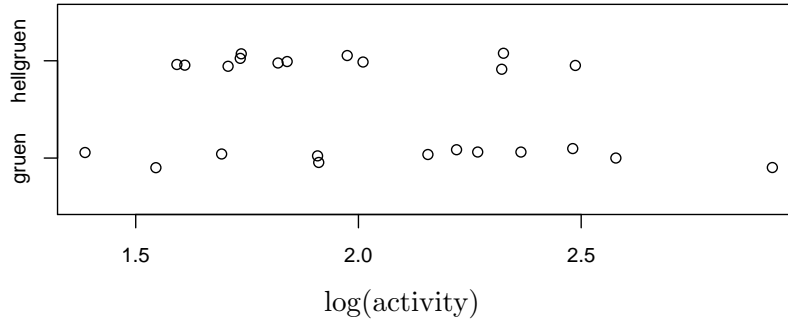


Der  $t$ -Test ist einer der wichtigsten Tests. Sie werden ihn später oft benutzen, um Unterschiede in Ihren Daten zu bewerten. Deshalb üben wir nun seine Anwendung an verschiedenen Beispielen.

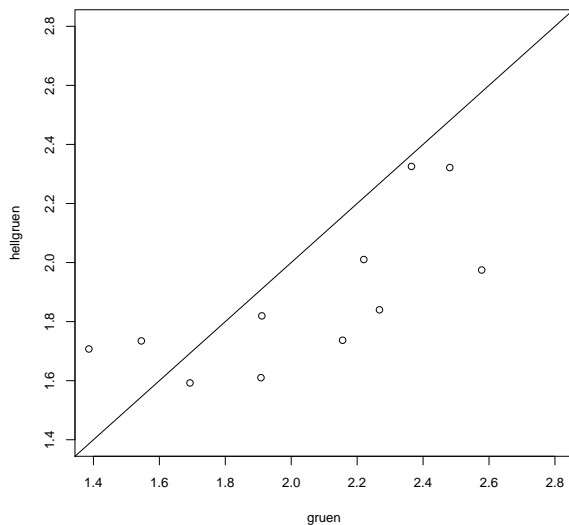
**1. Aufgabe** Sie haben bei 12 Vögeln die Aktivität bei verschiedenen monochromatischen Lichtquellen gemessen. Nun möchten Sie Ihre Messungen auswerten.

**a)** Sind folgende Werte plausibel, oder sind einige offensichtlich falsch?

$\bar{x}(\text{hellgrün}) = 1.93$ ,  $s(\text{hellgrün}) = 0.30$ ,  $\bar{x}(\text{grün}) = 2.12$  und  $s(\text{grün}) = 0.45$



**b)** Anton behauptet, die Aktivität bei grün ist signifikant größer als die Aktivität bei hellgrün. Was ist Ihr Eindruck? Zur Begründung ist kein Test nötig.



**c)** Berta widerspricht: Keiner der Mittelwerte ist signifikant von 2.0 verschieden. Hat sie recht? Zur Begründung ist kein Test nötig. (Berechnen Sie die Standardfehler aus den Werten in Teil a. Tragen Sie Mittelwerte  $\pm$  Standardfehler in die Abbildung ein.)

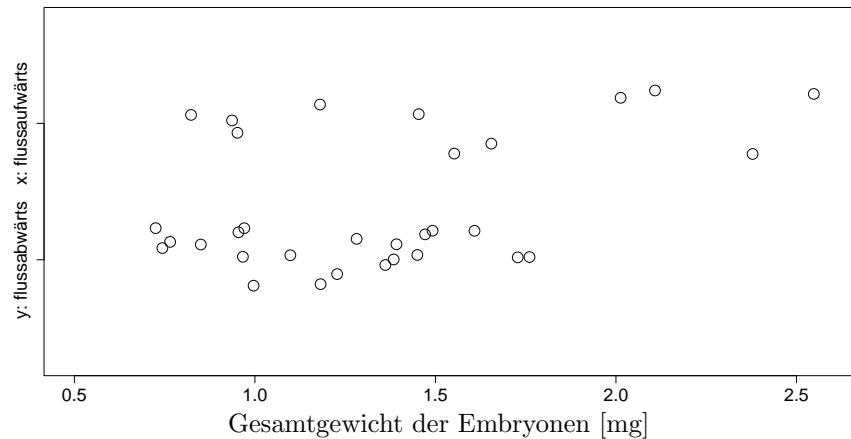
Die Daten oben stammen von 12 Vögeln: von jedem Vogel ein Grünwert und ein Hellgrünwert. Diese Paarung der Werte wird in dem Graphen oben (unzulässigerweise) unterdrückt. Im rechts abgebildeten Graphen ist die Paarung der Werte sichtbar.

**d)** Berechnen Sie die  $t$ -Statistik für einen Test der Gleichheit der Mittelwerte. (Sie können die Mittelwerte in Aufgabe 1 benutzen. Die Standardabweichung der Differenzen ist  $s = 0.27$ .)

**e)** Ist die Differenz auf dem 5%-Niveau signifikant? Benutzen Sie den R-Befehl `pt()` für die Antwort.

**f)** Formulieren Sie das Ergebnis in einem Satz.

**2. Aufgabe** In den Flüssen Trinidads sind die Guppys *Poecilia reticulata* verschiedenen Räubern ausgesetzt. Flussabwärts gibt es den großen und gefährlichen *Crenicichla alta*, flussaufwärts den weniger gefährlichen *Rivulus hartii*. Nach evolutionär-ökologischen Überlegungen sollten die Nachkommen flussabwärts kleiner sein. Um dies zu prüfen, erhob man in beiden Flussabschnitten Stichproben von schwangeren Weibchen (simulierte Daten):



- Flussabwärts war das mittlere Gesamtgewicht der Embryonen  $\bar{y} = 1,21$  mg mit Standardabweichung  $s_y = 0,32$  mg. Wie groß ist der Standardfehler  $f_y$ ?
- $\bar{x}$  ist ungefähr 0.8 1.0 1.2 1.4 1.6 1.8 2.0 mg.
- $s_x$  ist ungefähr 0.2 0.4 0.6 0.8 1.0 1.2 mg.
- $f_x$  ist ungefähr wie groß?
- Tragen Sie die Intervalle  $\bar{x} \pm f_x$  und  $\bar{y} \pm f_y$  in den Graphen ein.
- Wie groß sind (ungefähr) die geschätzte Standardfehler  $f$  von  $\bar{x} - \bar{y}$  und die  $t$ -Statistik zur Prüfung der Hypothese, dass die Populationsmittelwerte von  $x$  und  $y$  gleich sind?
- Ihr Fazit? Formulieren Sie das Ergebnis in einem Satz.

**3. Aufgabe** Man muss sich immer bewusst sein, dass die Wahrheit anders sein kann als ein Test suggeriert. Dies sollen Sie nun mit Hilfe von Simulationen selber beobachten.

**a)** Simulieren Sie mittels `rnorm(n, mean=0)` unabhängige, standard-normalverteilte Beobachtungen für  $n = 10$  und  $n = 100$ . Wie groß ist die Wahrscheinlichkeit, dass Sie anhand der Beobachtungen die Hypothese „der wahre Mittelwert ist gleich 0“ mittels eines zweiseitigen  $t$ -Tests zum Signifikanzniveau 0.01 ablehnen (obwohl sie zutrifft)? Approximieren Sie diese Wahrscheinlichkeit als relative Häufigkeit, indem Sie den Test sehr oft (z.B. 10000 mal) wiederholen (mittels einer `for()`-Schleife). Vergleichen Sie dabei den  $p$ -Wert `t.test()`\$`p.value` mit 0.01.

**b)** Simulieren Sie (beispielsweise mittels `rnorm(n, mean= $\mu$ )`) unabhängige, normalverteilte Beobachtungen mit Mittelwert  $\mu = 1$  und  $n = 10$ . Wie groß ist die Wahrscheinlichkeit, dass Sie anhand der Beobachtungen die Hypothese „der wahre Mittelwert ist gleich 0“ mittels eines zweiseitigen  $t$ -Tests zum Signifikanzniveau 0.01 *nicht* ablehnen (obwohl sie nicht zutrifft)? Approximieren Sie diese Wahrscheinlichkeit als relative Häufigkeit, indem Sie den Test sehr oft (z.B. 10000 mal) wiederholen. Wiederholen Sie die Berechnung nun mit wahren Mittelwert  $\mu = 0.1$ ,  $\mu = 1$  und  $\mu = 5$  und  $n = 10$  bzw  $n = 100$ .

**4. Aufgabe** Student's klassischer sleep-Datensatz: Zwei Schlafmittel wurden an 10 Patienten getestet. Dabei wurde längere Schlafdauer in Stunden relativ zu einer Kontrollgruppe gemessen. Negative Werte bedeuten dabei eine im Vergleich zur Kontrollgruppe verringerte Schlafdauer. Der Datensatz steht Ihnen als `sleep` nach Eingabe von `data(sleep)` zur Verfügung. Die Variable `group` identifiziert dabei, um welches der beiden Medikamente es sich handelt.

- Berechnen Sie die Mittelwerte der Extra-Schlafstunden bei den beiden Medikamenten. Stellen Sie ausserdem den Datensatz grafisch dar (in geeigneter Weise).
- Begründen Sie, warum man hier den gepaarten t-Test anwenden kann.
- Führen Sie den gepaarten t-Test durch.
- Formulieren Sie das Ergebnis in einem Satz.

**5. Aufgabe** Um Vertrauen zum t-Test zu gewinnen, wollen wir das Verhalten des t-Tests bei Daten untersuchen, die nicht normalverteilt sind und deren Stichprobenlänge klein ist. Dazu simulieren wir Daten aus einer U-förmigen Verteilung und berechnen die Wahrscheinlichkeit, dass der t-Test fälschlicherweise Signifikanz anzeigt.

- Plotten Sie `dbeta(x,0.5,0.5)` als Funktion von `x <- seq(from=0,to=1,by=0.01)`. Was ist wohl der Mittelwert dieser Verteilung?
- Samplen Sie einen Datenvektor der Länge  $n = 5$  mit dem R-Befehl `rbeta(n,0.5,0.5)`. Testen Sie, ob der Stichprobenmittelwert signifikant von 0.5 abweicht. Den p-Wert erhalten Sie mit `t.test()$p.value`. Wiederholen Sie dies 10000 mal mittels einer `for()`-Schleife und zählen Sie dabei, wie oft der p-Wert unter 5 Prozent liegt.
- Was die theoretisch zu erwartende Häufigkeit, mit der sich der Stichprobenmittelwert signifikant von 0.5 abweicht. Vergleichen Sie diesen theoretischen Wert mit der relativen Häufigkeit aus obiger Berechnung und erklären Sie die Abweichung.
- Wiederholen Sie die letzte Berechnung mit  $n = 10, 15, 20, 100$ .
- Formulieren Sie ihre Beobachtung in einem Satz.

**6. Aufgabe** Der  $t$ -Test setzt voraus, dass die Stichprobe aus unabhängigen Beobachtungen besteht. Untersuchen Sie empirisch, wie robust der  $t$ -Test gegenüber einer Verletzung dieser Annahme ist. Simulieren Sie dazu einen Beobachtungsvektor `x` mittels

```
x <- sample(rnorm(10),20,replace=TRUE)
```

- Was bedeuten diese R-Befehle? Was ist der tatsächliche Mittelwert der so generierten `x`-Werte? Argumentieren Sie theoretisch oder simulieren Sie sehr oft mit R.
- Testen Sie mittels eines (zweiseitigen)  $t$ -Tests zum Signifikanzniveau 0.05, ob das beobachtete `x` mit der Hypothese, dass der wahre Mittelwert gleich 0 ist, verträglich ist. (Hinweis: mittels `t.test(x)$p.value` können Sie R den  $p$ -Wert ausgeben lassen.) Wiederholen Sie dies 10000 mal mit jeweils neu simuliertem `x`. In wie vielen Fällen lehnen Sie die Hypothese ab, dass der wahre Mittelwert gleich 0 ist?
- Mit welcher Häufigkeit sollte theoretisch die Nullhypothese abgelehnt werden. Überlegen Sie sich, warum obige Häufigkeit von der theoretische Häufigkeit abweicht.
- Formulieren Sie ihr Ergebnis in einem Satz.