

**Aufgabe 1** Aus drei verschiedenen Seen wurden jeweils fünf Welse gefangen. Es soll untersucht werden, ob es zwischen den Seen Unterschiede in Hinblick auf die Größe der Welse gibt. Die Datei `welse.txt` enthält die Gewichte der Welse in kg.

- Visualisieren Sie die Daten in angemessener Weise.
- Führen Sie Schritt für Schritt, also ohne Verwendung des R-Befehls `kruskal.test()`, einen Kruskal-Wallis-Test durch, um zu untersuchen, ob sich die Größenverteilungen der Welse in den Seen unterscheiden.
- Überprüfen Sie Ihr Ergebnis mit dem R-Befehl `kruskal.test()`.
- Führen Sie nun auch eine Varianzanalyse durch und vergleichen Sie die Ergebnisse der beiden Tests. Wie lassen sich die Unterschiede erklären?
- Welches Fazit ziehen Sie in Hinblick auf die Welse? Formulieren Sie Ihr Ergebnis in einem Satz.

**Aufgabe 2** In einem Teich befindet sich eine unbekannte Anzahl  $N$  von Karpfen. Um  $N$  zu schätzen, wurden 50 Karpfen gefangen, markiert und wieder zurückgesetzt. Eine Woche später wurden 60 Karpfen gefangen, darunter waren 13 markierte.

- Wie wahrscheinlich ist es, beim Fang von 60 Karpfen aus  $N - 50$  nicht-markierten und 50 markierten Fischen genau 13 markierte zu finden (wir nehmen dabei an, dass die markierten und die unmarkierten Fische sich wieder gut durchmischt haben)?
- Schätzen Sie die Anzahl Karpfen  $N$  im Teich anhand der Informationen aus der Aufgabenstellung (bestimmen Sie beispielsweise den Maximum-Likelihood-Schätzer).

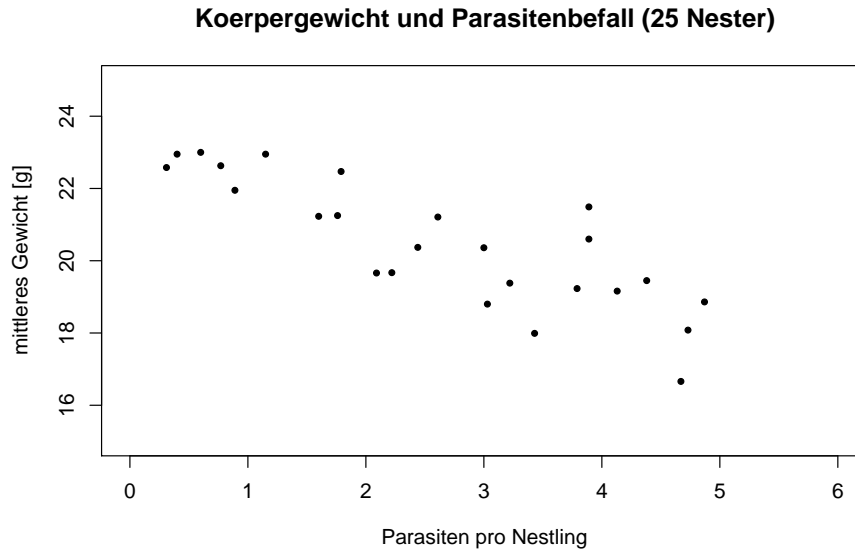
**Aufgabe 3** (simulierte Daten, freundlicherweise von Prof. S. Diehl zur Verfügung gestellt)

Man vermutet, dass es genetische Variabilität der Blattgröße in drei Spitzwegerich-Populationen gibt. Um dies zu prüfen, wurden aus jeder der drei Populationen zufällig Samen gesammelt. 30 Pflanzen, jeweils 10 Pflanzen pro Populationen, wurden gemeinsam im Gewächshaus gezogen. Nach 90 Tagen wurde die Blattfläche der Pflanzen in  $\text{mm}^2$  (als Mittelwert der jeweils drei größten Blätter einer Pflanze) bestimmt, mit folgendem Ergebnis:

Population A:	255	345	312	517	377	458	984	442	487	742
Population B:	344	389	517	598	734	1099	835	465	387	685
Population C:	649	756	484	527	534	388	734	645	865	1077

Stützen diese Daten obige Vermutung?

**Aufgabe 4** Nestlinge der Schwalbe *Hirundo pyrrhonota* werden oft von dem Parasiten *Oeciacus vicarius* befallen. In 25 Nestern mit Nestlingen im Alter von 10 Tagen wurde die Anzahl Parasiten pro Nestling und das mittlere Gewicht der Nestlinge bestimmt. Die Beziehung zwischen dem Gewicht  $G$  und der Parasitenzahl  $P$  scheint ungefähr linear zu sein:  $G = a + b \cdot P$ .



- Zeichnen Sie die Regressionsgerade nach Augenmaß ein.
- Schätzen Sie nach Augenmaß den Achsenabschnitt  $a$ .
- Was ist die biologische Bedeutung von  $a$ ? (z.B.: “ $a$  ist das mittlere Gewicht...”)
- Schätzen Sie nach Augenmaß die Steigung  $b$ .
- Was ist die biologische Bedeutung von  $b$ ? (z.B.: “Jeder zusätzliche Parasit...”)
- Schätzen Sie (zur nächsten ganzen Zahl), den Standardfehler  $s$  der Regression.
- In einem Nest mit leichtem Parasitenbefall gibt es 1.5 Parasiten pro Nestling. Schätzen Sie mit Hilfe Ihrer Schätzungen von  $a$  und  $b$  das mittlere Gewicht der Nestlingen. Mit einem wie großen typischen Fehler müssen Sie bei Ihrer Vorhersage rechnen?

<b>Aufgabe 5</b>	$x$	$y$	Berechnen Sie mit Papier, Bleistift und Taschenrechner die Regressionsgerade für die in der Tabelle gegebenen vier Wertepaare für Merkmal $x$ und Zielgröße $y$ und fertigen Sie eine Skizze an. Hängt $y$ signifikant von $x$ ab?
	1.2	5.4	
	2.3	2.1	
	3.4	3.1	
	5.0	3.2	

**Aufgabe 6** Folgende Tabelle zeigt Dauer des Studiums (in Semestern) und Einstiegsgehalt (in Tausend €) der Absolventen eines Jahres am Fachbereich Mathematik und Informatik der Yule-Simpson-Paradoxon-Universität:

Semester	12	14	16	12	15	14	13	14	11	13	10	12	14	13	14	15
Gehalt	39.4	38.2	37.4	39.5	32.8	35.3	39.1	35.2	37.9	35.7	41	40.9	34.2	38.4	36.2	38.4
Semester	9	11	9	9	12	13	11	10	10	10	9	10	12	10		
Gehalt	33.7	35.9	36.1	34.2	29.9	31.9	33.3	36.2	33.8	32.9	33.3	35.1	34.2	35.3		

- a)** Schlägt sich (für diese Absolventen) ein längeres Studium in einem höheren Anfangsgehalt nieder? Bestimmen Sie die Regressiongerade für Studiendauer gegen Anfangsgehalt.
- b)** Ändert sich Ihr Befund, wenn Sie zusätzlich erfahren, dass die oberen beiden Zeilen der Tabelle sich auf die Absolventen des Fachs Informatik, die unteren beiden sich auf die Absolventen des Fachs Mathematik beziehen, und Sie dieselbe Regression jeweils innerhalb dieser beiden Gruppen durchführen?