

Ihre Namen: _____

Übung 4: Quantitative Genetik 1, Heritabilität

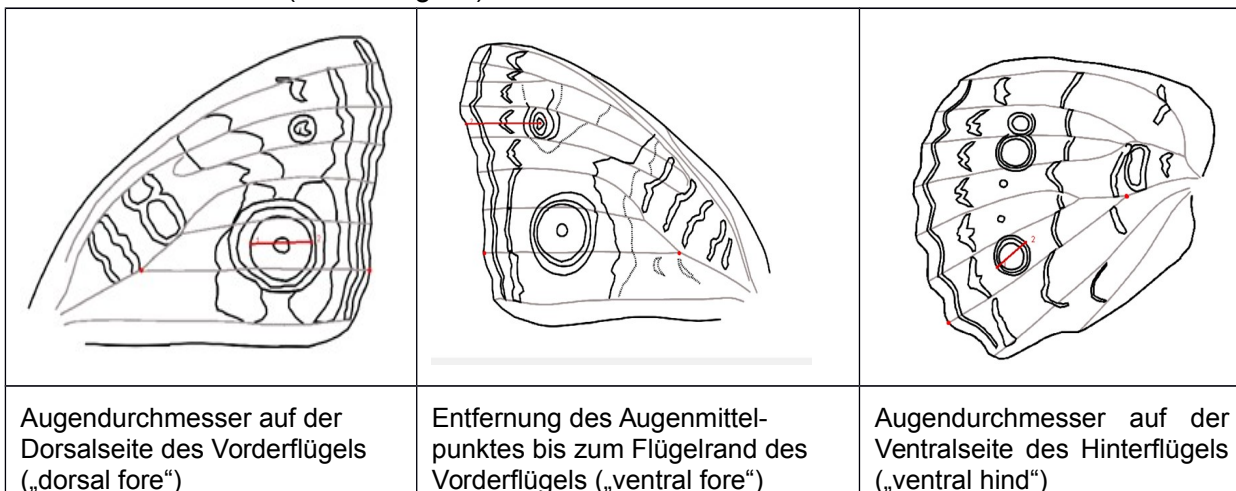
In dieser Übung sollen Sie lernen, die Heritabilität von Merkmalen zu ermitteln. Eine Möglichkeit dies zu tun, besteht darin Merkmale zwischen Eltern und ihren Kindern zu vergleichen. Für diese Übung wollen wir uns morphologische Merkmale von Schmetterlingsflügeln der Art *Precis coenia* (Nordamerikanisches Pfauenauge) ansehen. Die Daten werden anschließend mit dem Statistikprogramm R ausgewertet.

Diese Übung beruht auf der Dissertation von Susan Paulsen, für die sie viele Schmetterlingsflügel per Hand vermessen hat. Mit Ihrer Hilfe wurde auch das Programm entwickelt, welches Sie für die heutige Übung benutzen werden.

Teil 1) Datenerhebung an Schmetterlingsflügeln und Import der Daten in R

a. Computerprogramm starten. Suchen Sie bitte auf dem D-Drive nach einem Ordner, der das Wort Schmetterling im Namen hat. Darin sollte das Programm pcqgl.jar sein. (Falls nicht, bitte das Programm von der Webseite <http://www.ibiblio.org/butterflylab/> herunterladen und im Dokumenten-Ordner auf dem D-drive installieren und den Ordner „Schmetterlinge-Übung“ nennen). Wenn Sie das Programm gefunden/installiert haben, starten Sie es bitte durch einen Doppelklick. Es dauert eine Weile bis das Programm geladen ist. Klicken Sie auf „**New Lab**“ um eine neue Session zu starten.

b. Merkmale auswählen. Unter „Lab Setup“ kann nun das Merkmal ausgesucht werden, dessen Heritabilität bestimmt werden soll. Anhand der beiden bereits eingetragenen roten Marker-Punkte wird die Flügelgröße gemessen, womit die gemessenen Werte normalisiert werden können. Entscheiden Sie sich zuerst für **eines** der drei unten abgebildeten Merkmale (die rot eingezeichneten Distanzen). Gut wäre es wenn alle Gruppen in Ihrer Tischreihe verschiedene Merkmale benutzen (wenn möglich).



Falls Sie das mittlere oder rechte Merkmal gewählt haben, müssen Sie rechts im Programm auf „ventral fore“ bzw. „ventral hind“ umschalten. Wählen Sie im Menü auf der rechten Seite aus, dass Sie 30 Familien bearbeiten wollen. Jede Familie besteht aus Vater, Mutter, einem Sohn und einer Tochter. Setzen Sie nun per Mausclick Ihre zwei Marker um dem Programm mitzuteilen, welche Distanz Sie vermessen werden. Klicken Sie auf „Proceed“ und im erscheinenden Pop-up Fenster auf „Proceed anyway“.

c. Datenerhebung. Jetzt können Sie mit den Messungen beginnen. Mit dem 1. und 2. Klick messen Sie bei jedem Flügel zuerst die Flügellänge. Vergleichen Sie dabei nochmal mit der Schemazeichnung unter dem Tab „**Lab Setup**“, um sicherzugehen, dass Sie die Marker richtig setzen. (Während der Messungen kann man immer zwischen den Tabs „Lab Setup“ und „Lab Measurement“ hin und herklicken). Ignorieren Sie bei dieser Längenmessung eventuell auftretende Fransen am hinteren Flügelende. Mit dem 3. und 4. Klick messen Sie die Länge Ihres ausgesuchten Merkmals. Klick auf „**Next**“ bringt Sie zum nächsten Flügel. Sollten Sie sich mal verklicken, können Sie mit „Redo“ einen oder mehrere der zuletzt gesetzten Marker löschen. Sollte das Merkmal einmal nicht zu messen sein (z.B. weil ein Flügel stark beschädigt ist), setzen Sie ein Häkchen bei „Missing Trait“ und klicken „Next“. Rechts oben stehen die bereits vermessenen Individuen und die Flügelgrößen und die Längen des Merkmals. Jeder von Ihnen sollte 30 Familien vermessen. Wenn alle 30 Familien vermessen sind (also 120 Individuen), speichern Sie die Daten auf dem Desktop und wählen Sie als `daten.csv` als Dateiname.

Bevor Sie mit den Daten weiterarbeiten, duplizieren Sie bitte die Datei `daten.csv` für den Fall, dass später etwas schief geht. Die von Ihnen erstellten Daten werden nämlich nicht in der Schmetterlings-Software gespeichert! Nennen Sie die Kopie `daten_sicherung.csv`

d. Import in R. Öffnen Sie R und gehen Sie im Menü unter „Datei->Verzeichnis wechseln...“ auf den Desktop (dem Verzeichnis indem Ihre Datei `daten.csv` liegt). Lesen Sie nun Ihre Daten in R ein:

```
data <- read.csv("daten.csv",header=T,na.strings="N/A")
```

Die Daten werden so in einer Tabelle namens `data` gespeichert. Mit `header=T` werden die Spalten-Überschriften eingelesen und `na.strings="N/A"` sorgt dafür, dass fehlende Werte korrekt erkannt werden. Sie können sich den Inhalt dieser Tabelle nun ansehen, indem Sie in R einfach `data` eingeben.

Sie sollten nun eine Tabelle mit 5 Spalten und den Überschriften **family_id**, **generation** (p=parent, o=offspring), **sex** (m=male, f=female), **wing_size**, **trait_measurement** erhalten. Machen Sie sich klar, was die Angaben in jeder Spalte bedeuten.

Um uns die zukünftige Arbeit zu erleichtern, sortieren wir unsere Daten nach Familie, Generation und Geschlecht:

```
attach(data)
data <- data[order(family_id,generation,sex),]
```

Der Befehl `attach(data)` erlaubt es R die Überschriften in unserer Tabelle zu erkennen und für andere Funktionen zu nutzen, so dass wir sie einfach zum sortieren benutzen können. Geben Sie erneut `data` ein. Ihre Daten sollten nun sortiert sein.

Teil 2) Ermittlung der Heritabilität mehrerer Merkmale

e. Berechnung der Mittelwerte jedes Elternpaares und jedes Geschwisterpaares. Um die Heritabilität eines Merkmals von Eltern zu Kindern zu bestimmen, muss ein Diagramm erstellt werden, in dem für jede Familie die durchschnittliche Merkmalsgröße der Eltern gegen die durchschnittliche Merkmalsgröße der Kinder aufgetragen wird. Die Steigung der Regressionsgeraden, die durch diese Punktmenge läuft, gibt die Heritabilität dieses Merkmals an. Hierzu müssen wir unseren Datensatz in Eltern (parents) und Nachkommen (offspring) unterteilen:

```
par <- subset(data,generation=="p")
off <- subset(data,generation=="o")
```

Diese beiden Befehle erstellen zwei neue Tabellen die nur die jeweiligen Generationen enthalten. Vergewissern Sie Sich, dass es funktioniert hat, indem Sie Sich durch Eingabe von `par` und `off` den Inhalt der Tabellen ansehen. Als nächstes bilden wir den Mittelwert für unser Merkmal von jedem Männchen und Weibchen einer Familie:

```
par_mean_trait <- tapply(par$trait_measurement, par$family_id, mean)
off_mean_trait <- tapply(off$trait_measurement, off$family_id, mean)
```

Dieser Befehl gruppiert alle Einträge mit derselben **family_id** und Die entstehenden Tabellen haben zwei Zeilen: Die erste enthält die **family_id** und die zweite die durchschnittliche Merkmalsgröße.

f. Erstellen des Diagramms und der Regressionslinie

Mit den erhaltenen Daten können wir nun eine lineare Regression durchführen. Sehen Sie Sich dies zunächst graphisch an:

```
plot(par_mean_trait, off_mean_trait)
```

Dies trägt die mittleren Elternwerte auf der X-Achse gegen die mittleren Nachkommenwerte auf der Y-Achse auf. Falls Sie möchten, dass in Ihrem Graph die Achsen bei 0 starten, können Sie dies mit den Parametern "xlim" und "ylim" definieren. Z.B.:

```
plot(par_mean_trait, off_mean_trait, xlim=c(0, 20), ylim=c(0, 25))
```

In diesem Beispiel geht die X-Achse von 0 bis 20 und die Y-Achse von 0 bis 25. Sie sollten die oberen Grenzen natürlich an Ihren Datensatz anpassen, so dass alle Datenpunkte dargestellt werden. Führen Sie nun eine lineare Regression durch:

```
reg_trait <- lm(off_mean_trait ~ par_mean_trait)
```

Vorsicht: beim obigen Befehl ist die Reihenfolge von Eltern und Nachkommen wichtig und im Gegensatz zum Befehl `plot` vertauscht! Die Variable `reg_trait` enthält als Ergebnis der Regression zwei Parameter: den Achsenabschnitt (intercept) und die Steigung (slope) der Gerade. Sehen Sie Sich Ihr Ergebnis durch eintippen von `reg_trait` an. Sie können die Regressionsgerade nun auch in Ihrem Graphen sichtbar machen:

```
abline(reg_trait)
```

Aufgabe 1)

Wie hoch ist also die Heritabilität des Merkmals? Was hat dieser Wert für eine Bedeutung in der quantitativen Genetik? Was sagt dieser Wert über das vermessene Merkmal aus?

Aufgabe 2)

Wie würde die Regressionsgerade aussehen wenn das gewählte Merkmal gar nicht erblich wäre? Erstellen Sie eine Skizze.

Aufgabe 3)

Wie groß ist die Heritabilität der Flügelgröße? Tipp: Die Daten sind in der Spalte **wing_size** zu finden. Wiederholen Sie das Vorgehen zur Berechnung Ihres Merkmals von oben, wobei sie einfach den Parameter **trait_measurement** durch **wing_size** ersetzen.

Aufgabe 4)

Sie haben sich gerade zwei unterschiedliche Merkmale angesehen: Eines das im Bezug zur Augengröße bzw. Augenposition steht, sowie die Flügelgröße. Denken Sie, dass diese beiden Merkmale von einander unabhängig sind? Wie könnte ein eventueller Zusammenhang zwischen dem Augenmerkmal und der Flügelgröße die Berechnung der Heritabilität des Merkmals beeinflussen und wie könnte man das korrigieren?

Aufgabe 5)

Betrachten wir nun die Geschlechter einzeln. Wie hoch ist die geschlechtsspezifische Heritabilität Ihres Merkmals? Also die Heritabilität zwischen Vätern und Söhnen und die zwischen Müttern und Töchtern? Sehen Sie Unterschiede zwischen den Geschlechtern? Wie groß sind diese Heritabilitäten im Verhältnis zur Heritabilität zwischen Eltern und Kindern? Haben Sie eine Erklärung für eventuell auftretende Unterschiede?

Tipp: Für diese Aufgabe benötigen Sie die Messgrößen von Vater und Sohn, sowie Mutter und Tochter jeder Familie. Sie können diese einfach aus Ihrem ursprünglichen Datensatz extrahieren:

```
fathers <- subset(data, generation=="p" & sex=="m")
sons <- subset(data, generation=="o" & sex=="m")
mothers <- subset(data, generation=="p" & sex=="f")
daughters <- subset(data, generation=="o" & sex=="f")
```

Nun können wir die Regression durchführen.

Für Väter/Söhne:

```
plot(fathers$trait_measurement, sons$trait_measurement)
reg_male <- lm(sons$trait_measurement ~ fathers$trait_measurement)
reg_male
abline(reg_male)
```

Und analog dazu für Mütter/Töchter:

```
plot(mothers$trait_measurement, daughters$trait_measurement)
reg_female <- lm(daughters$trait_measurement ~ mothers$trait_measurement)
reg_female
abline(reg_female)
```

Vater/Sohn:

Mutter/Tochter:

Teil 3) Assortative Mating

Aufgabe 6)

Finden Sie in den Flügelgrößen einen Hinweis auf assortative mating? In anderen Worten: paaren sich Weibchen mit großen Flügeln eher mit Männchen mit großen Flügeln und Weibchen mit kleinen Flügeln eher mit Männchen mit kleinen Flügeln? Wie würden Sie vorgehen um das herauszufinden?

Teil 4) Sexueller Dimorphismus

Aufgabe 7)

Bei vielen Organismen mit zwei Geschlechtern, lässt sich sexueller Dimorphismus beobachten, d.h. Merkmale sind zwischen Männchen und Weibchen unterschiedlich ausgeprägt. Ist dies für Ihr Merkmal der Fall? Wie sieht es mit der Flügelgröße aus? Trennen Sie Ihren Datensatz in Männchen und Weibchen (Eltern und Nachkommen zusammengenommen) und führen Sie einen t-Test durch. Was ist das Ergebnis? Unterscheiden sich die Geschlechter?

```
males <- subset(data, sex=="m")
females <- subset(data, sex=="f")
t.test(males$trait_measurement, females$trait_measurement)
t.test(males$wing_size, females$wing_size)
```

Teil 5) Ermittlung der Verteilung Ihres Merkmals

Quantitative Merkmale folgen üblicherweise einer Normalverteilung. Wir wollen sehen, ob dies bei unserem Merkmal auch der Fall ist. Graphisch lässt sich dies veranschaulichen, indem man ein Histogramm erstellt. Da Männchen und Weibchen bezüglich des Merkmals unterschiedlich sein können, behandeln wir die Geschlechter getrennt.

Normalverteilungen werden durch ihren Mittelwert und die Standardabweichung definiert. Berechnen Sie zunächst diese Parameter für Ihre Datensätze. Für Männchen:

```
male_m <- mean(males$trait_measurement, na.rm=T)
male_sd <- sd(males$trait_measurement, na.rm=T)
```

Für Weibchen:

```
female_m <- mean(females$trait_measurement, na.rm=T)
female_sd <- sd(females$trait_measurement, na.rm=T)
```

Der Parameter `na.rm=T` sorgt dafür, dass fehlende Werte in der Berechnung ignoriert werden.

Notieren Sie Ihre Werte:

	Männchen	Weibchen
Mittelwert		
Standardabweichung		

Erstellen Sie nun ein Histogramm für die Männchen:

```
hist(males$trait_measurement, prob=T)
```

Ihre Daten sollten im Histogramm nun mehr oder weniger einer Normalverteilung folgen. Um eine bessere Vorstellung zu bekommen, können sie eine tatsächliche Normalverteilung in Ihr Histogramm einfügen:

```
curve(dnorm(x, mean=male_m, sd=male_sd), add=T)
```

Wiederholen dieses Vorgehen für Ihre Daten von Weibchen:

```
hist(females$trait_measurement, prob=T)
curve(dnorm(x, mean=female_m, sd=female_sd), add=T)
```

Würden Sie sagen Ihre Daten sind normal verteilt? Skizzieren Ihre beiden Histogramme grob in die beiden leeren Kästchen:

Männchen:



Weibchen:



Was könnte die Ursache für die Variabilität in der Merkmalsgröße unter den Schmetterlingen sein?

Ein interessanter Artikel zum Thema:

http://www.phschool.com/science/science_news/articles/butterfly_gets_spots.html