

**Regressionsmodelle mit Anwendungen in der  
Versicherungs- und Finanzwirtschaft**

**Probeklausur Wintersemester 2017/2018 06.12.2018**

BITTE LESERLICH IN DRUCKBUCHSTABEN AUSFÜLLEN

Nachname: .....

Vorname: .....

Matrikelnummer: 

--	--	--	--	--	--	--	--

Studienfach: .....

**Unterschrift der/des Studierenden:**

Überprüfen Sie die Klausur auf Vollständigkeit, sie besteht aus 12 Seiten.

**Bemerkungen:**

Aufgabe	max. Pkt.	err. Pkt.
1	20	
2	20	
3	6	
4	16	
5	6	
6	10	
7	12	
8	10	
<b>Summe</b>	<b>100</b>	
<b>Bonus</b>	<b>6</b>	
<b>Note</b>		

### Aufgabe 1: Multiple Choice (20 Punkte)

Kreuzen Sie an, welche der nachfolgenden Aussagen wahr bzw. falsch sind. Beachten Sie:

Jede Aussage ist entweder wahr oder falsch.

Richtig gesetztes Kreuz = 2 Punkte

Kein gesetztes Kreuz = 1 Punkt

Falsch gesetztes Kreuz = 0 Punkte

	wahr	falsch
(a) Bei linksschiefen Verteilungen nimmt das dritte zentrierte theoretische Moment besonders große Werte an.	<input type="checkbox"/>	<input type="checkbox"/>
(b) Die Power-Transformation wird verwendet um rechtsschiefe Verteilungen so zu transformieren, damit sie annähernd symmetrisch sind.	<input type="checkbox"/>	<input type="checkbox"/>
(c) Sei $\mathbf{X} = (X_1, X_2, \dots, X_n) \sim \mathbb{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ und $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)$ eine Zerlegung von $\mathbf{X}$ in zwei Zufallsvektoren, dann ist $X_1 X_2$ ebenfalls multivariat Normalverteilt.	<input type="checkbox"/>	<input type="checkbox"/>
(d) Das adjustierte Bestimmtheitsmaß $R_a^2$ kann größere Werte annehmen als das Bestimmtheitsmaß $R^2$ .	<input type="checkbox"/>	<input type="checkbox"/>
(e) Das Lognormal-Verfahren ist ein auf dem verallgemeinerten-linearen Modell basierendes Verfahren.	<input type="checkbox"/>	<input type="checkbox"/>
(f) Jeder schwach stationäre Prozess weist eine über die Zeit konstante Varianz auf.	<input type="checkbox"/>	<input type="checkbox"/>
(g) Der Lag-Operator verschiebt den Zeitindex des stochastischen Prozesses um eine Einheit in die Zukunft.	<input type="checkbox"/>	<input type="checkbox"/>
(h) Ein White-Noise Prozess ist ein Spezialfall aus der Klasse der MA-Prozesse.	<input type="checkbox"/>	<input type="checkbox"/>
(i) Der Lagrange-Multiplikator-Test prüft bei einer vorgegebenen Zeitreihe das Vorhandensein von Heteroskedastizität.	<input type="checkbox"/>	<input type="checkbox"/>
(j) Bei der Parameterschätzung von $AR(p)$ -Prozessen liefern die Kleinste-Quadrate-, Momenten- und Maximum-Likelihood-Methode immer die selben Schätzungen.	<input type="checkbox"/>	<input type="checkbox"/>

---



e) Was versteht man unter einem saturierten Modell und was ist das besondere an den Schätzungen  $\hat{y}_i$  in diesem Modell? (2 Punkte)

f) Was würde mit dem KQ-Schätzer passieren, wenn man bei einer kategorialen Variablen keine Referenzklasse bilden würde und das Modell für jede Klasse eine Binäre Variable enthält. (2 Punkte)

g) Was ist ein stochastischer Prozess? (2 Punkte)

h) Nennen Sie eine Erweiterung für das GARCH-Modell. (2 Punkte)

i) Nennen Sie einen Test mit dem überprüft werden kann, ob ein ARMA-Prozess nicht schwach stationär ist. (2 Punkte)

j) Welche Bedingungen muss ein stochastischer Prozess  $(\varepsilon_t)_{t \in \mathbb{Z}}$  erfüllen, damit es sich bei dem Prozess um einen White-Noise Prozess handelt? (2 Punkte)

**Aufgabe 3: Multivariate Normalverteilung (6 Punkte)**

Der Zufallsvektor  $\mathbf{X} = (X_1, X_2, X_3)^T$  sei dreidimensional  $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ -verteilt mit

$$\boldsymbol{\mu} = \begin{pmatrix} 1 \\ 4 \\ 5 \end{pmatrix} \quad \text{und} \quad \boldsymbol{\Sigma} = \begin{pmatrix} 4 & 2 & -1 \\ 2 & 25 & -10 \\ -1 & -10 & 5 \end{pmatrix}$$

Bestimmen Sie die Verteilung von  $X_2$  gegeben  $x_1 = -2$  und  $x_3 = 7$ .

---

#### Aufgabe 4: Residuenanalyse (16 Punkte)

Es wird das klassische lineare Modell der Form

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

mit  $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{E})$  betrachtet. Für den Beobachtungsvektor  $\mathbf{y}$  und für die Designmatrix  $\mathbf{X}$  hat man

$$\mathbf{y} = \begin{pmatrix} 6 \\ 10 \\ 12 \\ 30 \end{pmatrix} \quad \text{und} \quad \mathbf{X} = \begin{pmatrix} 1 & 3 \\ 1 & 4 \\ 1 & 5 \\ 1 & 10 \end{pmatrix}$$

erhalten. Mit  $\mathbf{y}$  und  $\mathbf{X}$  erhält man für den Kleinste-Quadrate-Schätzer

$$\hat{\boldsymbol{\beta}} = \begin{pmatrix} \hat{\beta}_0 & \hat{\beta}_1 \end{pmatrix}^T = \begin{pmatrix} -\frac{124}{29} & \frac{99}{29} \end{pmatrix}^T$$

und für die hat-Matrix

$$\mathbf{H} = \frac{1}{116} \begin{pmatrix} 54 & 44 & 34 & -16 \\ 44 & 38 & 32 & 2 \\ 34 & 32 & 30 & 20 \\ -16 & 2 & 20 & 110 \end{pmatrix}.$$

- Berechnen Sie das standardisierte Residuum  $r_4$  und beurteilen Sie anhand einer gängigen Daumenregel, ob es sich um einen Ausreißer handelt. (7 Punkte)
  - Bestimmen Sie das studentisierte Residuum  $r_4^*$  und beurteilen Sie anhand einer gängigen Daumenregel, ob es sich bei der Beobachtung  $y_4$  um einen Ausreißer handelt. (9 Punkte)
-

### Aufgabe 5: Modellüberprüfung und Modellwahl (6 Punkte)

Gegeben ist ein Datensatz mit der Designmatrix

$$X = \begin{pmatrix} 1 & 4 \\ 1 & 5 \\ 1 & 10 \end{pmatrix}.$$

Bestimmen Sie für diesen Datensatz die Hat-Matrix  $H$ . Ist eine der Beobachtungen selbst-schätzend? Verwenden Sie hierfür eine Daumenregel und beurteilen sie, ob die Anwendung dieser Daumenregel in diesem Fall sinnvoll ist.

Hinweis:

$$(\mathbf{X}^T \mathbf{X})^{-1} = \frac{1}{62} \begin{pmatrix} 141 & -19 \\ -19 & 3 \end{pmatrix}$$

---

### Aufgabe 6: Verallgemeinerte lineare Modelle (10 Punkte)

Die Finanzabteilung eines großen Unternehmens verwendet ein Verallgemeinertes lineares Modell mit  $k = 3$  erklärenden Variablen. Einige der Experten schlagen eine Erweiterung des bisherigen Modells um 4 zusätzliche Variablen vor. Zum Vergleich der beiden Modelle wurde eine Stichprobe von  $n = 12$  Beobachtungen erhoben. Für das kleinere Modell  $M_1$  hat man einen Log-Likelihood Wert von

$$l\left(\widehat{\beta}_{M_1; y_1, \dots, y_n}\right) = -10,6$$

und für das größere Modell  $M_2$  einen Wert von

$$l\left(\widehat{\beta}_{M_2; y_1, \dots, y_n}\right) = -6,4$$

erhalten. Der Log-Likelihood Wert des saturierten Modells beträgt

$$l\left(\widehat{\beta}_{S; y_1, \dots, y_n}\right) = -2.$$

- a) Prüfen Sie bei einem Signifikanzniveau von  $\alpha = 0,1$ , ob die Aufnahme der zusätzlichen Variablen eine signifikante Verbesserung liefert. (3 Punkte)
  - b) Für welches der beiden Modelle würde sich das Unternehmen entscheiden, wenn es das Informationskriterium von Akaike(AIC) als Entscheidungskonzept heranziehen würde? (2 Punkte)
  - c) Für welches der beiden Modelle würde sich das Unternehmen entscheiden, wenn es das Bayesianische Informationskriterium(BIC) als Entscheidungskonzept heranziehen würde? (2 Punkte)
  - d) Angenommen, das Unternehmen überlegt weder  $M_1$  noch  $M_2$  zu verwenden und stattdessen das saturierte Modell S. Für welches der drei Modelle sollte sich das Unternehmen entscheiden, wenn das Informationskriterium von Akaike(AIC) bzw. das Bayesianische Informationskriterium(BIC) verwendet wird? (3 Punkte)
-

### Aufgabe 7: ARMA-Prozesse (12 Punkte)

Betrachtet wird der stochastische Prozess  $(X_t)_{t \in \mathbb{Z}}$  mit

$$X_t = X_{t-1} - 1,25X_{t-2} + \varepsilon_t - 0,25\varepsilon_{t-2}$$

mit  $\varepsilon_t \sim WN(0, \sigma^2)$ .

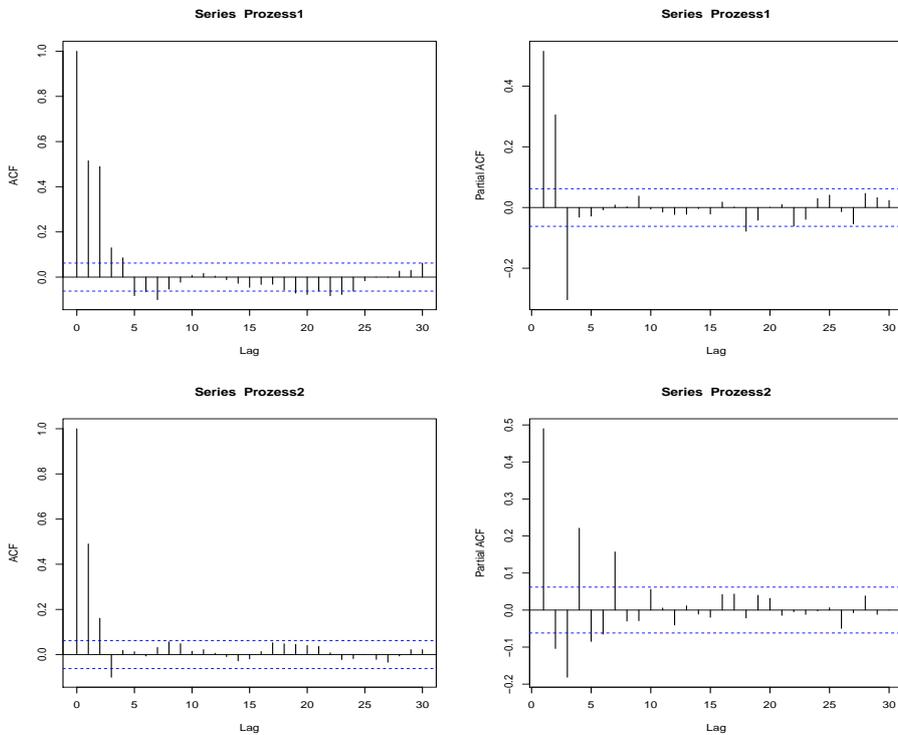
- Um was für einen Prozess handelt es sich bei  $(X_t)_{t \in \mathbb{Z}}$ ? (1 Punkt)
- Untersuchen Sie den Prozess auf Kausalität und Invertierbarkeit. (5 Punkte)

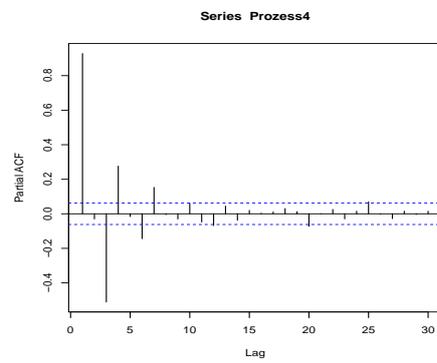
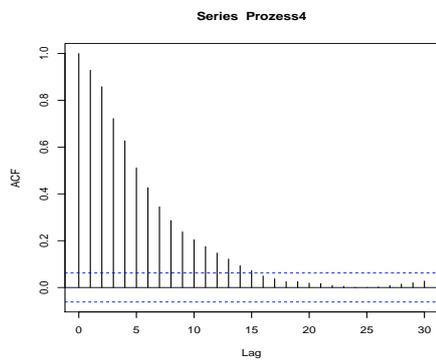
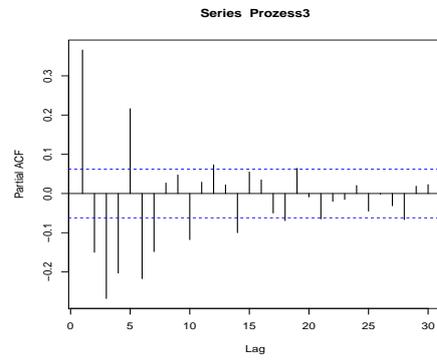
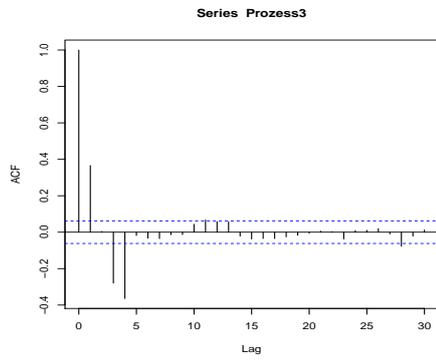
Betrachten Sie nun die stochastischen Prozesse  $(W_t)_{t \in \mathbb{Z}}$  und  $(Z_t)_{t \in \mathbb{Z}}$  mit

$$\begin{aligned} W_t &= \varepsilon_t + 0,8\varepsilon_{t-1} - 0,8\varepsilon_{t-4}, \\ Z_t &= 0,45Z_{t-1} + 0,4Z_{t-2} - 0,3Z_{t-3} + \varepsilon_t \end{aligned}$$

für  $t \in \mathbb{Z}$  und  $\varepsilon_t \sim WN(0, \sigma^2)$ .

- Ordnen Sie den Prozessen  $(W_t)_{t \in \mathbb{Z}}$  und  $(Z_t)_{t \in \mathbb{Z}}$  die passenden empirischen Autokorrelationsfunktionen und empirischen partiellen Autokorrelationsfunktionen zu und begründen Sie ihre Zuordnungen. (3 Punkte)
- Geben Sie zu den zwei nicht zugeordneten Prozessen die jeweils passende Prozessklasse an. (3 Punkte)





### Aufgabe 8: Schätzung und Prognose (10 Punkte)

Betrachtet wird der stochastische Prozess  $(X_t)_{t \in \mathbb{Z}}$  mit

$$X_t = \varepsilon_t + 0,3\varepsilon_{t-1} - 0,5\varepsilon_{t-2}$$

wobei  $\varepsilon \sim WN(0, \sigma^2)$  gilt.

- a) Geben Sie die Autokorrelationsfunktion  $\rho(h)$  des stochastischen Prozesses  $(X_t)_{t \in \mathbb{Z}}$  an. (4 Punkte)
- b) Bestimmen Sie den besten affin-linearen Prädiktor

$$\hat{X}_{T+h|T}^{opt} = a_0 + \sum_{t=1}^T a_{T+1-t} X_t$$

für  $T = 2$  und  $h = 2$ . (6 Punkte)

---