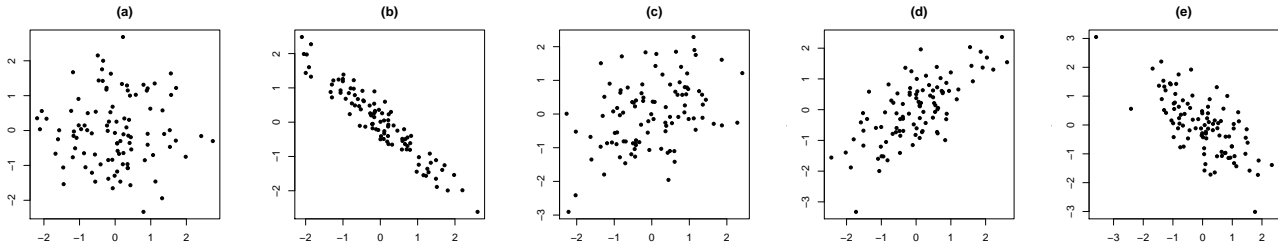


ÜBUNGEN ZUR WAHRSCHEINLICHKEITSRECHNUNG UND STATISTIK FÜR BIOLOGEN

Blatt 12

Alle Aufgaben sollen (ohne Verwendung von R) mit Taschenrechner gelöst werden. Quantiltafeln finden Sie auf der Vorlesungs-Homepage.

**1. Aufgabe a)** Die unten stehenden Scatterplots (a)–(e) zeigen Datensätze mit verschiedenen Korrelationskoeffizienten. Ordnen Sie die Bilder und Korrelationskoeffizienten einander zu.



Korrelationskoeffizient	-0.95	-0.7	0	0.4	0.7
Bild					

**b)** Berechnen Sie für folgenden Mini-Datensatz jeweils Mittelwert und Stichprobenvarianz der  $x$ - und der  $y$ -Werte sowie die Kovarianz von  $x$  und  $y$  und den Korrelationskoeffizienten:  $\begin{array}{c|ccc} x & 1 & 3 & 2 \\ y & 0 & 2 & -2 \end{array}$  (Hinweis: Sowohl Skalierung mit  $n - 1$  als auch mit  $n$  sind sinnvoll.)

**2. Aufgabe** Es könnte sein, dass mit Chlor versetztes Wasser (wie in Schwimmbädern üblich) den Zahnschmelz beeinträchtigt. Dazu wurden 200 Schwimmer je nach dem, ob sie mehr oder weniger als 6 Stunden wöchentlich trainieren, in zwei Gruppen eingeteilt und der Zahnschmelz untersucht, mit folgendem Ergebnis:

Schwimmzeit pro Woche	Zahnschmelz angegriffen		
	Ja	Nein	Gesamt
mehr als 6 h	29	71	100
weniger als 6 h	19	81	100
Gesamt	48	152	200

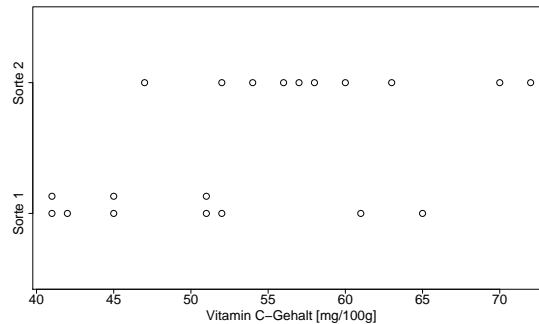
- a)** Stützen die Beobachtungen die Hypothese des ersten Satzes der Aufgabenstellung? Formulieren Sie eine Nullhypothese, testen Sie mittels eines  $\chi^2$ -Tests und formulieren Sie einen Antwortsatz.
- b)** Schätzen Sie den Anteil  $\theta$  der Personen mit angegriffenem Zahnschmelz unter den Viel-Schwimmern. Geben Sie auch ein 95%-Konfidenzintervall an. (Hinweis: Ein Konfidenzintervall könnte sowohl mit Quantilen der Normalverteilung als auch mit Quantilen der t-Verteilung berechnet werden.)
- c)** Nehmen wir an, analoge Ergebnisse wären in einer Studie des doppelten Umgangs beobachtet worden: Je 200 Viel- und 200 Wenig-Schwimmer wurden untersucht, 58 der Viel- und 38 der Wenig-Schwimmer hatten angegriffenen Zahnschmelz. Welchen Wert würde die  $\chi^2$ -Statistik dann annehmen, und wie lautete dann Ihr Befund?
- d)** Welches Konfidenzintervall für  $\theta$  würde sich mit den Daten aus c) ergeben?

**3. Aufgabe** In einer Studie wurde die Wirksamkeit eines Schlafmittels geprüft und für zehn Patienten die zusätzliche Anzahl Stunden Schlaf pro Tag (gemittelt über den Zeitraum der Medikamentengabe) im Vergleich zu einem Referenzzeitraum ohne Medikament bestimmt:

1.9 0.8 1.1 0.1 -0.1 4.4 5.5 1.6 4.6 3.4

Der Mittelwert der zusätzlichen Stunden Schlaf ist  $\mu = 2,33$  mit Varianz  $s^2 = 4,01$ . Testen Sie (unter einer Normalverteilungsannahme an die Daten) die Wirksamkeit des Schlafmittels. Hinweise: Wie immer zuerst die Nullhypothese formulieren. Geben Sie auch ein 95%-Konfidenzintervall für die mittlere zusätzliche Anzahl Stunden Schlaf mit Medikament an.

**4. Aufgabe** Der Vitamin C-Gehalt zweier Kohlsorten wurde in jeweils 10 Stichproben pro Sorte untersucht:



Dabei wurde für die Proben von Sorte 1 ein mittlerer Gehalt  $\mu_1 = 49.4$  mg pro 100g mit Standardabweichung  $s_1 = 8.33$  gefunden, für Sorte 2 waren die Werte  $\mu_2 = 58.9$  und  $s_2 = 7.74$ . Testen Sie die Hypothese, dass der mittlere Vitamin C-Gehalt für beide Sorten gleich ist, mittels eines  $t$ -Tests zum Irrtumsniveau 5% und formulieren Sie Ihr Ergebnis in einem Satz. Nehmen Sie dazu an, dass die wahren Varianzen gleich sind.

In einem Experiment mit verschiedenen Futtersorten wurden drei Gruppen von Küken mit jeweils einer Sorte gefüttert und nach 3 Wochen gewogen. Es kam zu nebenstehendem Ergebnis (Gewicht in Dekagramm).

Sorte A	21.5	20.5	15.7	17.5	
Sorte B	33.1	17.5	16.7	19.2	23.3
Sorte C	32.2	23.7	23.8	26.4	

Führen Sie eine Varianzanalyse durch und testen Sie die Hypothese, dass das mittlere Gewicht bei allen Futtersorten gleich ist.

**6. Aufgabe a)** Phenylketonurie ist eine autosomal-rezessiv vererbte menschliche Stoffwechselkrankheit (d.h. nur homozygote Träger des „kranken“ Allels  $a$  erkranken), an der (in Deutschland) ca. eines von 8000 Neugeborenen leidet.

Nehmen wir an, die Population befinde sich (bezüglich dieses Gens) im Hardy-Weinberg-Gleichgewicht. Welchen Anteil der Population haben dann die Genotypen  $AA$ ,  $aA$  und  $aa$ ?

Ein gesundes Paar (gebildet gemäß der „Hardy-Weinberg-Regeln“) habe ein gesundes Kind. Wie wahrscheinlich ist es, gegeben diese Information, dass beide Eltern Genotyp  $AA$  haben?

**b)** Zum Public Viewing beim Viertelfinale gegen Argentinien kamen 8% der Männer mit einer Fahne und 1% der Frauen mit einer Fahne in das Olympiastadion in München. Von allen Fans im Stadion waren 30% Frauen. Welcher Anteil unter allen Fahnenträgern war weiblich?