



Prof. Hanspeter Herzel
Prof. Nils Blüthgen
Dr. Manuela Benary

www.sys-bio.net/teaching
h.herzel@biologie.hu-berlin.de
nils.bluthgen@charite.de
manuela.benary@biologie.hu-berlin.de

COMPUTATIONAL BIOLOGY / BIOINFORMATICS SS 2017

Gib deine Lösung bitte am 6.6. im Kurs ab. Alternativ kannst du die Lösung auch per E-mail an manuela.benary@charite.de schicken.

1. Unabhängigkeit

Blutgruppen einer Sample der Größe 1000 Personen aus der deutschen Bevölkerung sind folgendermassen verteilt:

Blutgruppe	Anteil	Blutgruppe	Anteil
A, rhesus positiv	37%	0, rhesus negativ	6%
0, rhesus positiv	35%	AB, rhesus positiv	4%
B, rhesus positiv	9%	B, rhesus negativ	2%
A, rhesus negativ	6%	AB, rhesus negativ	1%

(a): Stelle eine Kontingenztabelle mit den absoluten Häufigkeiten auf (für A,B,AB,0 auf der einen und Rhesus auf der anderen Seite).

(b): Wie groß ist die Wahrscheinlichkeit, dass eine Person die Blutgruppe 0 hat? Wie groß ist die Wahrscheinlichkeit, wenn man weiss, dass eine Person rhesus negativ ist, dann diese Person die Blutgruppe 0 hat?

(c): Werden die beiden Blutgruppensysteme (A,B,AB,0 und Rhesus) voneinander unabhängig vererbt (Test auf Unabhängigkeit!)?

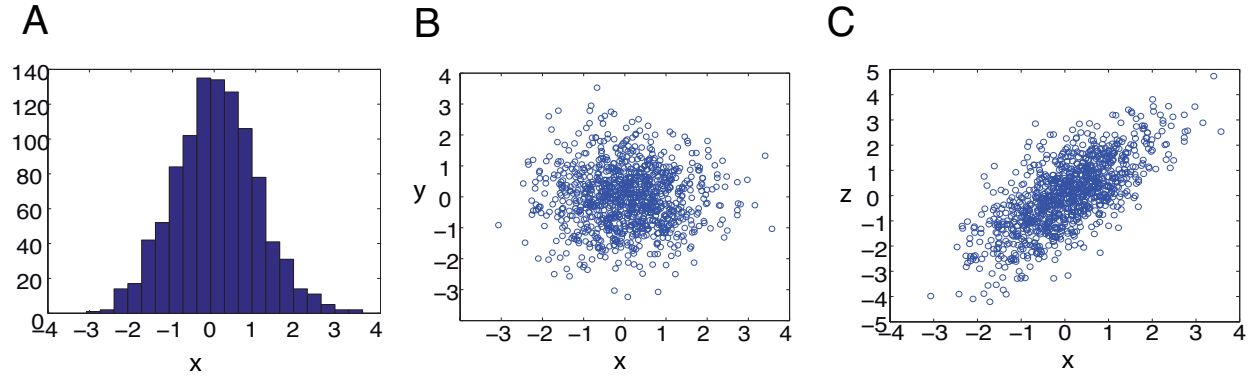
2. Gene Ontology

(a): In welcher Ontologie ist der Term *formaldehyde assimilation via ribulose monophosphate cycle* definiert? (Benutze dazu z.B. die Website www.geneontology.org)?

(b): Welche Terme sind generellere Beschreibungen für diesen Term (also oberhalb in der Hierarchie)?



3. Korrelation



(a): Was für einer Verteilung liegt den Daten aus dem Histogramm in der Abbildung (A) zugrunde? Wie groß sind in etwa Mittelwert und Standardabweichung?

(b): Wie groß sind die Korrelationskoeffizienten der Scatterplot in (B) und (C) in etwa? Bitte mit kurzer Begründung.