

Name:	Vorname:	Matrikel-Nr.:

## Klausur zur Vorlesung Statistik I für Bachelor-Studenten, SS 2009

### Aufgabe 1

**(30 Punkte)**

Um den Verkehr zu lenken, führten seit wenigen Jahren verschiedene Städte wie London, Stockholm oder San Francisco ein sog. *Mobility Pricing* ein.

Ziele dieser Maßnahmen können z.B. Verkehrsvermeidung oder eine räumliche bzw. zeitliche Verkehrsverlagerung (Entzerrung der Spitzenstunden) sein. Auch eine Verschiebung des Modal-Splits (hin zu ÖV) ist denkbar. Schließlich kann die Motivation rein finanzieller Natur sein (mehr Geld in die Stadtkasse). Die Gebühren können sowohl für den fließenden als auch für den ruhenden Verkehr (Parkgebühren) erhoben werden, und zwar für PKW, LKW, Taxis, ÖV usw. Ferner hängt die Höhe der Gebühren i.A. von Nutzungszeit und -raum, der Schadstoffklasse, der Nutzergruppe und des Besetzungsgrads ab.

Die Erhebung der Gebühren kann manuell oder automatisch (Durchfahren von Kontrollpunkten mit Nummernschilderkennung oder GPS gestützt) erfolgen. Je nach System sind unterschiedlich viele Teilnehmer betroffen. Die Gesamteinnahmen können den Verkehrssystemen zugute kommen oder aber nicht zweckgebunden sein.

- (a) Geben Sie die statistische Gesamtheit und den Merkmalsträger an. Grenzen Sie die statistische Gesamtheit ab.
- (b) Geben Sie mindestens drei häufbare Merkmale und die entsprechenden häufbaren Merkmalsausprägungen an.
- (c) Begründen Sie, warum häufbare Merkmale immer nominalskaliert sein müssen
- (d) Geben Sie mindestens drei Bestandsmassen und eine Bewegungsmasse an.
- (e) Geben Sie je ein dichotomes (binäres) häufbares und ein dichotomes nichthäufbares Merkmal an.
- (f) Geben Sie schließlich noch
  - ein absolutskaliertes Merkmal,
  - drei weitere kardinalskalierte Merkmale
  - und ein ordinalskaliertes Merkmal
 an.

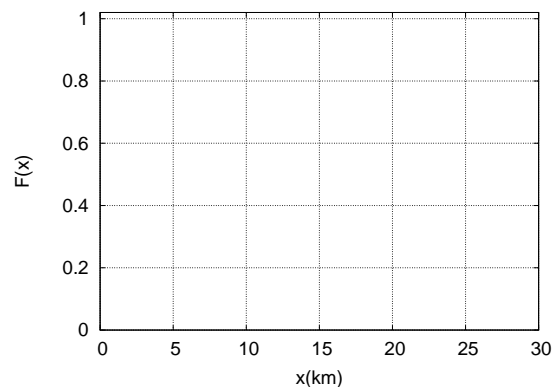
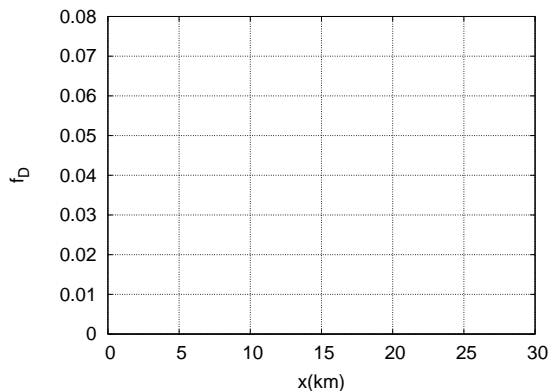
Name:	Vorname:	Matrikel-Nr.:
-------	----------	---------------

**Aufgabe 2****(55 Punkte)**

Im Rahmen einer Mobilitätsbefragung (Umfang: 1000 Personen) wurden für einen Stichtag die Reiseweiten aller durchgeführten Wege erfasst:

Reiseweitenklasse (km)	0–1	1–2	2–3	3–5	5–10	10–25	25–50
Häufigkeiten	32	165	247	488	770	1365	670

- (a) Wieviel Wege wurden insgesamt erfasst? Welcher Mobilitätskennziffer (Wege pro Person und pro Tag) entspricht dies?
- (b) Bestimmen Sie die relativen Häufigkeiten, Summenhäufigkeiten und die empirische Dichtefunktion für jede Klasse.
- (c) Zeichnen Sie die Dichte- und Verteilungsfunktion in die beiden folgenden Diagramme dieses Aufgabenblatts ein. Warum ist die Häufigkeitsdichte der kürzesten Wege eher gering, obwohl diese doch sicherlich die attraktivsten sind?



- (d) Bestimmen Sie das arithmetische Mittel, den Median und den Modus (ohne Feinberechnung). Zeichnen Sie in die Grafik das 60. Perzentil ein.
- (e) Als sog. “Pendlerpauschale” wurde in den Jahren 2007 und 2008 ein Satz von 30 Cent/km für den einfachen Arbeitsweg ab Beginn des 21. Kilometers auf die Steuer angerechnet. Wie viele Prozent kommen in den Genuss dieser Pauschale falls sich die Arbeitswege statistisch nicht von den anderen Wegen unterscheiden?
- (f) Ab 2009 (und vor 2007) gibt es die Pendlerpauschale für Arbeitswege beliebiger Länge. Es wurde aber auch über eine komplette Abschaffung der Pendlerpauschale nachgedacht. Wie hoch sind die steuerlichen Mindereinnahmen pro Erwerbstätigen und Jahr durch (i) die Pendlerpauschale ab Beginn des 21. Kilometers, (ii) für alle Wege im Vergleich zu ihrer Abschaffung? Nehmen Sie dabei 240 Arbeitstage und einen mittleren Grenzsteuersatz von 30% an.

Name:	Vorname:	Matrikel-Nr.:
-------	----------	---------------

**Aufgabe 3****(35 Punkte)**

Aus einer Mobilitätsuntersuchung wurde der Modal Split (Verkehrsmittelwahl) in Abhängigkeit der Weglänge festgestellt. Insbesondere ergaben sich die Häufigkeiten der Wahl "zu Fuß" in Abhängigkeit der Entfernungsklassen gemäß folgender Tabelle:

Reiseweitenklasse (km)	0-0.5	0.5-1	1-1.5	1.5-2	2-3	3-5
Häufigkeiten Wege	30	50	120	140	50	40
Häufigkeiten Wege zu Fuß	28	44	64	62	11	1

- (a) Bestimmen Sie zunächst für jede Entfernungsklasse die Anteilswerte der Fußwege.
- (b) Gemäß eines Verkehrsmittelwahlmodells ist der Fußweganteil  $y$  in Abhängigkeit der Weglänge  $x$  (in km) gegeben durch

$$\hat{y}(x) = \frac{1}{1 + e^{a+bx}}.$$

Zeigen Sie, dass durch die Transformation

$$y \rightarrow v = \ln\left(\frac{1}{y} - 1\right)$$

dass Modell in den transformierten Variablen linear ist.

- (c) Wurde die unabhängige oder die abhängige Variable transformiert? Ist demnach die Regression des transformierten Modells äquivalent zu einer direkten Regression des nichtlinearen Modells oder nicht?
- (d) Führen Sie die lineare Regression in den transformierten Variablen durch.  
*Hinweis:* Die Kovarianz ist hier, ebenso wie die Mittelwerte und Varianzen, eine gewichtete Einfachsumme.
- (e) Das Modell kann als Grenzfall einer logistischen Regression aufgefasst werden (vgl. die Vorlesungsunterlagen):

$$\hat{y}(x) = \frac{y_s}{1 + e^{-\tilde{b}(x-x_0)}}$$

Geben Sie allgemein die Bedeutung der drei Parameter  $y_s$ ,  $\tilde{b}$  und  $x_0$  und für die vorliegende Situation die Zahlenwerte an. Gehen Sie dabei vom Regressionsergebnis  $a = -2.61$  und  $b = 1.61$  aus. Was würde ein negativer Wert von  $\tilde{b}$  ausdrücken?