Übungsklausur zu „Statistik II für Statistiker, Mathematiker und Informatiker“ (SS 06)

Aufgabe 1	$4 + 10 + 16 = 30$ Punkte

Durch eine statistische Erhebung soll ein Konfidenzintervall der Länge 0.02 zum Niveau $1 - \alpha = 0.98$ für den Anteil p der Raucher in einer gegebenen Bevölkerungsgruppe ermittelt werden. Sei n die Anzahl der befragten Personen und X_n die Anzahl der Raucher darunter.

(a) Wie ist X_n exakt verteilt? Welche Annahmen müssen Sie dazu treffen?

(b) Geben Sie einen erwartungstreuen Schätzer für p an (Nachweis).

(c) Wieviele Personen der als sehr groß vorausgesetzten Bevölkerungsgruppe müssen befragt werden, wenn keinerlei Informationen über p vorliegen und das oben beschriebene Konfidenzintervall ermittelt werden soll? Berechnen Sie eine Approximation mit Hilfe der Tschebyschev-Ungleichung!

Aufgabe 2	$5 + 8 + 7 = 20$ Punkte

In einer Werkstatt werden Kraftfahrzeuge repariert. Die zufällige Reparaturzeit für die Behebung eines bestimmten Schadentyps kann als eine mit dem Parameter $\lambda > 0$ exponentialverteilte Zufallsgröße betrachtet werden. Die über einen konkreten Zeitraum unabhängig voneinander erfaßten Zeiten für n Reparaturen ergaben eine mittlere Reparaturdauer von 15 Zeiteinheiten.
Bestimmen Sie einen Schätzer für den unbekannten Parameter λ

(a) nach der Momentenmethode und

(b) nach der Maximum-Likelihood-Methode!

Untersuchen Sie die ermittelten Schätzer auf Erwartungstreue! Benutzen Sie dazu die *Jensen'sche Ungleichung*:
Sei X eine Zufallsvariable und $g : D \subseteq \mathbb{R} \to \mathbb{R}$ eine konvexe Funktion. Dann gilt

$$E[g(X)] \geq g(E[X]).$$

Ist g strikt konvex auf D und $Var(X) > 0$, so gilt sogar $E[g(X)] > g(E[X])$.
Klären Sie dabei zunächst die Anwendbarkeit der Jensen'schen Ungleichung ab!

Aufgabe 3	$10 + 2 + 8 = 20$ Punkte

In einer Fußball-Liga wurden für eine Saison die Torerfolge pro Spiel in folgender Häufigkeitstabelle zusammengefasst:

Torerfolge pro Spiel	0	1	2	3	4	> 4
Häufigkeit	18	24	56	63	61	78

Es soll mit einem statistischen Testverfahren überprüft werden, ob die Torerfolge pro Spiel einer Poissonverteilung mit Parameter $\lambda = 3.4$ folgen. Als Signifikanzniveau sei $\alpha = 0.1$ vorgegeben.

(a) Formulieren Sie die obige Aufgabenstellung als statistisches Testproblem! Benennen Sie den Test, der zur Lösung adäquat erscheint!

(b) Geben Sie den Ablehnungsbereich für diesen Test explizit an!

(c) Führen Sie den Test als Signifikanztest zum oben angegebenen Niveau α durch! Wie lautet die Testentscheidung?

Aufgabe 4	$10 + 10 + 10 = 30$ Punkte

Für die Nachfragefunktion eines Haushaltes nach Butter wurde folgendes lineares Modell aufgestellt

$$y_t = x_{t1}\beta_1 + x_{t2}\beta_2 + x_{t3}\beta_3 + \epsilon_t,$$

dabei bezeichne y_t die logarithmierte Nachfrage nach Butter im Monat t, x_{t2} den logarithmierten Preis (in Euro) von Butter im Monat t und x_{t3} den logarithmierten Preis von Margarine (in Euro), x_{t1} nimmt in allen Perioden den Wert eins an. Folgende Zwischenergebnisse liegen vor

$$(X'X)^{-1} = \begin{bmatrix} 0.05 & 0 & 0 \\ 0 & 0.1 & -0.2 \\ 0 & -0.2 & 0.5 \end{bmatrix}, \; X'y = \begin{bmatrix} 10 \\ -40 \\ -14 \end{bmatrix}, \; \sigma^2 = 4, \; T = 20.$$

Beachten Sie im folgenden, dass σ^2 bekannt ist.

(a) Schätzen Sie die Parameter β_1, β_2 und β_3 mit Hilfe der KQ-Methode. Welche Annahmen bezüglich der ϵ_t müssen Sie dazu treffen? Interpretieren Sie die geschätzten Parameter $\hat{\beta}_2$ und $\hat{\beta}_3$.

(b) Testen Sie die Hypothese $H_0 : \beta_1 = 0$ gegen $H_1 : \beta_1 \neq 0$ auf einem Signifikanzniveau von $\alpha = 0.05$ anhand eines geeigneten Tests. Bei welchem Signifikanzniveau hätten Sie die Nullhypothese gerade noch nicht verwerfen können?

(c) Testen Sie die Hypothese $H_0 : \beta_2 = -\beta_3$ gegen $H_1 : \beta_2 \neq -\beta_3$ auf einem Signifikanzniveau von $\alpha = 0.05$.

Aufgabe 5	$5 + 10 + 5 = 20$ Punkte

Die Produktion von Mikroprozessoren entwickelte sich in einem Unternehmen der elektronischen Industrie in folgenden Stückzahlen (in Tsd.)

Jahr	2001	2002	2003	2004	2005
Stückzahl g_t	110	130	155	180	220

(a) Wie hat sich die Produktion von Mikroprozessoren im Mittel jährlich im Zeitraum 2001-2005 entwickelt?

(b) Bestimmen Sie eine Funktion mit exponentiellem Trend (d.h. $g_t = \beta_0\beta_1^t$, $t = 1$ entspricht 2001) für die langfristige Entwicklung der Produktion von Mikroprozessoren. Schätzen Sie die unbekannten Parameter und interpretieren Sie Ihr Ergebnis. Hinweis:

$$\sum_{t=1}^{5} t^2 = 55, \quad \sum_{t=1}^{5} \ln g_t = 25.1980, \quad \sum_{t=1}^{5} \ln g_t \cdot t = 77.3058.$$

(c) Erstellen Sie auf der Grundlage Ihrer Berechnungen in (a) und (b) je eine Prognose für die Produktion 2007 und vergleichen Sie die Ergebnisse.