

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 II — Statistik I (Statistik und Datenanalyse, SS 2009)

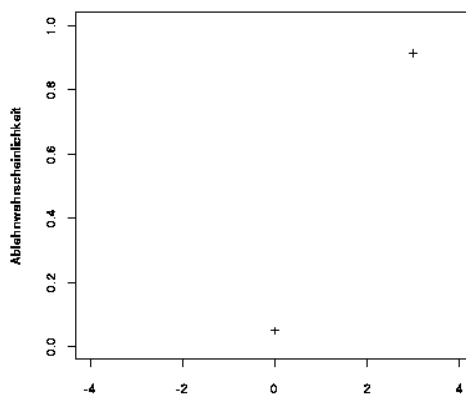
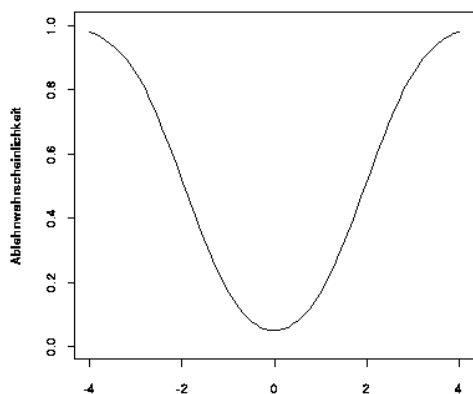
Übungsblatt 6

Abgabe: Mittwoch 3. Juni 2009, bis spätestens 10.00 Uhr; Briefkasten: Stochastik II

1. Betrachtet wird der Datensatz `crabs` im Paket `MASS` in R.
 - (a) Erstellen Sie Boxplots für jede der morphologischen Variablen gegen die vier Kombinationen aus `sp` (species) und `sex`. Was fällt Ihnen dabei auf?
 - (b) Eine Tierärztin behauptet, dass orange männliche Krebse andere Erwartungswerte für die beiden Variablen `CL` und `RW` als die übrigen Krebse haben.
 - i. Beurteilen Sie dies anhand der Plots aus Teilaufgabe (a).
 - ii. Beurteilen Sie das Ganze auch mit Hilfe von z -Tests zu einem Test-Niveau $\alpha = 0.01$. Gehen Sie hierbei von einer Varianz von 50 für `CL` und 6.8 für `RW` aus. Verwenden Sie als Nullhypothesen, dass die Erwartungswerte der beiden Variablen den Durchschnittswerten für `CL` und `RW` der übrigen Krebse entsprechen. Geben Sie auch jeweils den Annahmehbereich an.
2. Bei einem IQ-Test wurden die Werte 84, 88, 89, 92, 94, 96, 97, 99, 99, 101, 102, 102, 103, 106, 109, 111, 111, 113, 115, 116, 117, 120, 122, 126 und 132 erzielt.

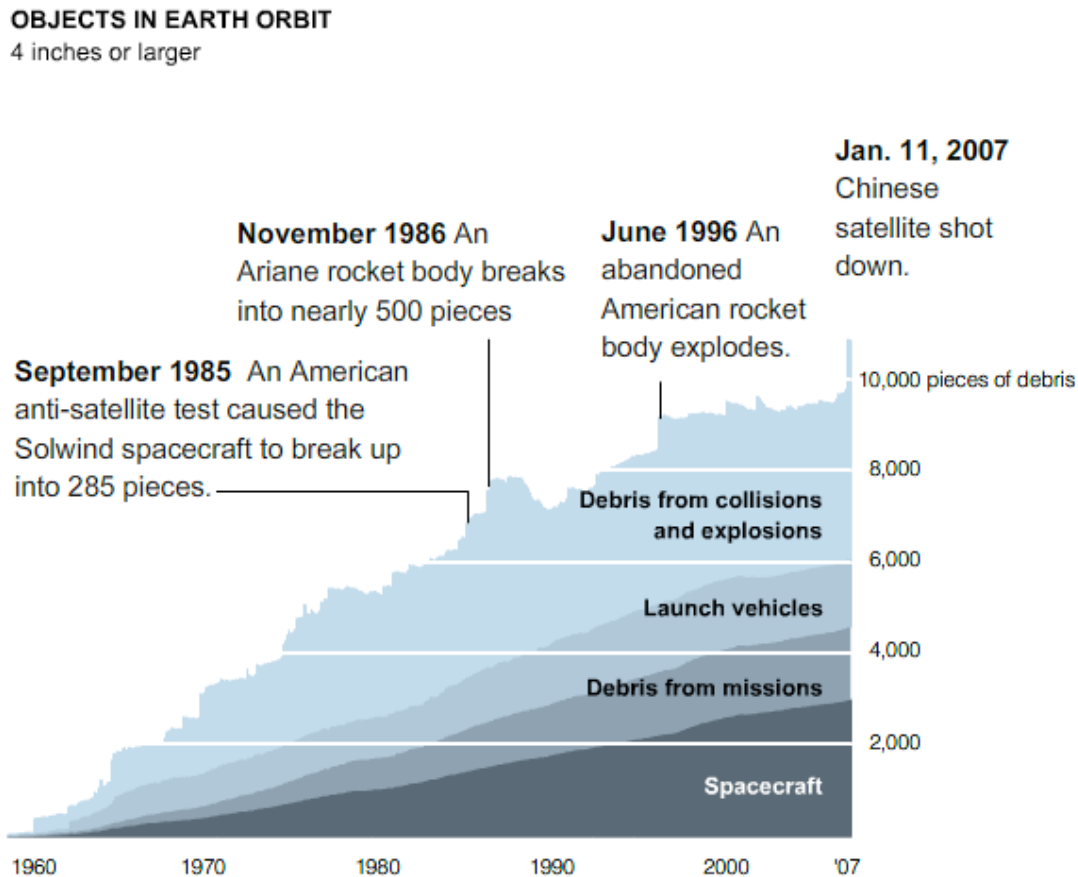
Bearbeiten Sie die nachfolgenden Aufgabenstellungen per Hand.

 - (a) Bestimmen Sie mittels eines z -Tests zu einem Test-Niveau von $\alpha = 0.05$, ob die getesteten Personen aus einer Population mit Erwartungswert 100 stammen, wenn eine Standardabweichung von 10 vorausgesetzt wird.
 - (b) Wie verhält es sich, wenn ein Test-Niveau von $\alpha = 0.001$ verlangt wird?
 - (c) Diskutieren Sie was bei wachsendem Stichprobenumfang passiert.
3. Abgebildet sehen Sie die Gütefunktionen zu zwei verschiedenen Tests. Die Nullhypothese ist jeweils $\mu = 0$.



- Wie lautet jeweils die Alternativhypothese?
- Geben Sie das Signifikanzniveau an.
- Lesen Sie den Typ-II Fehler im rechten Plot ab. Lässt sich der Typ-II Fehler im linken Plot ablesen?
- Wie würden die Gütefunktionen im optimalen (aber wohl unrealistischen) Fall aussehen?

4. Kommentieren Sie die folgende Graphik:



Von Edward Tufte Internetseite

http://www.edwardtufte.com/boardq-and-a-fetch-msg?msg_id=00004b
(abgerufen 16. Mai 2008)

- Was ist das Thema der Graphik und welche Informationen stellt sie bereit?
- Was stellt die y -Achse dar und wie beurteilen Sie die Tatsache, dass diese rechts abgetragen ist?
- Wie lassen sich die Flächen interpretieren? Ist es möglich die Gesamtanzahl aller Raumsonden (Spacecraft) im Zeitraum von 1960 bis 2007 aus der Graphik abzulesen?
- Sollten weitere Informationen in der Graphik gezeigt werden, wenn ja welche?
- Beurteilen Sie die Graphik in Bezug auf Suggestivität, Übersichtlichkeit und Ästhetik.

5. Auf der Internetseite <http://www.zeit.de/sport/tabellen/chancentod> finden Sie eine Tabelle zur Chancenverwertung der Bundesligateams der vergangenen Saison. Untersuchen Sie in folgender Weise in R, ob sich die Trefferwahrscheinlichkeit für den Karlsruher SC von den anderen Teams unterscheidet:

- (a) Schätzen Sie für die anderen Teams eine mittlere Trefferrate. Berechnen Sie den symmetrischen Annahmebereich um 50 für den Binomialtest, wenn Sie von einem Signifikanzniveau von 0.05 bzw. 0.01 und einer Wahrscheinlichkeit von p gleich der mittleren Trefferrate ausgehen. Wäre die Trefferwahrscheinlichkeit des Karlsruher SC jeweils signifikant verschieden?
- (b) Bestimmen Sie mittels des Befehls `binom.test()` den p -Wert. Lassen Sie sich hierzu einmal ein 95%-Konfidenzintervall und einmal ein 99%-Konfidenzintervall ausgeben.

6. *Einzeiler-Aufgaben.*

- (a) Erklären Sie was man unter dem Fehler 2. Art (Fehler Typ II) versteht?
- (b) Was wird hier in R realisiert?

```
x <- rnorm(10)
y1 <- mean(x) - qnorm(0.975) * (1 / sqrt(10))
y2 <- mean(x) + qnorm(0.975) * (1 / sqrt(10))
(y1 > 0) || (y2 < 0)
```