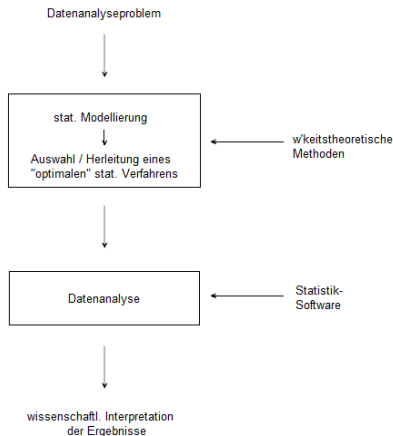


Anwendungsorientierte Beispiele aus der Statistikforschung - Analyse komplexer Datenbestände

Thorsten Dickhaus
Universität Bremen
Institut für Statistik

Chefrühstück
Bremen, 15.11.2016

Was ist Statistikforschung?



Das Problem des multiplen Testens

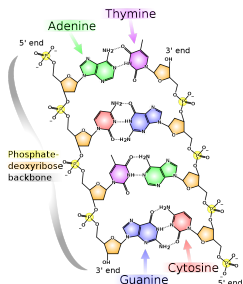
Ein Datensatz → viele Fragestellungen
viele (stat.) Hypothesen simultan prüfen
viele mögliche **Fehlentscheidungen!**

Aufgaben des multiplen Testens:

- Entwicklung von math. Kriterien zur Fehlerkontrolle
- Entwicklung von math. Strukturkriterien für multiple Tests
- vergleichende Beurteilung konkurrierender Verfahren
- Anwendung der Testprozeduren auf reale Daten

Beispiel 1: Genetische Assoziationsstudien

Fragestellung: Welche **Genorte** sind mit gewissen Phänotypen (typischerweise: Krankheiten) assoziiert?



SNiPs or SNPs =

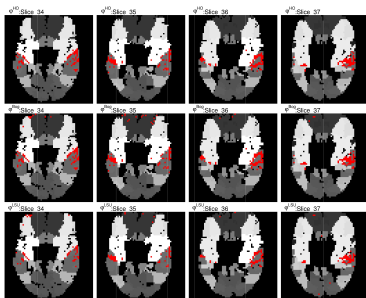
sites of variation in the genome
(spelling mistakes)

Karen	AGCTTGAC	TCCA	TGATGATT
Debo	AGCTTGAC	GCCA	TGATGATT
Jose	AGCTTGAC	TCC	TGATGATT
Thomas	AGCTTGAC	GCC	TGATGATT
Anupriya	AGCTTGAC	TCCA	TGATGATT
Robert	AGCTTGAC	GCCA	TGATGATT
Michelle	AGCTTGAC	TCC	TGATGATT
Zhijun	AGCTTGAC	GCC	TGATGATT

Typische Dimensionalität: **500.000 – 2.5 Mio.**
Genorte gleichzeitig analysieren!

Beispiel 2: Funktionelle Bildgebung

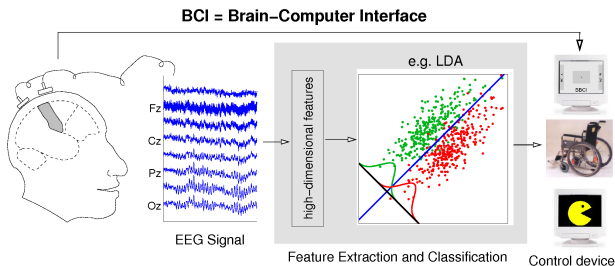
Fragestellung: Welche **Gehirnareale oder Voxel** werden bei gewissen Reizen oder kognitiven Aufgaben aktiviert?



Beispielhafte Dimensionalität: Räumliches Gitter von
 $64 \times 64 \times 64$ Voxels!

Beispiel 3: Brain-Computer Interfacing

Fragestellung: Welche Kombination von EEG-Kanälen liefert die besten Steuerungssignale für gewisse Anwendungen?



Beispielhafte Dimensionalität: 128 EEG-Kanäle!

Arbeitsgruppe "Mathematische Statistik"

Leitung	Prof. Dr. Thorsten Dickhaus
Sekretariat	Martina Titze
Wiss. Mitarb.	Dr. Rostyslav Bodnar Dr. Magdalena Hernández (DFG) André Neumann (DFG-Forschergruppe) Natalia Sirotko-Sibirskaya
Stud. HK	Jonathan von Schroeder
Techniker	Jannis Wilken