

**Statistische Aspekte der Analyse
molekularbiologischer und
genetischer Daten (WS 2016/17)**

Inhaltsverzeichnis

1	V1	1
1.1	Aufbau und Struktur der DNA	1
1.2	Genetischer Code	1
1.3	Replikation / Transkription / Translation	1
1.4	Mitose	1
1.5	Nicht-kodierende RNAs	1
1.6	Aufgaben zur Übung 1	1
2	V2	2
2.1	Mechanismen der epigenetischen Modifikation	2
2.2	Mechanismen der DNA Reparatur	2
2.3	Typische Mutationen	2
2.4	PCR	2
2.5	Sanger Sequenzierung	2
2.6	TaqMan	2
2.7	SNP-Microarray	2
2.8	Aufgaben zur Übung 2	2
3	V3	3
3.1	Meiose	3
3.2	Mendelsche Gesetze	3
3.3	Erbgänge / Stammbäume	3
3.4	Gründe für Abweichungen von Mendelschen Erbgängen	3
3.5	Aufgaben zur Übung 3	3
4	V4	4
4.1	Bias und Präzision	4
4.2	Frequentistischer und Bayesianischer Wahrscheinlichkeitsbegriff . .	4
4.3	Zufallsvariablen (Erwartungswert, Varianz, Standardabweichung, Covarianz, Unabhängigkeit, Randverteilung)	4
4.4	Bedingte Wahrscheinlichkeit, Bayessche Lernformel	4
4.5	Einige wichtige Verteilungsfunktionen	4
4.6	Aufgaben zur Übung 4	4
5	V5	5
5.1	Konfidenzintervall	5
5.2	Logik des statistischen Testens, Testdurchführung und Interpretation	5
5.3	Typ I und Typ II Fehler, Einfluß der Fallzahl	5
5.4	Problem des multiplen Testens und Korrekturmöglichkeiten	5
5.5	Faktoren für die Auswahl des richtigen Tests	5
5.6	Zusammenhangsmaße auf Vierfeldertafeln	5
5.7	Korrelation, Scheinkorrelation und Confounder	5
5.8	Aufgaben zur Übung 5	5

6	V6	6
6.1	Lineare Regression	6
6.1.1	Modellannahme	6
6.1.2	Schätzen der Betas („Intercept“ und „Slope“)	6
6.1.3	Varianzzerlegung und erklärte Varianz bei linearer Regression	6
6.1.4	Multivariate Regression	6
6.1.5	AIC	6
6.2	Multivariate Regression	6
6.2.1	Schätzen von Kontrasten	6
6.2.2	AIC	6
6.2.3	Interaktion	6
6.3	Auswahl einer passenden Regressionsmethode	6
6.4	Aufgaben zur Übung 6	6
7	V7	7
7.1	Motivation und Ansatz für gemischte Modelle	7
7.2	Feste und zufällige Effekte	7
7.3	Idee der Hauptkomponentenanalyse	7
7.4	Interpretation PCA-Plots und Eigenwerte	7
7.5	Aufgaben zur Übung 7	7
8	V8	8
8.1	Hardy-Weinberg Gleichgewicht incl. Test	8
8.2	Kinship-Koeffizient, Verwandtschaftsschätzung	8
8.3	Kopplungsungleichgewicht	8
8.3.1	Entstehung und Entwicklung	8
8.3.2	Bewertung (Maße)	8
8.3.3	Bedeutung (Interpretation, Tagging, LD-Heatmaps)	8
8.4	Aufgaben zur Übung 8	8
9	V9	9
9.1	Interpretation der Fixationsindices F_{st} und F_{is}	9
9.2	Bootstrap, Jackknife als Schätzverfahren für Standardfehler	9
9.3	Hauptkomponentenanalyse in der Genetik (Interpretation)	9
9.4	ROH: Definition und Interpretation	9
9.5	Aufgaben zur Übung 9	9
10	V10	10
10.1	Heritabilität, Definition + Möglichkeiten zur Schätzung	10
10.2	Genetische Assoziation (Prinzip)	10
10.3	Stratifikationsbias bei genetischen Studien	10
10.4	Genetische Modelle und deren Schätzung	10
10.5	Spezifik gonosomaler Markeranalysen	10
10.6	Genomweite Assoziationsstudie	10
10.6.1	Ansatz	10

10.6.2	Replikation	10
10.6.3	Mehrstufigendesign	10
10.6.4	Power	10
10.7	Aufgaben zur Übung 10	10
11	V11	11
11.1	Phänotyp, Genotyp-Phänotyp-Beziehung	11
11.2	Reliabilität, Validität	11
11.3	(Genetische) Studiendesigns:	11
11.3.1	Querschnitt	11
11.3.2	Kohorten	11
11.3.3	Fall-Kontroll	11
11.4	GxE Interaktion	11
11.5	Coverage von Microarrays	11
11.6	Aufgaben zur Übung 11	11
12	V12	12
12.1	Calling von SNP-Daten	12
12.2	Clusterplots + Interpretation	12
12.3	Maße zur Bewertung der Clusterplotirregularität	12
12.4	Typische SNP-QC Maße	12
12.5	Typische Sample-QC Maße	12
12.6	Aufgaben zur Übung 12	12

1 V1

1.1 Aufbau und Struktur der DNA

1.2 Genetischer Code

1.3 Replikation / Transkription / Translation

1.4 Mitose

1.5 Nicht-kodierende RNAs

1.6 Aufgaben zur Übung 1

2 V2

2.1 Mechanismen der epigenetischen Modifikation

2.2 Mechanismen der DNA Reparatur

2.3 Typische Mutationen

2.4 PCR

2.5 Sanger Sequenzierung

2.6 TaqMan

2.7 SNP-Microarray

2.8 Aufgaben zur Übung 2

3 V3

3.1 Meiose

3.2 Mendelsche Gesetze

3.3 Erbgänge / Stammbäume

3.4 Gründe für Abweichungen von Mendelschen Erbgängen

3.5 Aufgaben zur Übung 3

4 V4

4.1 Bias und Präzision

4.2 Frequentistischer und Bayesianischer Wahrscheinlichkeitsbegriff

4.3 Zufallsvariablen (Erwartungswert, Varianz, Standardabweichung, Covarianz, Unabhängigkeit, Randverteilung)

4.4 Bedingte Wahrscheinlichkeit, Bayessche Lernformel

4.5 Einige wichtige Verteilungsfunktionen

4.6 Aufgaben zur Übung 4

5 V5

5.1 Konfidenzintervall

5.2 Logik des statistischen Testens, Testdurchführung und Interpretation

5.3 Typ I und Typ II Fehler, Einfluß der Fallzahl

5.4 Problem des multiplen Testens und Korrekturmöglichkeiten

5.5 Faktoren für die Auswahl des richtigen Tests

5.6 Zusammenhangsmaße auf Vierfeldertafeln

5.7 Korrelation, Scheinkorrelation und Confounder

5.8 Aufgaben zur Übung 5

6 V6

6.1 Lineare Regression

6.1.1 Modellannahme

6.1.2 Schätzen der Betas („Intercept“ und „Slope“)

6.1.3 Varianzzerlegung und erklärte Varianz bei linearer Regression

6.1.4 Multivariate Regression

6.1.5 AIC

6.2 Multivariate Regression

6.2.1 Schätzen von Kontrasten

6.2.2 AIC

6.2.3 Interaktion

6.3 Auswahl einer passenden Regressionsmethode

6.4 Aufgaben zur Übung 6

7 V7

7.1 Motivation und Ansatz für gemischte Modelle

7.2 Feste und zufällige Effekte

7.3 Idee der Hauptkomponentenanalyse

7.4 Interpretation PCA-Plots und Eigenwerte

7.5 Aufgaben zur Übung 7

8 V8

8.1 Hardy-Weinberg Gleichgewicht incl. Test

8.2 Kinship-Koeffizient, Verwandtschaftsschätzung

8.3 Kopplungsungleichgewicht

8.3.1 Entstehung und Entwicklung

8.3.2 Bewertung (Maße)

8.3.3 Bedeutung (Interpretation, Tagging, LD-Heatmaps)

8.4 Aufgaben zur Übung 8

9 V9

- 9.1 Interpretation der Fixationsindices F_{st} und F_{is}
- 9.2 Bootstrap, Jackknife als Schätzverfahren für Standardfehler
- 9.3 Hauptkomponentenanalyse in der Genetik (Interpretation)
- 9.4 ROH: Definition und Interpretation
- 9.5 Aufgaben zur Übung 9

10 V10

10.1 Heritabilität, Definition + Möglichkeiten zur Schätzung

10.2 Genetische Assoziation (Prinzip)

10.3 Stratifikationsbias bei genetischen Studien

10.4 Genetische Modelle und deren Schätzung

10.5 Spezifik gonosomaler Markeranalysen

10.6 Genomweite Assoziationsstudie

10.6.1 Ansatz

10.6.2 Replikation

10.6.3 Mehrstufendesign

10.6.4 Power

10.7 Aufgaben zur Übung 10

11 V11

11.1 Phänotyp, Genotyp-Phänotyp-Beziehung

11.2 Reliabilität, Validität

11.3 (Genetische) Studiendesigns:

11.3.1 Querschnitt

11.3.2 Kohorten

11.3.3 Fall-Kontroll

11.4 GxE Interaktion

11.5 Coverage von Microarrays

11.6 Aufgaben zur Übung 11

12 V12

12.1 Calling von SNP-Daten

12.2 Clusterplots + Interpretation

12.3 Maße zur Bewertung der Clusterplotirregularität

12.4 Typische SNP-QC Maße

12.5 Typische Sample-QC Maße

12.6 Aufgaben zur Übung 12