Prof. Dr. Antony Unwin, Dr. Ali Ünlü
Lehrstuhl für Rechnerorientierte Statistik und Datenanalyse
Institut für Mathematik
Universität Augsburg
http://stats.math.uni-augsburg.de/

Stochastik I – Wahrscheinlichkeitstheorie (WS 2008/09)
Übungsblatt 6

Abgabe: Mittwoch 26. November 2008, bis spätestens 12.00 Uhr; Briefkasten: Wahrscheinlichkeitstheorie

1. Sei X eine stetige Zufallsvariable mit der Dichte

$$f(x) = \frac{a}{1+x^2} \qquad (x \text{ reelle Zahl}).$$

   Man berechne

   (a) den Koeffizienten a;

   (b) die Verteilungsfunktion von X;

   (c) die Wahrscheinlichkeit dafür, dass X einen Wert im Intervall $[0, 1]$ annimmt.

2. Radargeräte sind mit einem zufälligen "Rauschpegel" behaftet, dessen Verteilung die Dichte

$$f(x) = 2x \cdot \mathrm{e}^{-x^2} \text{ für } x > 0$$

   besitzt. Falls der Rauschpegel größer als eine gewisse Konstante d ist, so werden Geräte in Kategorie 2 eingestuft, andernfalls in Kategorie 1.

   (a) Wie groß ist die Wahrscheinlichkeit, dass ein zufällig ausgewähltes Gerät der Kategorie 1 einen Rauschpegel kleiner als $\frac{2}{3}d$ hat?

   (b) Bestimmen Sie eine Integralgleichung für d, so dass der durchschnittliche Rauschpegel in der Kategorie 1 um d Einheiten kleiner ist als der durchschnittliche Rauschpegel in Kategorie 2.

3. Der Datensatz in nachfolgender Tabelle beschreibt Blutplasma-Nikotin-Messungen (in Nanogramm [ng] pro Milliliter [ml]) an 55 Rauchern, in einer Studie über das Rauchen von Zigaretten [von D.J. Hand, The Open University]:

123	311	242	474	375	449	419	185	33	564	256	242
312	179	456	232	389	429	309	269	274	274	157	348
384	274	179	306	260	346	106	468	597	233	304	448
182	527	155	347	74	471	260	213	346	304	256	233
227	607	464	469	209	314	456					

   (a) Laden Sie den Datensatz `Nikotin` (abrufbar von Internetseite der Lehrveranstaltung) in `R`.

      i. Studieren Sie die `R`-Hilfeseite zur Funktion `hist()` und erstellen Sie ein Histogramm (in relativen Häufigkeiten) der Messungen.

      ii. Bestimmen Sie den Mittelwert und die Varianz der Messungen. Studieren Sie die `R`-Hilfeseite zur Funktion `dnorm()` und plotten Sie in das Histogramm die Dichte einer Normalverteilung mit den letzteren Größen als Erwartungswert und Varianz.

iii. Beurteilen Sie den Plot im Hinblick auf Anpassung. Wie gut beschreibt die Normalverteilung die Daten?

(b) Plotten Sie die Dichte der Verteilung aus vorheriger Teilaufgabe (a) und tragen Sie in den Plot die Stellen $\mu + k\sigma$ für $k = -3, -2, -1, 0, 1, 2, 3$ ein. Plotten Sie auch die zugehörige Verteilungsfunktion.

(c) Bestimmen Sie die Wahrscheinlichkeit für eine Nikotin-Messung zwischen $300\ ng/ml$ und $500\ ng/ml$, sowohl in `R` anhand der Funktion `pnorm()` als auch per Hand anhand einer Tabelle über die Standardnormalverteilung. Wieviele Standardabweichungen liegt der Nikotin-Wert $0\ ng/ml$ vom Erwartungswert entfernt? Wie groß ist daher die Wahrscheinlichkeit, eine negative Nikotin-Messung zu erhalten? Macht das Sinn und wie interpretieren Sie dies?

(d) Wie groß ist die Wahrscheinlichkeit, dass bei 20 zufällig getesteten Rauchern höchstens einer einen Nikotin-Wert größer als $500\ ng/ml$ aufweist? Vergleichen Sie das Ergebnis mit den Daten.

4. Ein Assistent wohnt in M und arbeitet in H. Er beginnt um 8:30 Uhr morgens mit der Arbeit. Er nimmt jeden Morgen den Zug von M nach A, der in A laut Fahrplan um 8:10 Uhr ankommt. Dort fährt alle 15 min. eine Tram nach H ab. Die Tram, die um 8:15 Uhr in A startet, trifft laut Fahrplan um 8:26 Uhr in H ein. Die Verspätung des Zuges von M nach A beträgt durchschnittlich 4 min., mit einer Standardabweichung von 2 min. Die Tram von A nach H fährt immer pünktlich ab, erreicht H aber mit durchschnittlich 2 min. Verspätung und einer Standardabweichung von 4 min.

   Behandeln Sie die nachfolgenden Teilaufgaben per Hand und als Vergleich auch in `R`.

   (a) Wie groß ist die Wahrscheinlichkeit, dass der Assistent nach 8:30 Uhr in H ankommt?

   (b) Ein Professor startet in A mit seinem Auto um 8:15 Uhr. Seine Fahrzeit nach H beträgt durchschnittlich 12 min., mit einer Standardabweichung von 3 min. Wie groß ist die Wahrscheinlichkeit, dass der Professor nach 8:30 Uhr in H ankommt?

   (c) Berechnen Sie die Wahrscheinlichkeit dafür, dass der Professor und Assistent beide nach 8:30 Uhr in H ankommen. Welche Annahmen haben Sie zur Berechnung der Wahrscheinlichkeiten getroffen?

5. Studieren Sie die `R`-Hilfeseiten zu den Funktionen `par()` und `rnorm()` und behandeln Sie in `R` die nachfolgenden Teilaufgaben.

   (a) Simulieren Sie jeweils $100, 1\,000, 10\,000, 100\,000$ und $1\,000\,000$ Zufallszahlen, die normalverteilt sind mit den Parametern $\mu = 5$ und $\sigma^2 = 4$, und wenden Sie in jedem dieser fünf Fälle die Verteilungsfunktion der Normalverteilung $N(5, 4)$ auf die generierten Zufallszahlen an.

   (b) Plotten Sie, in einem einzigen Graphikfenster, jeweils Histogramme geeigneter Klassenbreite der nach Anwendung der Verteilungsfunktion erhaltenen Zahlenwerte (in relativen Häufigkeiten). Zeichnen Sie in die Histogramme die Dichte einer Gleichverteilung im Intervall $[0, 1]$ in Farbe ein. Vergleichen und interpretieren Sie die Ergebnisse.

   (c) Beweisen Sie ganz allgemein: Ist Y eine Zufallsvariable mit der stetigen, streng monotonen Verteilungsfunktion G, so besitzt die Zufallsvariable $G(Y)$ eine Gleichverteilung im Intervall $[0, 1]$. Sehen Sie diese Aussage in der obigen Simulationsstudie bestätigt (beschreiben Sie)?