

Züchtungslehre - Übung 5

Peter von Rohr

October 28, 2015

Aufgabe 1 (8)

Diese Aufgabe soll einerseits eine Einführung in das Thema der Varianzkomponentenschätzung geben und andererseits die Vertrautheit mit dem R-package **pedigreemm** fördern. Bis jetzt hatten wir nur lineare Modelle gesehen mit fixen Effekten, so wie zum Beispiel die Effekte an verschiedenen Loci. Eine Annahme in diese Modellen waren, dass die Beobachtungen unkorreliert waren.

In der Tierzucht arbeiten wir mit Beobachtungen von verwandten Tieren. Deshalb können wir nicht annehmen, dass die Beobachtungen unkorreliert sind. Für ein Tier liegen auch mehrere Beobachtungen vor, welche in der Regel nicht unabhängig sind voneinander.

Die Abhängigkeiten zwischen den Beobachtungen werden in sogenannten “linearen gemischten Modellen” analysiert. Die Modelle heissen gemischt, da sie neben fixen Effekten auch noch sogenannte “zufällige” Effekte enthalten. Bei den zufälligen Effekten wird angenommen, sie hätten Erwartungswert 0 und eine bekannte Co-Varianzmatrix.

Beim Thema der Varianzkomponentenschätzung geht es genau darum diese Varianzen aus den Daten zu schätzen.

Als Beispiel für ein lineares gemischtes Modell verwenden wir das Pedigree aus der Dokumentation von “pedigreemm” und generieren daraus einen einfachen Datensatz.

Das Pedigree für die Analyse sieht wie folgt aus und kann in der Hilfedatei von “pedigreemm” kopiert werden. Die Hilfedatei kann mit “?pedigreemm” oder `help("pedigreemm")` aufgerufen werden.

```
sire dam
1 <NA> <NA>
2 <NA> <NA>
3 1 2
4 1 <NA>
5 4 3
6 5 2
```

Der Datensatz kann von der Webseite unter http://charlotte-ngs.github.io/LivestockBreedingAndGenomics/w7/z1_w7_u5_DataLmm.csv heruntergeladen werden. Im Datensatz befinden sich die drei Kolonnen

1. ID: Tier-Id
2. FixerFactor: Fixer Faktor
3. y: Beobachtungswerte

Anpassung eines linearen gemischten Modells

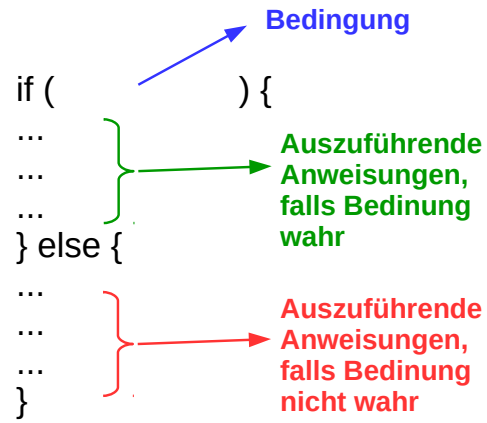
Ihre Aufgabe ist es, für den gegebenen Datensatz ein lineares gemischtes Modell anzupassen. Dazu können Sie die Funktion `pedigreemm()` verwenden. Als erstes Argument für die Funktion `pedigreemm()` muss eine Formel angegeben werden. Die Formel wird gleich angegeben, wie bei der Funktion `lm()`. Im Unterschied zu `lm()` muss die Variable "ID" als zufälliger Effekt angegeben werden. Der zufällige Effekt wird mit `(1 | ID)` angegeben. Das Pedigree wird als Argument "pedigree" angegeben. Das Ergebnis der Schätzung können Sie wieder mit Funktion `summary()` anschauen.

Zusatz

Als Zusatzaufgabe können Sie der Funktion `pedigreemm()` das zusätzliche Argument "REML = FALSE" angeben und dann das Ergebnis mit dem ursprünglichen Ergebnis vergleichen. Das Argument "REML" steuert, ob die Varianzkomponenten mit Maximumlikelihood (ML, d.h. REML = FALSE) oder mit Restricted Maximum Likelihood (REML = TRUE) geschätzt werden soll.

Aufgabe 2 (5)

In der zweiten Aufgabe behandeln wir ein Programmierkonzept in R, womit wir den Programmfluss steuern können. Es geht dabei um sogenannte "if-else"-Anweisungen. Diese haben folgende Struktur.



Als kleine Übung sei eine Liste mit zwei Kolonnen zu erzeugen. Die erste Kolonne enthalte die Zahlen von 1 bis 10 und die zweite die Farben rot und grün. Bei allen Werten der ersten Kolonne, welche kleiner sind als 8 soll "grün" stehen, bei allen anderen Werten soll "rot" stehen. Das Resultat der Liste soll wie folgt aussehen.

```
1, gruen  
2, gruen  
3, gruen  
4, gruen  
5, gruen  
6, gruen  
7, gruen  
8, rot  
9, rot  
10, rot
```