

# **Klausur + Musterlösung** **zur Vorlesung** **Adaptive Modellierung und Simulation** **Sommersemester 2015**

Datum: 21.07.2015

Vorname:
Name:
Matrikelnummer:
Geburtsdatum:
Studiengang:

Die Punktzahl einer Aufgabe entspricht ungefähr der maximalen Bearbeitungsdauer der Aufgabe in Minuten. Die Klausur enthält 16 Aufgaben (160 Punkte), aus denen 120 Punkte erzielt werden müssen, um 100% der Klausur zu bestehen. Sie müssen also nicht alle Aufgaben lösen, um die Klausur vollständig zu bestehen.

Durch die Übungspunkte können maximal 10% der Klausurleistung erbracht werden. Als Hilfsmittel sind nur ein Taschenrechner und dokumentenechte Stifte erlaubt. Bitte benutzen Sie für Notizen die Rückseiten der Aufgabenblätter.

*Viel Erfolg!*

Wird vom Prüfer ausgefüllt:

1	2	3	4	5	6	7	8		Σ
/9	/8	/6	/15	/9	/7	/8	/8		
9	10	11	12	13	14	15	16		
/5	/7	/6	/40	/8	/6	/10	/8		/120

% Klausur:                      % Übungen:                      % Gesamt:                      Note:

**Aufgabe 1 Lineare Modellierung****9 Pkte**

- a) Wie lautet die Gleichung eines stochastischen linearen Modells? (1P)
- b) Was versteht man unter „Linearisierung“? Geben Sie ein Beispiel. (2P)
- c) Bei der Parameterschätzung: Was ist der Unterschied zwischen „erwartungswerttreu“ und „konsistent“? (4P)
- d) Warum ist es sinnvoll, die gemessenen Variablen vor der Analyse zu normieren? Und wie geht das? (2P)

*Lösungen:*

- a)  $f(x) = a_0 + a_1x + u$  mit den Parametern  $a_0$ ,  $a_1$  und der Zufallsabweichung  $u$
- b) Werden die Messdaten vorher einer nichtlinearen Operation unterworfen, so kann der Gesamtzusammenhang oftmals wieder linear dargestellt und die Parameter mit den Methoden der linearen Systeme bestimmt werden. (1P)

*Beispiel: der Zusammenhang  $f(x) = a_0 + a_1x_1 + a_2x_2^2 + u$  kann mit der Linearisierung  $x_2^* = x_2^2$  linear zu  $f(x) = a_0 + a_1x_1 + a_2x_2^*$  dargestellt werden. (1P)*

- c) erwartungswerttreu: der echte Parameter ist gleich dem Erwartungswert aller bestimmten Parameterwerte.  $\langle a(t) \rangle = a$  (2P)

*konsistent: die bestimmten Parameterwerte konvergieren zum echten Wert mit Wahrscheinlichkeit eins.  $\lim \text{Prob}(|a(t) - a| < \varepsilon) = 1$  für alle  $\varepsilon > 0$  (2P)*

- d) Da die verschiedenen Variablen in verschiedenen (willkürlichen) Einheiten gemessen werden, sind sie mit einem (willkürlichen) Skalierungsfaktor behaftet. (1P)

*Es ist deshalb sinnvoll, alle Variablen vorher durch eine Normierung  $x^* = (x - \mu) / \sigma$  auf Mittelwert null und Varianz 1 zu bringen. (1P)*

**Aufgabe 2 TLMSE****8 Pkte**

- a) Was ist der Total Least Mean Squared Error? (2P)
- b) Warum ist er ein besseres Maß für die Güte einer linearen Approximation als der „normale“ quadratische Fehler MSE? (2P)
- c) Wie wird die approximierende Hyperebene mit kleinstem TLMSE berechnet? (4P)

*Lösungen:*

- a) Der TLMSE misst die quadratische Abweichung als kürzesten Abstand der Datenpunkte zu der approximierenden Hyperebene.
- b) Im Unterschied zum MSE ist er unabhängig von der Lage des Koordinatensystems, ist also invariant bezüglich der Rotation der Basisvektoren.
- c) Zur Berechnung wird der Eigenvektor der Kovarianzmatrix der Daten gebildet mit dem kleinsten Eigenwert. Dieser Vektor steht senkrecht auf der approximierenden Hyperebene und dient der Beschreibung der Hyperebene mittels der Hesseschen Normalform.

**Aufgabe 3 Online vs. Offline-Learning****6 Pkte**

- a) Was ist der Unterschied zwischen offline-learning und online-learning? (3P)  
 b) Welche Methode konvergiert meist schneller und warum? (3P)

Lösungen:

- a) *Beim Offline-Learning sind alle Muster zum Training vorhanden. Beim Online-Algorithmus treten die Trainingsmuster nacheinander auf und werden nicht gespeichert. (1P)*  
*Zum Lernen werden beim Online-Learning nach jedem Trainingsmuster sofort die Gewichte verbessert. Im Gegensatz dazu werden beim Offline-Learning die benötigten Gewichtsverbesserungen nur erfasst und aufsummiert. Nach dem Durchlauf aller Trainingsdaten wird dann in einem Schritt die akkumulierte Gewichtsverbesserung angewendet, so dass hier die Verbesserung den Erwartungswert darstellt. (2P)*
- b) *Meist konvergiert Online-Learning schneller, da bei jedem Schritt das System verbessert wird und der Fehler gegenüber dem verbesserten System erfasst wird und nicht gegenüber dem Originalsystem. (2P)*  
*Allerdings kann es bei stark schwankender Statistik der Eingaben zu Fehladaptationen beim Online-Learning kommen, so dass in diesem Fall das Offline-Learning mit der Erwartungswertbildung besser abschneiden kann. (1P)*

**Aufgabe 4 Training und Testen****15 Pkte**

Angenommen, Sie haben von 3 Klassen je 10 Datenquellen (Individuen) und von jeder Quelle 3 Muster, also insgesamt 90 Muster für Training, Testen und Validieren.

- a) Wie müssen Sie Trainings- und Testmengen bilden, um die Klassifizierung zu lernen, und was müssen Sie dabei beachten? (5P)
- b) Wie funktioniert die Kreuzvalidierung? (5P)
- c) Wie funktioniert „stopped training“? (5P)

Lösungen:

- a) *Die Mustermenge wird in 2 (Training und Testen) oder 3 (Training, Testen, Validieren) Mengen unterteilt, je nach Aufgabenstellung. (2P)*  
*Dabei ist zu beachten, dass alle Samples einer Quelle in derselben Menge liegen, um einen Transfer der Information über die Quelle zwischen den Mengen, etwa Training und Testen, zu verhindern. (3P)*
- b) *Haben wir zu wenig Muster in einer Menge, so können wir die Anzahl dadurch erhöhen, dass wir fast alle Muster zum Training verwenden und nur wenige (minimal 1) zum Testen. (2P)*  
*Allerdings müssen wir dann das Training wiederholen, wobei jeweils die Testmenge anders zusammengestellt wird. (1P)*  
*Am Schluss werden alle Testergebnisse über die Anzahl der Testläufe gemittelt. (2P)*
- c) *Bemerken wir einen starken Unterschied zwischen den Trainingsergebnissen und den Testergebnissen, so liegt ein overfitting vor. In diesem Fall können wir nach jeder Iteration einen Test mit allen Mustern einer Validierungsmenge durchführen und das Training genau dann abbrechen, wenn der Validierungsfehler nicht mehr ab-, sondern zunimmt. (3P)*  
*Allerdings muss der Erfolg noch durch eine weitere, vom Training unabhängige Testmenge überprüft werden. (2P)*

**Aufgabe 5 RBF-Netze****9 Pkte**

- a) Was ist eine Glockenfunktion? (2P)
- b) Was sind die wesentlichen Vorteile und Nachteile von RBF-Netzen gegenüber Multi-Layer-Perzeptrons? (5P)
- c) Wie bestimmen Sie die Parameter der ersten und der zweiten Schicht? (2P)

*Lösungen:*

- a) *Eine Glockenfunktion  $f(x) \geq 0$  ist null im neg. Unendlichen, monoton steigend mit  $x$  bis zu einem endl. Wert  $f(x^*) > 0$ , und fällt bei  $x^* < x$  zum pos. Unendlichen wieder monoton bis auf null. Außerdem ist das Integral  $> 0$ .*
- b) *Vorteile: schichtweises, einfaches Training (1P), schnellere Konvergenz (1P), inkrementierbares Netz (2P), weniger Neuronen pro Klassenentscheidung bei geclusteter Eingabe. (1P)*  
*Nachteile: schlechte Extrapolationsmöglichkeiten (2P)*
- c) *erste Schicht: z.B. Rasterung im gleichen Abstand;  $c_i$  ist dann der Mittelpunkt jedes Rasterelements. (1P) Zweite Schicht: linear über TLMSE, s. Aufgabe 2 (1P)*

**Aufgabe 6 Approximationseigenschaften Neuronaler Netze 7 Pkte**

Angenommen, Sie haben ein neuronales Netz.

- a) Wie gut kann das Netz eine vorgegebene Funktion approximieren? (2P)
- b) Was sind die Vorbedingungen dafür? (5P)

*Lösungen:*

- a) *Ein zweischichtiges Netz kann jede beliebige Funktion beliebig dicht approximieren. (2P)*
- b) *Dies ist aber nur bei genügend großer Anzahl von Neuronen der Fall. (2 P). Weiterhin muss die erste Schicht nicht-lineare Einheiten haben, etwa sigmoidale oder Glocken-Ausgabefunktionen. (2P) Die zweite Schicht kann linear sein, muss aber nicht. (1P)*

**Aufgabe 7 Wissensbasierte vs. black-box-Modellierung****8 Pkte**

- a) Was sind die Vor- und Nachteile einer wissensbasierten Modellierung gegenüber einer black-box-Modellierung mit neuronalen Netzen? (4P)
- b) In welchen Fällen ziehen Sie eine wissensbasierte Modellierung einer black-box-Modellierung vor? Nennen Sie Beispiele. (2P)
- c) Angenommen, Sie haben ein Gütekriterium  $R(\theta)$  für die beobachtete Modelleleistung definiert. Mit welchen Gleichungen (Algorithmen) erhalten Sie die Werte für die optimalen Modellparameter  $\theta$ ? Welche beiden Verfahren wurden dazu in der Vorlesung eingeführt? (2P)

*Lösungen:*

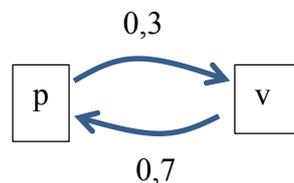
- a) *Vorteile: weniger Parameter sind nötig und damit weniger Trainingsbeispiele für eine gute Bestimmung der Parameter (1P). Außerdem ist sie meist plausibler. (1P)*  
*Nachteile: Es gibt nicht bei allen Problemen ausreichendes Wissen über die zugrunde liegenden Zusammenhänge (1P). Wenn bei der Modellierung Fehler gemacht werden, kann das Modell die Realität nicht gut abbilden und es erscheinen starke Abweichungen trotz Parameterabgleich. (1P)*

- b) Wenn wenige, gut dokumentierte quantitativ messbare Beziehungen existieren, etwa Ursache-Wirkungszusammenhänge im naturwissenschaftlich-technischen Bereich. Beispiele: chem. Reaktionen, physikalische Effekte, biologische beobachtbare Einflüsse, geologische Wirkungen. (2P)
- c) Für die iterative Schätzung der Parameter  $\theta$  wurden das Gradientenverfahren mit  $\theta(t+1) = \theta(t) - \gamma R(\theta)^t$  und das Newton-Verfahren  $\theta(t+1) = \theta(t) - R(\theta)^t / R(\theta)^t$  eingeführt. (2P)

## Aufgabe 8 Dynamische Systeme

8 Pkte

Gegeben sei folgendes System



- a) Wie lauteten die Differenzialgleichungen dafür? (1P)
- b) Fassen Sie beide Zustände zu einem Zustandsvektor zusammen. Wie lautet dann die Differenzialgleichung und wie die Systemmatrix? (1P)
- c) Ist das System stabil, labil, oszillierend? (6P)

Lösungen:

a)  $v' = 0,3p + v, \quad p' = 0,7v + p$

b)  $z = (v, p), \quad z' = \begin{bmatrix} 1 & 0,3 \\ 0,7 & 1 \end{bmatrix} z = Az$

- c) Es muss das char. Polynom gebildet werden, um die Bedingung  $|A - I\lambda| = 0$  zu erfüllen. (2P)

Es ist  $|A - I\lambda| = (1 - \lambda)(1 - \lambda) - 0,7 \cdot 0,3 = 0$  oder  $\lambda^2 - 2\lambda + 1 - 0,21 = 0$  und damit  $\lambda_1 = 1 \pm \sqrt{1 - 0,21} = 1,88$  und  $\lambda_2 = 0,12$ . (2P)

Die Eigenwerte sind reell und das System ist also nicht oszillierend. Da beide Eigenwerte  $> 0$ , ist das System labil divergent. (2P)

## Aufgabe 9 Gekoppelte Systeme

5 Pkte

Seien zwei Populationen  $N_1$  und  $N_2$  gegeben mit ihren Geburts- und Todesraten  $\alpha$  und  $\beta$ . Die Änderungen seien  $N_1' = (\alpha_1 - \beta_1)N_1$  und  $N_2' = (\alpha_2 - \beta_2)N_2$ .

Nun sei  $N_1$  eine Schmetterlingsart und  $N_2$  eine Schlupfwespenart. Die Schlupfwespen pflanzen sich fort, indem sie ihre Eier in die Raupen der Schmetterlinge legen. Die Raupen sterben und bilden das Anfangsfutter für die Larven der Wespen.

Zusammen bilden Sie also ein Parasiten-Wirts-System, bei dem die Geburtenrate des Wirts  $\alpha_1$  durch die Menge der Infektion des Parasiten  $N_2$  bestimmt wird und umgekehrt:  $\alpha_1 = \gamma_1 - c_1 N_2$  und  $\alpha_2 = \gamma_2 + c_2 N_1$ .

- a) Wie lauten also die Differentialgleichungen für dieses Parasiten-Wirts-System? (1P)
- b) Wie verhält sich das System dynamisch? Vergleichen Sie es dazu mit dem Ihnen bekannten Räuber-Beute-System. (4P)

Lösungen:

a) Die Differentialgleichungen sind

$$N_1' = (\alpha_1 - \beta_1)N_1 = (\gamma_1 - c_1N_2 - \beta_1)N_1 = (\gamma_1 - \beta_1)N_1 - c_1 N_1N_2 \quad (1P)$$

$$N_2' = (\alpha_2 - \beta_2)N_2 = (\gamma_2 + c_2N_1 - \beta_2)N_2 = (\gamma_2 - \beta_2)N_2 + c_2 N_1N_2$$

b) Diese Differentialgleichungen sind formal gleich mit denen des Räuber-Beute-Systems, nur die Konstanten sind andere. Also ist die Dynamik analog zum Räuber-Beute-System durch phasenverschobene, periodische Schwingungen gegeben, die durch die Wirtspopulation angeführt wird, wobei man die Parasiten als Räuber und die Wirte als Beute betrachten kann. (4P)

## Aufgabe 10 Gleichgewichtssysteme

7 Pkte

- a) Was ist ein Gleichgewichtssystem? Geben Sie ein Beispiel. (4P)  
 b) Welchen Vorteil bietet es bei einer Modellierung, und welchen Nachteil hat es? (2P)  
 c) Wann sollte man es verwenden, wann nicht? (1P)

Lösungen:

a) Ein Gleichgewichtssystem ist ein dynamisches System aus interagierenden Wirkungen, das einen Zustand des Gleichgewichts annehmen kann, in dem es nach außen hin konstant erscheint. Dieser konstante Zustand wird bei allen anderen Wechselwirkungen nach außen hin vorausgesetzt. (2P)

Beispiel: Das Gleichgewicht zwischen einer mit zunehmendem Preis  $x$  zunehmender Produktion  $f(x)$  und einer mit dem Preis abnehmenden Konsum  $g(x)$  tritt bei einem Marktpreis  $x^*$  ein, bei dem gilt  $f(x^*) = g(x^*)$ . (2P)

b) Vorteil: Anstelle einer zeitlichen Dynamik hat es einen konstanten Zustand. (1P)

Nachteil: Das Gleichgewicht wird erst nach einer gewissen Zeit  $T$  erreicht. Ändern sich innerhalb dieser Zeit die äußeren Bedingungen (Parameter), so ändert sich auch die Dynamik und das Gleichgewicht wird nicht erreicht. Das System wird in dieser Zeit durch den Gleichgewichtszustand falsch beschrieben. (1P)

c) Verwenden sollte man es, wenn die Zeit  $T$  deutlich kleiner ist als die Zeit, in der sich der Kontext ändert. (1P)

## Aufgabe 11 RAS

6 Pkte

Angenommen, Sie haben Produzentengruppen und Produktkategorien. Bekannt sind nur die Summenwerte der jeweiligen Gruppen und der Kategorien. Sie wollen eine Input-Output-Tabelle erstellen und verwenden dazu das RAS-Verfahren.

- a) Wie lauten die algorithmischen Schritte des Verfahrens? (4P)  
 b) Wann ist das Verfahren zu Ende? Gibt es immer ein eindeutiges Ergebnis? (2P)

Lösungen:

a) Die Schritte bestehen aus folgenden Schritten, wobei die gesuchte Matrix  $A_{ij}$  mit zufälligen Werten vorbelegt wird, um an die vorgegebenen Werte der Spalten  $S^*$  und Zeilen  $R^*$  anzupassen:

- Bilde die Zeilensumme  $R_i = \text{Summe}_j(A_{ij})$  für alle  $i$
- Normiere die Zeilen mit  $A_{ij} = A_{ij} (R_i^*/R_i)$
- Bilde die Spaltensumme  $S_j = \text{Summe}_i(S_{ij})$  für alle  $j$
- Normiere die Zeilen mit  $A_{ij} = A_{ij} (S_j^*/S_j)$

in einem Durchlauf.

b) Dies wird solange wiederholt, bis eine Konvergenz eingetreten ist, also die mittleren Abweichungen kleiner als ein gewünschtes  $\varepsilon$  sind:  $\text{Summe}_i |R_i - R_i^*| + \text{Summe}_j |S_j - S_j^*| < \varepsilon$ . (1P)

Allerdings ist dabei keine eindeutige Lösung für die  $n \times m$  Werten garantiert, da nur  $n+m$  Nebenbedingungen vorgegeben sind; sie kann von den Startwerten abhängen. (1P)

## Aufgabe 12 Kausalitätsanalyse

40 Pkte

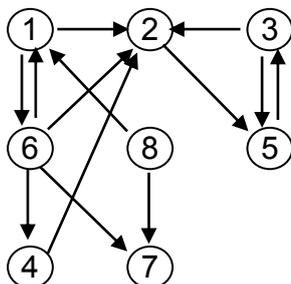
Angenommen, Sie haben ein Modell mit folgenden Abhängigkeiten:

$$\begin{aligned} z_1 &= F_1(z_6, z_8) \\ z_2 &= F_2(z_1, z_3, z_4, z_6) \\ z_3 &= F_3(z_5) \\ z_4 &= F_4(z_6) \\ z_5 &= F_5(z_2, z_3) \\ z_6 &= F_6(z_1) \\ z_7 &= F_7(z_6, z_8) \end{aligned}$$

- Stellen Sie dazu die Adjazenzmatrix und den dazu gehörenden Graphen für die kausale Struktur auf. (4P)
- Vereinfachen Sie die kausale Struktur durch Verschmelzung und Variablensubstitution. Wie sehen nun Graph und Adjazenzmatrix aus? (13P)
- Führen Sie statt Schritt b) eine Blockbildung stark zusammenhängender Knoten durch. Welche Adjazenzmatrix und Graph erhalten Sie nun? (8P)  
Was sind die Unterschiede zum Ergebnis von Schritt b)? (1P)
- Gliedern Sie den Graphen aus c) in Ebenen auf. Wie lautet die Adjazenzmatrix nun? Wie sieht die Erreichbarkeitsmatrix aus? (4P)
- Angenommen, der Zustand  $Z_8$  sei von einem früheren Zeitpunkt  $t-1$  abhängig:  $Z_8 = f(z_2(t-1))$ . Wie sehen dann Graph und Erreichbarkeitsmatrix für zwei Zeitpunkte aus? (10P)

Lösungen:

a) (4P)



	1	2	3	4	5	6	7	8
1	0	1	0	0	0	1	0	0
2	0	0	0	0	1	0	0	0
3	0	1	0	0	1	0	0	0
4	0	1	0	0	0	0	0	0
5	0	0	1	0	0	0	0	0
6	1	1	0	1	0	0	1	0
7	0	0	0	0	0	0	0	0
8	1	0	0	0	0	0	1	0



**Aufgabe 13 Simulation****8 Pkte**

- a) Wie lauten die vier Anforderungen an einen guten Pseudo-Zufallszahlengenerator? (2P)  
 b) Wie lautet die Erzeugendengleichung für den quadratischen LCG? (1P)  
 c) Zeigen Sie: Die Varianz der Summe aus 12 Zufallszahlen von  $U(0,1)$  ist eins. Beachten Sie dazu, dass die Varianz der Summe von unabhängigen Zufallszahlen gerade gleich der Summe aller einzelnen Varianzen ist. (5P)

Lösungen:

- a) - *Unabhängigkeit der Zahlen*  
 - *konstante Besetzungsdichte bei  $U(0,1)$*   
 - *hohe Besetzungsdichte, nur kleine Lücken*  
 - *Reproduzierbarkeit (Pseudo-Zufallszahlen)*
- b)  $x(t+1) = (a x(t) + b x^2(t) + c) \bmod m$
- c) Es ist  $x(t) = \sum_{i=1}^{12} (x_i - 0.5)$ . Die Varianz davon ist

$$\text{var}(x(t)) = \sum_{i=1}^{12} \text{var}((x_i - 0.5))$$

Dabei ist mit  $y_i = x_i - 0.5$  und  $y_i$  aus  $U(-0.5, +0.5)$  auch  $p(y_i) = 1$

$$\text{var}(y_i) = \int_{-0.5}^{+0.5} y_i^2 p(y_i) dy_i = \int_{-0.5}^{+0.5} y_i^2 dy_i = \frac{1}{3} y_i^3 \Big|_{-0.5}^{+0.5} = \frac{1}{3} \left( \frac{1}{8} + \frac{1}{8} \right) = \frac{1}{12}$$

Also ist

$$\text{var}(x(t)) = \sum_{i=1}^{12} \frac{1}{12} = 1 \quad \text{q.e.d.}$$

**Aufgabe 14 RUN-Tests****6 Pkte**

Beschreiben Sie die beiden in der Vorlesung vorgestellten RUN-Tests.

- a) Welche Schritte sind nötig? (je 2 P)  
 b) Welche Testkriterien verwenden Sie? (je 1P)

Lösungen:

**RUN-TEST1**

- a) Bilden von Zeichen  $s_i = 0, 1$  je nach Zahlenfolge

$$s_i = \begin{cases} 1 & x(t+1) \geq x(t) \\ 0 & \text{sonst} \end{cases}$$

- Bilden von Zeichenketten. Es entstehen Zeichenketten gleicher Zahlen der Längen  $r_i$ .
- Zählen: Wie oft kommt die Länge  $r_i$  vor?

- b) Test auf Normalverteilung der normalisierten Beobachtungen  $y_i = (r_i - \langle r_i \rangle) / \sigma_i$   
 Ist  $\langle y_i \rangle$  im Intervall  $y_\alpha < \langle y_i \rangle < y_{1-\alpha}$  ?

**RUN-TEST2**

- a) Bilden von Zeichen  $s_i = 0, 1$  je nach Zahlenfolge

$$s_i = \begin{cases} 1 & x(t) \geq 0.5 \\ 0 & \text{sonst} \end{cases}$$

- Bilden von Zeichenketten. Es entstehen Zeichenketten gleicher Zahlen der Längen  $r_i$ .
- Zählen: Wie oft kommt die Länge  $r_i$  vor?

- b) Vergleich der relativen Häufigkeiten der Beobachtungen  $P(r_i)$  mit den errechneten Wahrscheinlichkeiten  $0.5^{r_i}$ . Bestehen große Unterschiede?

**Aufgabe 15 Monte-Carlo-Simulation****10 Pkte**

- a) Angenommen, Sie möchten den Wert des Integrals der Funktion  $f(x) = \sin(x)$  im Intervall  $[0, \pi]$  durch *Monte-Carlo-Simulation* bestimmen. Welche Schritte muss Ihre Simulation dafür enthalten? Welche Zufallszahlenverteilung benutzen Sie? (5P)
- b) Angenommen, Sie wollen dies durch die *Crude Monte-Carlo-Simulation* bestimmen. Wie lautet dann Ihre Simulationsvorschrift? Welche Verteilungsfunktion müssen Ihre Zufallszahlen dafür haben? (5P)

Lösungen:

- a) Schritte: # Zähle die rel. Anzahl der Zufallszahlen im Sinusbereich

Für  $N$  Iterationen tue

- Bilde Zufallszahlen  $x_1, x_2$  aus  $U(0, 1)$  (1P)
- Skalieren  $x_1$  auf  $[0, \pi]$  durch  $y = x_1 \pi$
- Teste: Ist  $x_2 < \sin(y)$ ? Wenn ja, erhöhe  $h = h + 1$  (3P)

Bilde  $I = h/N$  als Wert des Integrals (1P)

- b) Schritte:  $sum = 0$ ; # Bilde den Erwartungswert der Funktion

Für  $N$  Iterationen tue

- Bilde Zufallszahl  $x$  aus  $U(0, 1)$
- Skalieren auf  $[0, \pi]$  durch  $y = x \pi$  (1P)
- Bilde  $f = \sin(y)$
- Erhöhe  $sum = sum + f$  (3P)

Bilde  $I = \pi \text{ sum} / N$  als Wert des Integrals. (1P)

**Aufgabe 16 Simulationsauswertung****8 Pkte**

- a) Beim Vergleich der stochastischen Simulationen zweier Systeme ist für eine Beurteilung die Varianz der Ergebnisse oft zu hoch. Hier hilft die Methode der korrelierten Zufallsvariablen.
- Woraus besteht die Methode genau? (2P)
  - Beschreiben Sie dies an einem Beispiel. (1P)
  - Was sind die Grenzen der Methode? (1P)
- b) Bei der Beurteilung der stochastischen Simulationen eines Systems ist für eine Beurteilung die Varianz der Ergebnisse oft zu hoch. Hier hilft die Methode der antithetischen Zufallsvariablen.
- Woraus besteht die Methode genau? (2P)
  - Beschreiben Sie dies an einem Beispiel. (1P)
  - Was sind die Grenzen der Methode? (1P)

**Lösungen:**

- a) *Wenn die Zufallszahlen beider Systeme korreliert sind, verringert sich die Varianz der Differenz der Ergebnisse  $X_A - X_B$ .*
- Beispiel: Beide Systeme starten ihren Zufallszahlengenerator mit derselben Saat. Dann folgen alle Ereignisse beider Systeme korreliert und damit auch die Ergebnisse.*
- Grenzen: Haben beide Systeme keine vergleichbaren Einzelkomponenten, denen die Zufallszahlen zugeordnet sind, so nützt die Korrelation wenig, da sie nicht äquivalent benutzt wird.*
- b) *Bei einem einzigen System nützt es nichts, dieselben Zufallszahlen zu verwenden. Stattdessen kann man Paare von Simulationen bilden. Benutzt man in dem einen Simulationslauf eine Folge von Zufallszahlen, so kann man im anderen Lauf dazu negativ korrelierte Zufallszahlen benutzen, um die Kovarianz der Ergebnisdifferenz zu erhöhen und damit die Varianz der Ergebnisdifferenz beider Systeme zu verringern.*
- Beispiel: Ist  $u$  eine Zufallsvariable, so kann man  $1-u$  als antithetische Zufallsvariable im zweiten Lauf verwenden.*
- Grenzen: Ist ein alternativer Programmablauf möglich, so kann der Strom der Ereignisse auseinanderklaffen; die antithetischen Zufallsvariablen werden an der falschen, zum ersten Lauf nicht-korrespondierenden Stelle verwendet.*