. Sie und ein Bekannter ziehen per Zufall je eine Zahl und halten sie geheim.

(a) Unter der Annahme, dass das Gesetz von Benford gilt, was ist die Wahrscheinlichkeit, dass Sie eine größere Zahl als Ihr Bekannter gezogen haben, wenn er preisgibt, dass seine Zahl kleiner 4 ist? Erläutern Sie die Annahmen, die Sie in Ihrer Berechnung machen.

Das Gesetz von Benford kann auf mehrere Ziffer erweitert werden. Die Verteilung der beiden ersten Ziffern ist

$$P(X_1 = n \text{ und } X_2 = m) = \log_{10}(1 + \frac{1}{10n + m})$$
$$n = 1, \dots, 9 \quad m = 0, \dots, 9$$

(b) Was ist $P(X_2 = m)$?

(c) Was ist $P(X_2 = m | X_1 = n)$?

(d) Was ist $P(X_1 = n | X_2 = m)$?

(e) Welche Wahrscheinlichkeit ist größer:
$P(X_2 = 0 | X_1 < 5)$ oder $P(X_2 = 0 | X_1 \geq 5)$?

4. Ausschlafen

Dr. T hat 2 Kinder. Es ist zu erwarten, dass Tochter A morgens bis 7:30 ausschläft, während Sohn S erwartungsgemäß um 7:40 aufwacht. Die Standardabweichung der Aufwachzeit der Tochter betrage 15 min, die des Sohnes 30 min.

(a) Dr. T verlässt um 7:05 das Haus. Mit welcher Wahrscheinlichkeit kann er sich
   i) von seiner Tochter verabschieden?
   ii) von mindestens einem Kind verabschieden?

(b) Mit welcher Wahrscheinlichkeit wacht S vor A auf?

(c) Am Wochenende möchte Familie T aussschlafen. Mit welcher Wahrscheinlichkeit wacht keines der beiden Kinder vor 8:00 auf?

(d) Skizzieren Sie die verwendeten Dichtefunktionen — welche Annahmen haben Sie für (a)–(c) getroffen?

5. Übungsleitertreffen

Hiwi O und S wollen sich Donnerstag zwischen 12 und 13 Uhr treffen um Übungsblätter zu korrigieren. Da sie beide an der Uni sehr beschäftigt sind, können sie beide keinen exakten Termin einhalten. Daher kann man davon ausgehen, dass sie zufällig zwischen 12 und 13 Uhr am Treffpunkt ankommen.

Da es sich beide nicht leisten können, länger auf den anderen zu warten, vereinbaren O und S, dass sie maximal 10 min aufeinander warten, maximal jedoch bis 13 Uhr (um 13 Uhr verlassen beide den Treffpunkt).

(a) Mit welcher Wahrscheinlichkeit verpassen sich O und S nicht?

(b) Welche Wartezeit müssen beide ausmachen, damit sie sich mit einer Wahrscheinlichkeit von 0.75 treffen?

(c) Gegeben, dass sie sich getroffen haben, was ist die Wahrscheinlichkeit, dass O vor 12:45 angekommen ist?

(d) Interpretieren Sie das Problem geometrisch.

6. Pizza

Alle Angehörigen des Mathematikinstituts gehen im Il Porcino essen. Da die Pizza dort so lecker ist, bestellen alle Pizza. Um die Bestellung für den Kellner zu erleichtern werden die möglichen Beläge auf 3 beschränkt. 60 Prozent der Personen wählen Salami, 70 Prozent Pilze und 40 Prozent Thunfisch.

(a) Wieviele Personen haben mindestens sowohl Salami als auch Pilze auf ihrer Pizza?
(b) Falls die Wahl von Salami und Thunfisch voneinander unabhängig sind, wie hoch ist der Anteil Personen die nicht Salami und Thunfisch gleichzeitig auf ihrer Pizza haben?

Beim Betrachten der Rechnung wird deutlich, dass 45 Prozent der Personen eine Pizza mit Salami und Pilzen und 55 Prozent der Personen eine Pizza mit Thunfisch und Pilzen gewählt haben.

(c) Wie groß ist die Wahrscheinlichkeit Salami oder Thunfisch auf der Pizza zu finden, wenn man weiß, dass die Person auf jeden Fall Pilze auf ihrer Pizza hat? Es ist bekannt, dass 65 Prozent der Personen nicht alle drei Zutaten gleichzeitig auf der Pizza haben.

Dr. T ist auf Diät und wählt deshalb eine Pizza ohne Belag. Da auch einige figurbewußte Damen anwesend sind, schließen diese sich Dr. T an und bestellen auch Pizzas ohne Belag.

(d) Wie hoch ist der Anteil der Personen, die eine Pizza ohne Belag bestellen, wenn davon ausgegengen werden kann, dass die Entscheidung für Salami von der für Thunfisch unabhängig ist?

7. Erwartungswert

Die Zufallsvariablen $X_1, X_2, \ldots, X_n$ seien unabhängig und jede habe folgende Verteilungsfunktion:

$$F(x) = \begin{cases} 0 & x < 0; \\ 2x - x^2 & 0 \leq x \leq 1; \\ 1 & 1 < x. \end{cases}$$

(a) Geben Sie die Dichte von X_1 an.
(b) Berechnen Sie $E[e^{2X_1}]$.
(c) Sei $S_n = \sum_{i=1}^{n} X_i$. Berechnen Sie Erwartungswert und Varianz von S_n.
(d) Finden sie die Folgen (a_n) und (b_n) so, dass die Verteilung von $\frac{S_n - a_n}{b_n}$ gegen $N(0,1)$ konvergiert.

8. R Output

Eine veröffentlichte Studie der Leukämie Fälle unter deutschen Kindern hat viel Aufsehen erregt. Einige Berechnungen in R sind hier gemacht worden, um die Ergebnisse der Studie zu überprüfen.

Erläutern Sie den Input und interpretieren Sie den Output. Welche Berechnungen sind gemacht worden und warum? Formulieren Sie Ihre Antworten in kurzer Berichtform.

Für einen Zeitraum von zwanzig Jahren sind 17 Fälle erwartet worden, und es gab in der Tat 37 Fälle.

```
> 1-ppois(36,17)
[1] 1.802502e-05

> ppois(25,17)-ppois(8,17)
[1] 0.96216

> p1<-1-ppois(1,17/20)
> p1
[1] 0.2092824

> p2<-1-ppois(2,17/20)
> p2
[1] 0.05487873

> 1-(1-p2)^20
[1] 0.676592
> 1-(1-p1)^20
[1] 0.9908709

> pbinom(2,20,p2)
[1] 0.9061585
> dbinom(0,20,p2)
[1] 0.323408

> y20<-rpois(20,17/20)
> mean(y20)
[1] 0.7
> table(y20)
y20
 0  1  2  3
10  7  2  1
```

```
> y1000<-rpois(1000,17/20)
> mean(y1000)
[1] 0.852

> table(y1000)
y1000
  0   1   2   3   4   5   6
410 385 162  34   6   1   2

> z1000<-1000*dpois(0:6,17/20)
> z1000
[1] 427.4149319 363.3026922 154.4036442
[4]  43.7476992   9.2963861   1.5803856
[7]   0.2238880
```