Grundbegriffe der medizinischen Biometrie und Epidemiologie (Sommer 2013)
Übungsblatt 2 für den 16.5.2013

1 Aufgabe: Deskriptive Statistik

1. Die folgende Tabelle zeigt Nettokaltmieten pro m^2 für 1- und 2-Raum Wohnungen.

	1	2	3	4	5	6	7	8	9	10
1-Raum	8.70	11.28	13.24	8.37	12.16	11.04	10.47	11.16	4.28	19.54
2-Raum	3.36	18.35	5.19	8.35	13.10	15.65	4.29	11.36	9.09	

Tabelle 1: Nettokaltmieten für Ein- und Zweiraumwohnungen

   (a) Bestimmen Sie für beide Gruppen den Mittelwert und die Varianz.
   (b) Berechnen Sie zudem Median und den Interquartilsabstand (IQR).
   (c) Interpretieren Sie ihre Ergebnisse aus den ersten beiden Aufgaben. Gehen Sie dabei besonders auf das Verhältnis von Mittelwert und Median sowie von IQR und Varianz ein.

2. Die Stadt Leipzig will die Altersstruktur ihrer Einwohner untersuchen. In der folgenden Tabelle finden Sie die Daten aus dem Jahr 2008.

Alter	Absolute H'keit	Relative H'keit	Kumulierte relative H'keit
0 bis 10	40529		
10 bis 20	33585		
20 bis 30	85450		
30 bis 40	71229		
40 bis 50	77920		
50 bis 60	65970		
60 bis 70	62639		
70 und älter	78147		
Gesamt	515469		

Tabelle 2: Einwohner nach Altersgruppen: Stand 31.12.2008

   (a) Berechnen Sie die relativen und kumulierten relativen Häufigkeiten.
   (b) Zeichnen Sie das Histogramm und die empirische Verteilungsfunktion.

2 Aufgabe: Labortest

Ein Labortest zur Erkennung einer Krankheit K, an der 5% der Gesamtbevölkerung leiden, besitze die folgenden Eigenschaften:

- Hat eine Person die Krankheit K, so erkennt dies der Test mit einer Wahrscheinlichkeit von 95% (*True Positiv Value*).
- Hat eine Person die Krankheit K nicht, so ist das Testergebnis mit einer Wahrscheinlichkeit von 16% dennoch positiv (*False Positiv Value*).

Die Ereignismenge für die tatsächlicher Erkrankung ist gegeben als $\{K, \bar{K}\}$ und für das Testergebnis als $\{T, \bar{T}\}$. Man berechne die Wahrscheinlichkeit, dass eine zufällig gewählte Person aus der Gesamtbevölkerung:

1. An Krankheit K leidet (Ereignis: K), obwohl der Test keine Erkrankung erkennt (Ereignis: $\bar{T}$).
2. Nicht an Krankheit K leidet (Ereignis: $\bar{K}$), obwohl der Test eine Erkrankung erkennt (Ereignis: T).
3. Nicht an Krankheit K leidet (Ereignis: $\bar{K}$), wenn der Test auch Entwarnung bezüglich der Erkrankung gibt (Ereignis: $\bar{T}$).

3 Aufgabe: Gleichverteilung

Eine stetige Zufallsvariable X heißt gleichverteilt auf dem Intervall $[a, b]$, kurz $X \sim U(a, b)$ falls für ihre Dichte gilt:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{für } a \leq x \leq b; \\ 0 & \text{sonst.} \end{cases}$$

(a) Weisen Sie nach, dass die Verteilungsfunktion einer gleichverteilten Zufallsvariable X wie folgt gegeben ist:

$$F(x) = \begin{cases} 0 & \text{für } x < a; \\ \frac{x-a}{b-a} & \text{für } a \leq x \leq b; \\ 1 & \text{für } x > b. \end{cases}$$

(b) Zeigen Sie, dass $E(X) = (b + a)/2$ und $\mathrm{Var}(X) = (b - a)^2/12$ gilt.

Hinweis: Benutzen Sie zur Berechnung der Varianz den Zusammenhang $\mathrm{Var}(X) = E(X^2) - E(X)^2$.

(c) Gegeben sei eine unabhängige Stichprobe $x_1, x_2, \ldots, x_n$ einer Gleichverteilung auf dem Intervall $[-a, a]$ $(a > 0)$. Weisen Sie nach, dass $T = \frac{3}{n}(X_1^2 + \ldots + X_n^2)$ ein unverzerrter Schätzer für a^2 ist.