

Bitte für die Korrektur freilassen!					
Aufgabe	1 + Zusatz	2	3	4	Σ
Punkte	30 + 10*	30	30	30	120 + 10*
erzielt					

Wiederholungsklausur zur Vorlesung „Statistik II für Statistiker, Mathematiker und Informatiker“ (WS 06/07)

5. Oktober 2006

Hinweise:

- (a) Überprüfen Sie bitte, ob Ihre Angabe vollständig ist. Diese Angabe sollte (inklusive dieser Seite) 4 Seiten umfassen.
- (b) Die Bearbeitungszeit beträgt 120 Minuten. Es können 120 Punkte erreicht werden.
- (c) Als Hilfsmittel sind ausschließlich ein Taschenrechner (nicht programmierbar, ohne Plot-Funktion) sowie die Formelsammlung zugelassen.
- (d) Bei Unterschleif gilt die Klausur als nicht bestanden und es erfolgt eine Meldung an das Prüfungsamt.
- (e) Verwenden Sie für Ihre Notizen und Lösungen ausschließlich die Ihnen zur Verfügung gestellten Papierbögen.
- (f) Alle Ausarbeitungen müssen nachvollziehbar sein. Die Ergebnisse müssen klar ersichtlich sein (unterstreichen bzw. Antwortsatz). Es erfolgt eine detaillierte Bepunktung des Lösungsweges.
- (g) **Geben Sie bitte am Ende der Klausur alle von Ihnen zur Korrektur vorgesehenen Blätter ab und kennzeichnen Sie jedes abgegebene Blatt mit Name und Matrikelnummer.**

Bitte ausfüllen und unterschreiben!!!

Name (in Druckbuchstaben): _____

Matrikelnummer: _____ Studienfach: _____

Geburtstag: _____ Geburtsort: _____

Ich bestätige, dass ich obige Hinweise zur Kenntnis genommen habe und sie befolgen werde.
Ich bin mit einer Veröffentlichung meines Klausurergebnisses im Internet in der Form
<Matrikelnummer><Note> einverstanden. (Falls nicht, den letzten Satz bitte streichen!)

Unterschrift: _____

Viel Erfolg!

Aufgabe 1

13 + 17 + 10* = 30 + 10* Punkte

Sei $X_1, \dots, X_n$ eine einfache Zufallsstichprobe aus einer Grundgesamtheit mit folgender Wahrscheinlichkeitsfunktion

$$P_\theta(X = x) = \theta^x(1 - \theta)^{1-x}, \quad x \in \{0, 1\}, \quad 0 < \theta < 1.$$

- (a) Bestimmen Sie einen Schätzer für den unbekannten Parameter θ nach der Maximum-Likelihood-Methode $\hat{\theta}_1$ und einen Schätzer nach der Momentenmethode $\hat{\theta}_2$.
- (b) Ermitteln Sie für beide Schätzer, ob sie MSE-konsistent sind. Wie kann man das Konzept der MSE-Konsistenz anschaulich interpretieren?
- (c*) **Zusatzaufgabe:** Man möchte θ so schätzen, dass der Abstand zwischen Maximum-Likelihood-Schätzer $\hat{\theta}_1$ und wahren Parameterwert θ höchstens 0.05 beträgt. Wie groß muss der Stichprobenumfang mindestens sein, damit dieser Abstand mit einer Wahrscheinlichkeit von 0.9 eingehalten wird, dass also gilt

$$P(|\hat{\theta}_1 - \theta| \leq 0.05) \geq 0.9.$$

Betrachten Sie hierzu die Zufallsvariable $Y := \hat{\theta}_1 - \theta$ und verwenden Sie die Normalapproximation.

Aufgabe 2

4 + 3 + 20 + 3 = 30 Punkte

Eine Vertriebsgesellschaft vermutet, dass es hinsichtlich der Nachfrage nach bestimmten Wochenzeitschriften regionale Unterschiede gibt. Um diese Vermutung statistisch zu überprüfen, zieht sie aus der Menge der Käufer von Wochenzeitschriften zwei voneinander unabhängige einfache Zufallsstichproben, und zwar eine ($n_1 = 510$) aus der norddeutschen und eine ($n_2 = 450$) aus der süddeutschen Region.

Die Käufer in beiden Stichproben werden gefragt, welche Wochenzeitschrift sie bevorzugen (keine Mehrfachnennungen!). Das Ergebnis dieser Befragung zeigt folgende Tabelle:

Wochenzeitschrift	Norddeutschland	Süddeutschland
A	195	145
B	140	200
C	175	105
Summe	510	450

- (a) Formulieren Sie für die Vermutung der Betriebsgesellschaft „Es gibt regionale Unterschiede“ Null- und Alternativhypothese.
- (b) Begründen Sie, welcher Test - unter Berücksichtigung der erhobenen Daten - geeignet ist, um die Vermutung der Vertriebsgesellschaft zu überprüfen.
- (c) Führen Sie den Test auf einem Signifikanzniveau von $\alpha = 0.01$ durch und interpretieren Sie Ihr Ergebnis.
- (d) Kann man aus dem Ergebnis von Teilaufgabe (c) schließen, dass der Verkauf der Wochenzeitschriften unabhängig von der Region ist, in der sie verkauft werden? Begründen Sie Ihre Entscheidung.

Der Internetprovider Webspeed muss die Computerkapazitäten für das nächste Jahr planen. Hierfür muss insbesondere die Zahl der Internetkunden bei Webspeed prognostiziert werden. Dazu wird folgendes Modell verwendet:

$$y_t = x_{t1}\beta_1 + x_{t2}\beta_2 + x_{t3}\beta_3 + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma^2), \quad \text{Cov}(\epsilon_s, \epsilon_t) = 0 \quad \forall s \neq t,$$

wobei

- y_t die Anzahl der Internetkunden in Tsd.,
- x_{t1} eine Konstante mit Wert 1,
- x_{t2} den Preis pro Monat des größten Anbieters in Euro,
- x_{t3} den Preis pro Monat von Webspeed in Euro

angibt. Es liegen Daten für 24 Monate vor und

$$(X'X)^{-1} = \frac{1}{1000} \begin{bmatrix} 260.0 & 1.2 & 14.0 \\ 1.2 & 0.4 & -0.4 \\ 14.0 & -0.4 & 2.0 \end{bmatrix}, \quad X'y = 1000 \begin{bmatrix} 0.3 \\ 2.0 \\ -2.5 \end{bmatrix}.$$

- (a) Ein Firmenmitglied behauptet, dass sich bei einer Preissenkung beim größten Anbieter die Zahl der Internetkunden bei Webspeed substantiell verringern würde. Ist diese Aussage auf einem Signifikanzniveau von $\alpha = 0.05$ haltbar? Dabei sei Ihnen $\sigma^2 = 10$ gegeben.
- (b) Berechnen Sie ein Konfidenzintervall für β_3 zum Konfidenzniveau 0.95. Wie bisher sei $\sigma^2 = 10$ vorgegeben. Zu welchem Ergebnis kommen Sie hinsichtlich der Nullhypothese $H_0 : \beta_3 = -2$ auf Basis Ihres Konfidenzintervalls?
- (c) Begründen Sie **verbal**, ob sich Ihre Ergebnisse in (b) ändern würden, falls Sie die Fehlervarianz schätzen müssen und dabei einen Wert von $\hat{\sigma}^2 = 20$ erhalten.
- (d) Überprüfen Sie die Nullhypothese, dass eine Preissenkung des größten Anbieters um 1 Euro exakt den entgegengesetzten Effekt hat wie eine Preissenkung um 1 Euro durch Webspeed. Legen Sie dabei ein Signifikanzniveau von $\alpha = 0.05$ zugrunde. Die geschätzte Fehlervarianz ist wie bisher $\hat{\sigma}^2 = 20$.
- (e) Könnte sich Ihre Schlussfolgerung in (d) verändern, wenn Sie das Signifikanzniveau auf $\alpha = 0.01$ reduzieren? Wie lässt sich eine Reduktion des Signifikanzniveaus interpretieren?

Aufgabe 4

8 + 19 + 3 = 30 Punkte

Die Drahtseilbahn in einem alpinen Ausflugsgebiet hat zwischen 2002 und 2005 folgende Passagierzahlen (in Tsd.) verzeichnen können,

Jahr Periode	2002	2003		2004		2005	
	WS	SS	WS	SS	WS	SS	WS
Passagiere (in Tsd.)	197	191	213	211	234	232	256

wobei die Wintersaison (WS) vom 1.10 bis zum 31.3 des Folgejahres und die Sommersaison (SS) vom 1.4 bis zum 30.9 läuft.

- (a) Formulieren Sie ein additives Trend-Saison-Modell mit linearer Trendfunktion für den beschriebenen Sachverhalt. Erläutern Sie kurz, wie der Funktionstyp der Trendfunktion bestimmt werden kann.
- (b) Schätzen Sie die unbekannten Parameter des Trend-Saison-Modells, in dem Sie zunächst nur den Trend, ohne Berücksichtigung der Saisonkomponente, mittels KQ-Methode schätzen und anschließend die trendbereinigten Daten zur Schätzung der Saisonkomponente verwenden. Interpretieren Sie Ihre Schätzer für die Saisonkomponente.
- (c) Prognostizieren Sie die Passagierzahl für das Jahr 2006.