Mathematik IV f. Elektrotechnik
Mathematik III f. Informatik
1. Übungsblatt

TECHNISCHE UNIVERSITÄT DARMSTADT

Fachbereich Mathematik
Prof. Dr. Stefan Ulbrich
Dipl.-Math. Hannes Meinlschmidt
Dipl.-Math. Sebastian Pfaff

SoSe 2013
23. April 2012

Gruppenübung

Aufgabe G1 (Zweidimensionale Messreihen)
Um neue Aufgaben zur Statistik zu erhalten, wurde in der AG Optimierung eine Umfrage unter den Doktoranden zum Thema "Autofahren" durchgeführt. Dabei wurden unter anderem die Höchstgeschwindigkeit und der Kraftstoffverbrauch der jeweiligen Autos ermittelt und in nachfolgender Tabelle aufgetragen. *(Wir nehmen hier an, dass Optimierer von Natur aus an einer Minimierung der Tankkosten interessiert sind und daher keine "Spritfresser" fahren.)*

i	1	2	3	4	5	6
Höchstgeschwindigkeit (in km/h): x_i	174	170	172	167	170	175
Verbrauch (in l/100 km): y_i	6.9	6.4	6.8	5.7	6.3	7.1

a) Stellen Sie die Umfrageergebnisse in einem Punktediagramm dar.

b) Ausgerechnet der Mitarbeiter, der die Umfrage durchgeführt hat, konnte sich nicht mehr an den Verbauch seines Autos erinnern. Um trotzdem einen Wert angeben zu können, kommt er auf die Idee, einen plausiblen Verbrauchswert durch lineare Regression aus der Höchstgeschwindigkeit zu ermitteln.

   Berechnen Sie die empirischen Streuungen, die empirische Kovarianz und den empirischen Korrelationskoeffizienten dieser zweidimensionalen Messreihe. Ist die Annahme eines linearen Zusammenhangs zwischen Höchstgeschwindigkeit und Verbrauch hier gerechtfertigt? Warum?

c) Wir nehmen nun an, ein linearer Zusammenhang sei begründet. Berechnen Sie die Regressionsgerade zur Vorhersage des Verbrauchs anhand der Höchstgeschwindigkeit eines Autos und zeichnen Sie diese in das Punktediagramm.

d) Bestimmen Sie einen Vorhersagewert für den Verbrauch eines Autos mit einer Höchstgeschwindigkeit von 168 km/h.

e) Ein weiterer Mitarbeiter fährt ein älteres Modell. Sein Auto fährt nur maximal 155 km/h und schluckt dabei 8.4 Liter auf 100 km. Betrachten Sie nun die um dieses Wertepaar erweiterte Messreihe. Beurteilen Sie anhand geeigneter statistischer Maßzahlen, ob ein linearer Zusammenhang zwischen Höchstgeschwindigkeit und Verbrauch gerechtfertigt ist.

Aufgabe G2 (Verteilungsfunktion, Maßzahlen)
In einer Automobilfabrik wurden bei 20 Fahrzeugen eines Typs folgende Höchstgeschwindigkeiten gemessen:

141, 142, 143, 144, 144, 144, 145, 145, 146, 147,
147, 148, 150, 150, 151, 151, 152, 138, 140, 141.

(a) Zeichnen Sie die empirische Verteilungsfunktion der Stichprobe.

(b) Zeichnen Sie ein Histogramm mit Klasseneinteilung

$$(v-1, v+1], \qquad v = 138, 140, 142, 144, 146, 148, 150, 152.$$

(c) Berechnen Sie den Median, das arithmetische Mittel und das p-Quantil für $p = 0.25$ und $p = 0.75$.

(d) Berechnen Sie die empirische Varianz und die empirische Streuung für die **verkleinerte Stichprobe**, bestehend aus den ersten 7 Messwerten.

(e) Angenommen bei der Übertragung der Messdaten ist ein Fehler passiert und es wurde bei einer der Messungen statt 145 km/h 345 km/h übertragen. Welche Auswirkung hat das auf die in Aufgabe (c) und (d) berechneten Maßzahlen?

Aufgabe G3 (Kombinatorik)
Ein Skatspiel besteht aus 32 Karten, vier davon heißen Bauern (Buben). Nach dem Mischen der Karten erhalten die drei Spieler (Alex, Bodo und Carl) jeweils zehn Karten. Die verbleibenden zwei Karten bilden den sogenannten Skat. Berechnen Sie jeweils die Wahrscheinlichkeit der folgenden Ereignisse:

A: Mindestens ein Bauer befindet sich im Skat.

B: Carl hat genau einen Bauer.

C: Ein Spieler hat genau drei Bauern.

D: Jeder Spieler besitzt mindestens einen Bauern.

Hausübung

Aufgabe H1 (Zufallsexperiment und Wahrscheinlichkeit)
Eine Firma möchte ihren Kunden den Zugriff auf ihre persönlichen Daten über das Internet ermöglichen. Für den Zugang müssen die Kundennummer und eine PIN, bestehend aus den Ziffern 0-9, eingegeben werden. Nachdem die PIN dreimal hintereinander falsch eingegeben wurde, wird der Zugang gesperrt und der Kunde informiert.

(a) Angenommen, einem Hacker sei eine Kundennummer bekannt und er probiert nun zufällig gewählte PINs aus. Jede PIN wird mit der gleichen Wahrscheinlichkeit gewählt und unser Hacker ist so schlau, dass er keine PIN doppelt versucht. Wieviele Stellen muss die PIN mindestens haben, damit die Wahrscheinlichkeit, dass der Hackerangriff unbemerkt bleibt, höchstens 10^{-6} ist?

(b) Wieviele Stellen wären nötig, wenn anstelle der PIN ein Passwort verwendet würde? Dabei soll das Passwort aus Buchstaben (ohne Umlaute) und Ziffern bestehen, wobei Groß- und Kleinschreibung nicht beachtet wird.

Aufgabe H2 (Verteilungsfunktion, Histogramm)
Am Frankfurter Flughafen wurde im letzten Februar täglich um 8:00 Uhr die Windgeschwindigkeit gemessen. Die Messungen ergaben:

20.1	18.4	8.0	8.0	8.6	8.6	9.2	9.7	9.7	9.7	9.7	10.9	11.5	11.5	11.5
12	12	12	12.6	13.2	13.8	14.3	14.9	14.9	14.9	16.6	16.6	6.9	7.4	

(a) Skizzieren Sie die empirische Verteilungsfunktion der angegebenen Messreihe und zeichnen Sie ein Histogramm mit folgender Klasseneinteilung: $(5,7]$, $(7,9]$, $(9,11]$, $\cdots$ $(19,21]$.

(b) (Für Fleißige:) Berechnen Sie das arithmetische Mittel, den Median und die empirische Varianz.

Aufgabe H3 (Zufallsexperimente und Wahrscheinlichkeit)
Sepp, Hinz und Kunz schauen zusammen ein Fussball-WM-Spiel. Für den Fall, dass das Bier nicht reichen sollte, haben sie das folgende Verfahren verabredet um denjenigen zu ermitteln, der Nachschub besorgen muss:

Zunächst werfen Hinz und Kunz eine Münze. Zeigt diese Zahl, scheidet Hinz aus bei Kopf Kunz. Dann werfen Sepp und der Nichtausgeschiedene eine Münze. Ist das Ergebnis Zahl, dann muss Sepp das Bier holen ansonsten der Nichtausgeschiedene.

(a) Wählen Sie eine Bezeichung für die Ergebnisse des im Verfahren durchgeführten Zufallsexperiments und geben Sie die Ergebnismenge an.

(b) Geben Sie dann die Ereignisse

A_1 : Sepp muss Bier holen.
A_2 : Hinz muss Bier holen.
A_3 : Kunz muss Bier holen.

mit Hilfe der in (a) gewählten Bezeichnung an. Welche dieser Ereignisse sind Elementarereignisse?

(c) Berechnen Sie die Wahrscheinlichkeiten $P(A_1)$, $P(A_2)$ und $P(A_3)$. Ist das Verfahren gerecht?

Aufgabe H4 (Zweidimensionale Messreihen)
Eine Strecke wurde an 15 verschiedenen Tagen und zu unterschiedlichen Tageszeiten mit dem gleichen Fahrzeug abgefahren. Dabei wurde jeweils die Durchschnittsgeschwindigkeit v_i (in km/h) und die Verkehrsdichte d_i (in Anzahl Fahrzeuge pro km) ermittelt. Dies ergab die folgenden Daten:

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
v_i	29	40	42	47	50	56	57	60	60	62	63	67	69	74	82
d_i	40	37	34	30	25	19	23	21	13	16	21	13	16	11	7

(a) Stellen Sie die beobachteten Daten zunächst in einem Punktediagramm graphisch dar und berechnen Sie dann den empirischen Korrelationskoeffizienten.
Hinweis für Eilige: Für die Empirischen Streuungen gilt $s_v \approx 13.8471$ bzw. $s_d \approx 9.845$.

(b) Die Ergebnisse von Teil (a) legen nahe, dass der Zusammenhang zwischen Durchschnittsgeschwindigkeit v und Verkehrsdichte d durch eine Gerade beschrieben werden kann. Bestimmen Sie die Regressionsgerade

$$d = \hat{a}v + \hat{b}$$

zur Messreihe (v_i, d_i), $i = 1, \ldots, 15$, und zeichnen Sie diese in das Punktediagramm ein.

(c) Da die Durchschnittsgeschwindigkeit leichter zu ermitteln ist als die Verkehrsdichte, sollen mit Hilfe der in Teil (b) berechneten Regressionsgerade Schätzwerte für die Verkehrsdichte bestimmt werden. Geben Sie den Schätzwert für die Verkehrsdichte bei einer Durchschnittsgeschwindigkeit von 55 km/h an.