

# Lösungen 11 Probeproofung Züchtungslehre HS 2016

Peter von Rohr

DATUM *16. Dezember 2016*  
BEGINN *09:15 Uhr*  
ENDE

Name:

Legi-Nr:

Aufgabe	Maximale Punktzahl	Erreichte Punktzahl
1	12	
2	24	
3	29	
4	36	
5	12	
6	16	
Total	129	

## **Aufgabe 1: Tierzucht (12)**

- a) Nennen Sie die zwei Werkzeuge, welche in der Tierzucht bei der Auswahl potentieller Elterntiere verwendet werden.

**2**

### **Lösung:**

1. Selektion
2. gezielte Paarung

- b) Die Vermehrung in Wildpopulationen verläuft etwas anders als in Nutztierpopulationen. Nennen Sie die wichtigsten Unterschiede.

5

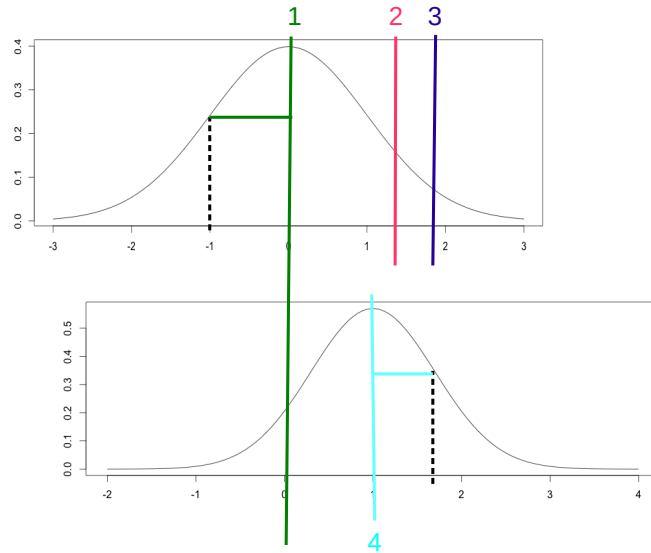
	Wildpopulation	Nutztierpopulation

**Lösung:**

	Wildpopulation	Nutztierpopulation
	natürliche Selektion	künstliche Selektion
	zufällige Paarungen	gezielte Paarung
	Vermehrung	gerichtete Selektion
	optimale Anpassung an Umwelt	Optimierung des vom Menschen geschaffenen Zuchtsystems

c) Was ist im nachfolgenden Diagramm dargestellt. Was bedeuten die Punkte 1-4?

5



**Lösung:**

Es zeigt die Verschiebung der Verteilung von der Elterngeneration zur Nachkommen-  
generation bei gerichteter Selektion. Dabei bedeuten die Punkte:

1. Elterndurchschnitt
2. Selektionsgrenze
3. Remontendurchschnitt
4. Nachkommendurchschnitt

## Aufgabe 2: Verwandtschaft und Inzucht (24)

Gegeben ist folgendes Pedigree

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

- a) Stellen Sie die additive genetische Verwandtschaftsmatrix für das oben dargestellte Pedigree auf.

18

### Lösung:

Die Berechnung erfolgt mit der Matrixmethode. Diese umfasst die folgenden Schritte:

1. Die Tiere werden dem Alter nach von links nach rechts (und von oben nach unten) in einer symmetrischen Matrix angeordnet. Als erstes Tier wird das unbekannte Tier 'NA' eingetragen.
2. Oberhalb jedes Tieres werden seine Eltern eingetragen
3. Die erste Zeile und die erste Kolonnen werden mit lauter Nullen aufgefüllt.
4. Als Diagonalelement für Tier  $x$  tragen wir  $1 + F_x$  ein
5. Die Offdiagonalelemente auf der Zeile für Tier  $x$  werden mit den Verwandtschaftsgraden  $a_{x,y}$  zwischen Tier  $x$  und Tier  $y$  aufgefüllt. Dabei gilt, dass

$$a_{xy} = \frac{1}{2}(a_{x,m(y)} + a_{x,v(y)}) \quad (1)$$

6. Damit die Matrix symmetrisch wird, werden die berechneten Werte aus der Zeile für Tier  $x$  in die Kolonnen für Tier  $x$  übertragen.

Die Verwandtschaftsmatrix für das gezeigte Pedigree lautet

```
> matA <- as.matrix(getA(ped = ped))  
> print(matA)
```

	1	2	3	4	5	6
1	1.00	0.00	0.50	0.50	0.500	0.750
2	0.00	1.00	0.50	0.50	0.500	0.250
3	0.50	0.50	1.00	0.50	0.750	0.500
4	0.50	0.50	0.50	1.00	0.750	0.750
5	0.50	0.50	0.75	0.75	1.250	0.625
6	0.75	0.25	0.50	0.75	0.625	1.250

b) Welche der Tiere 1-6 im anfangs gezeigten Pedigree sind ingezüchtet und wie gross ist deren Inzuchtkoeffizient?

4

**Lösung:**

- Tiere 5 und 6 sind ingezüchtet
- Der Inzuchtkoeffizient für Tier 5 beträgt 0.25 und für Tier 6 beträgt er 0.25.

- c) Angenommen die Tiere 5 und 6 sind ingezüchtet. Nach wievielen Berechnungsschritten in der rekursiven Berechnungsart der Verwandtschaftsmatrix kann der Inzuchtgrad der Tiere 5 und 6 aus der Matrix bestimmt werden? (Hinweis: Das Übertragen der Elemente von einer Zeile in die entsprechende Kolonne soll nicht als Berechnungsschritte gezählt werden.)

**2**

**Lösung:**

Generell kann der Inzuchtgrad eines Tieres aus der Verwandtschaft der Eltern bestimmt werden. Tier 6 hat Eltern 1 und 4, der Inzuchtgrad ist als nach 4 Berechnungsschritten ablesbar. Tier 5 hat Eltern 3 und 4. Der Inzuchtgrad ist dann erst nach 13 Schritten ablesbar.



### Aufgabe 3: Inverse der Verwandtschaftsmatrix (29)

Gegeben ist folgendes Pedigree

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

- a) Berechnen Sie die Inverse  $A^{-1}$  der additiv genetischen Verwandtschaftsmatrix für das oben dargestellte Pedigree.

16

#### Lösung:

Die Regeln zur Berechnung von  $A^{-1}$  sind in der Lösung von Aufgabe b) beschrieben. Die Regeln basieren auf der LDL-Zerlegung der Matrix  $A$  und diese ist in der Lösung von Aufgabe c) hergeleitet.

Die Inverse  $A^{-1}$  lautet

```
> matAInv <- as.matrix(getAInv(ped = ped))
> print(matAInv)
```

```
      1 2      3 4      5
1 2 1 -1.0000000 -1 0.0000000
2 1 2 -1.0000000 -1 0.0000000
3 -1 -1 2.3333333 0 -0.6666667
4 -1 -1 0.0000000 2 0.0000000
5 0 0 -0.6666667 0 1.3333333
```

b) Wie lauten die Regeln zur Berechnung der Inversen  $A^{-1}$

7

**Lösung:**

- Initialisiere alle Elemente der Matrix  $A^{-1}$  mit 0
- Für Tier  $i$  mit Eltern  $s$  und  $d$ ,
  - addiere  $\delta_i$  zum Element  $(i, i)$ ,
  - addiere  $-\delta_i/2$  zu den Elementen  $(s, i)$ ,  $(i, s)$ ,  $(d, i)$  und  $(i, d)$  und
  - addiere  $\delta_i/4$  zu den Elementen  $(s, s)$ ,  $(s, d)$ ,  $(d, s)$  und  $(d, d)$
- Für Tier  $i$  mit bekanntem Elternteil  $d$ ,
  - addiere  $\delta_i$  zum Element  $(i, i)$ ,
  - addiere  $-\delta_i/2$  zu den Elementen  $(d, i)$  und  $(i, d)$  und
  - addiere  $\delta_i/4$  zu den Elementen  $(d, d)$
- Für Tier  $i$  mit unbekanntem Eltern
  - addiere  $\delta_i$  zum Element  $(i, i)$

Die Elemente  $\delta_i$  sind die Diagonalelemente der Matrix  $D^{-1}$  aus der LDL-Zerlegung der Verwandtschaftsmatrix  $A$ . In einem Pedigree ohne Inzucht können die Elemente  $\delta_i$  nur die folgenden drei Werte annehmen

$$\delta_i = \begin{cases} 2 & , \text{ falls beide Eltern bekannt} \\ 4/3 & , \text{ falls nur ein Elternteil bekannt} \\ 1 & , \text{ falls beide Eltern unbekannt} \end{cases}$$

- c) Auf welchem Prinzip basiert die Berechnung der Inversen Verwandtschaftsmatrix und weshalb ist das verwendete Prinzip einfacher als die direkte Invertierung der Verwandtschaftsmatrix?

6

**Lösung:**

- Die Berechnung der Inversen der Verwandtschaftsmatrix basiert auf der LDL-Zerlegung der Verwandtschaftsmatrix  $A$ . Diese Zerlegung erlaubt es die Matrix  $A$  als Produkt von drei Matrizen  $L$ ,  $D$  und  $L^T$  zu schreiben. Dabei gilt zu beachten, dass die Matrizen  $L$  und  $D$  eine einfache Struktur haben.  $L$  ist eine linke untere Dreiecksmatrix, welche die Pfade von den Tieren zu den Foundertieren enthält. Die Matrix  $D$  ist eine Diagonalmatrix, welche ein Bestandteil der Kovarianzmatrix der Mendelian-Sampling Effekte ( $m$ ) darstellt. Somit gilt

$$A = L * D * L^T \quad (2)$$

- Die Inverse Verwandtschaftsmatrix kann aus den Inversen dieser drei Matrizen geschrieben werden. Aus (2) folgt

$$A^{-1} = (L^{-1})^T * D^{-1} * L^{-1} \quad (3)$$

- Die Inversen  $L^{-1}$  und  $D^{-1}$  sind einfach zu berechnen.  $D^{-1}$  ist auch eine Diagonalmatrix, deren Elemente den inversen Elementen aus  $D$  entsprechen. Die Matrix  $L^{-1} = I - P$ , wobei  $I$  die Einheitsmatrix ist und  $P$  die Matrix aus der Zerlegung Zuchtwerte in die Zuchtwerte der Eltern plus die Mendelian Sampling Effekte  $a = P * a + m$  entspricht.

## Aufgabe 4: BLUP-Tiermodell (36)

Gegeben ist der folgende Datensatz

Tochter	Herde	Vater	Mutter	Leistung
4	1	C	NA	110.00
5	1	A	4	100.00
6	2	B	5	110.00

Die genetisch-additive Varianz beträgt  $\sigma_a^2 = 12.5$  und die Restvarianz ist gegeben als  $\sigma_e^2 = 50$

- a) Wie lautet das BLUP-Tiermodell in Matrix-Vektor-Schreibweise? Beschreiben Sie die einzelnen Modellkomponenten und geben Sie auch die Erwartungswerte und die Varianzen der Modellkomponenten an.

14

### Lösung:

Das BLUP-Tiermodell lautet:

$$y = Xb + Za + e \quad (4)$$

- mit  $y$  Vektor der phänotypischen Leistungen, entspricht der Kolonne Leistung im Datensatz  
 $b$  Vektor der fixen Herden-Effekte  
 $X$  Inzidenzmatrix verknüpft  $b$  mit  $y$   
 $a$  Vektor der Zuchtwerte  
 $Z$  Inzidenzmatrix verknüpft  $a$  mit  $y$   
 $e$  Vektor der Resteffekte

Der Erwartungswert  $E[b]$  für die fixen Effekte  $b$  ist  $b$  selber. Die Varianz eines fixen Effektes können wir als 0 angeben.

Die Erwartungswerte und die Varianzen für die zufälligen Teile des Modells lauten:

$$E[a] = 0 \quad \text{und} \quad E[e] = 0$$

Somit ist

$$E[y] = Xb$$

Die Varianzen von  $e$  und  $a$  sind definiert als

$$\text{var}(a) = A * \sigma_a^2 \quad \text{und} \quad \text{var}(e) = I * \sigma_e^2$$

Die Kovarianzen zwischen  $a$  und  $e$  werden auf 0 gesetzt. Daraus können wir die Kovarianzmatrix von  $y$  berechnen als

$$\text{var}(y) = ZGZ^T + R$$

- b) Welche Inzidenzmatrizen gibt es im BLUP-Tiermodell und welche Funktion haben sie? Stellen Sie die Inzidenzmatrizen für das BLUP-Tiermodell für den gegebenen Datensatz auf.

6

**Lösung:**

Im BLUP-Tiermodell gibt es die Inzidenzmatrizen  $X$  und  $Z$ . Die Matrix  $X$  verknüpft die fixen Effekte  $b$  und die Beobachtungen  $y$ . Die Matrix  $Z$  verknüpft die Zuchtwerte  $a$  und die Beobachtungen  $y$ . Für unseren Datensatz sind die Matrizen  $X$  und  $Z$  wie folgt definiert:

$$X = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Diese Definition setzt aber voraus, dass wir die zwei Herden als fixe Effekte betrachten und der Vektor  $b$  der fixen Effekte, wie folgt aussieht

$$b = \begin{bmatrix} b_{Herde1} \\ b_{Herde2} \end{bmatrix}$$

Die Matrix  $Z$  verknüpft Beobachtungen und Zuchtwerte.

$$Z = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

- c) Stellen Sie die Mischmodellgleichungen für das BLUP-Tiermodell zuerst allgemein und dann für das gegebene Zahlbeispiel auf.

16

**Lösung:**

Allgemein lauten die Mischmodellgleichungen für das BLUP-Tiermodell

$$\begin{bmatrix} X^T X & X^T Z \\ Z^T X & Z^T Z + A^{-1} * \alpha \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{a} \end{bmatrix} = \begin{bmatrix} X^T y \\ Z^T y \end{bmatrix}$$

Für unser Zahlenbeispiel sehen die einzelnen Komponenten wie folgt aus:

$$X^T X = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

$$X^T Z = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$Z^T Z = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Bei der Inversen  $A^{-1}$  wurde angenommen, dass die Väter "A", "B" und "C" auf die Positionen 1, 2 und 3 in der Matrize gesetzt wurden. Die Verwandtschaftsmatrix sieht dann wie folgt aus.

$$A = \begin{bmatrix} 1.000 & 0.000 & 0.000 & 0.000 & 0.500 & 0.250 \\ 0.000 & 1.000 & 0.000 & 0.000 & 0.000 & 0.500 \\ 0.000 & 0.000 & 1.000 & 0.500 & 0.250 & 0.125 \\ 0.000 & 0.000 & 0.500 & 1.000 & 0.500 & 0.250 \\ 0.500 & 0.000 & 0.250 & 0.500 & 1.000 & 0.500 \\ 0.250 & 0.500 & 0.125 & 0.250 & 0.500 & 1.000 \end{bmatrix}$$

Die Inverse  $A^{-1}$  kann mit den Regeln aus Aufgabe 3b) direkt aufgestellt werden. Aufgrund der Matrix  $A$  sehen wir, dass kein Tier ingezüchtet ist.

$$A^{-1} = \begin{bmatrix} 1.50 & 0.00 & 0.00 & 0.50 & -1.00 & 0.00 \\ 0.00 & 1.50 & 0.00 & 0.00 & 0.50 & -1.00 \\ 0.00 & 0.00 & 1.33 & -0.67 & 0.00 & 0.00 \\ 0.50 & 0.00 & -0.67 & 1.83 & -1.00 & 0.00 \\ -1.00 & 0.50 & 0.00 & -1.00 & 2.50 & -1.00 \\ 0.00 & -1.00 & 0.00 & 0.00 & -1.00 & 2.00 \end{bmatrix}$$

Die linke Handseite der Mischmodellgleichungen sehen wie folgt aus.

$$X^T y = \begin{bmatrix} 210 \\ 110 \end{bmatrix}$$

$$Z^T y = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 110 \\ 100 \\ 110 \end{bmatrix}$$

Nun können wir alle Komponenten zusammensetzen. Dabei müssen wir berücksichtigen, dass das Varianzverhältnis  $\alpha = \sigma_e^2/\sigma_a^2 = 4$  ist.

$$\begin{bmatrix} 2.00 & 0.00 & 0.00 & 0.00 & 0.00 & 1.00 & 1.00 & 0.00 \\ 0.00 & 1.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 1.00 \\ 0.00 & 0.00 & 6.00 & 0.00 & 0.00 & 2.00 & -4.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 6.00 & 0.00 & 0.00 & 2.00 & -4.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 5.33 & -2.67 & 0.00 & 0.00 \\ 1.00 & 0.00 & 2.00 & 0.00 & -2.67 & 8.33 & -4.00 & 0.00 \\ 1.00 & 0.00 & -4.00 & 2.00 & 0.00 & -4.00 & 11.00 & -4.00 \\ 0.00 & 1.00 & 0.00 & -4.00 & 0.00 & 0.00 & -4.00 & 9.00 \end{bmatrix} \begin{bmatrix} \hat{b}_{Herde1} \\ \hat{b}_{Herde2} \\ \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \\ \hat{a}_4 \\ \hat{a}_5 \\ \hat{a}_6 \end{bmatrix} = \begin{bmatrix} 210 \\ 110 \\ 0 \\ 0 \\ 0 \\ 110 \\ 100 \\ 110 \end{bmatrix}$$

## Aufgabe 5: Selektionsindex (12)

In der Schweinezucht sind Merkmale der Fleischqualität ein wichtiger Bestandteil des Zuchtziels. Nehmen wir an, wir definieren einen Selektionsindex für die Fleischqualität, welcher die Fleischqualitätsmerkmale im Zuchtziel charakterisieren soll. Als Fleischqualitätsmerkmale im Gesamtzuchtwert sollen die Merkmale

- Saftverlust beim Kochen und
- sensorischer Geschmack

berücksichtigt werden.

Da diese Merkmale nicht einfach routinemässig messbar sind, verwenden wir im Selektionsindex die zwei Hilfsmerkmale

1. pH-Wert des Fleisches eine Stunde nach der Schlachtung (pH1)
2. intra-muskulärer Fettgehalt (ImF)

Die Merkmale im Gesamtzuchtwert (Zuchtziel) und die verfügbaren Informationsquellen sind in folgender Tabelle aufgelistet

	Merkmale im Gesamtzuchtwert	Merkmale als Informationsquellen
1	Saftverlust (SV)	pH1
2	Geschmack (GS)	ImF

Für die Merkmale im Gesamtzuchtwert wurden folgende Populationsparameter geschätzt:

Parameter	Wert
Standardabweichung Saftverlust	2
Standardabweichung Geschmack	4
Korrelation zwischen Saftverlust und Geschmack	0.4

Für die Merkmale im Selektionsindex wurden folgende Populationsparameter geschätzt:

Parameter	Wert
Standardabweichung pH1	5
Standardabweichung ImF	4
Korrelation zwischen pH1 und ImF	0

Die Korrelationen zwischen den Merkmalen im Selektionsindex und im Gesamtzuchtwert sind in folgender Tabelle zusammengefasst.

	Saftverlust	Geschmack
pH1	0.7	0.2
ImF	0.1	0.6



Die wirtschaftlichen Gewichte der Merkmale im Gesamtzuchtwert betragen

Merkmal	Wirtschaftliches Gewicht
Saftverlust	1
Geschmack	2

### Ihre Aufgaben

- a) Stellen Sie die Indexgleichungen für die Berechnung der Gewichtung der Merkmale im Selektionsindex auf. Verwenden Sie dazu die Matrix-Vektor-Schreibweise.

**2**

### Lösung:

Die Indexgleichungen lauten

$$\mathbf{Pb} = \mathbf{Gv}$$

wobei  $\mathbf{P}$  die Kovarianzmatrix zwischen den Merkmalen im Selektionsindex ist,  $\mathbf{b}$  der Vektor der unbekanntes Indexgewichte ist,  $\mathbf{G}$  die Matrix der Kovarianzen zwischen den Merkmalen im Index und im Gesamtzuchtwert und  $\mathbf{v}$  der Vektor der wirtschaftlichen Gewichte ist.

b) Stellen Sie die folgenden beiden Covarianz-Matrizen auf:

4

- 1) die Matrix  $G$  als Covarianz-Matrix zwischen den Merkmalen im Selektionsindex und im Gesamtzuchtwert
- 2) Matrix  $P$  als Covarianzmatrix zwischen den Merkmalen im Selektionsindex

**Lösung:**

Die Covarianz-Matrix zwischen den Merkmalen im Selektionsindex und den Merkmalen im Gesamtzuchtwert wird mit  $\mathbf{G}$  bezeichnet und wird wie folgt berechnet

$$\mathbf{G} = \begin{bmatrix} Cov(pH1, SV) & Cov(pH1, GS) \\ Cov(ImF, SV) & Cov(ImF, GS) \end{bmatrix}$$

Die einzelnen Komponenten werden berechnet als  $Cov(M1, M2) = Cor(M1, M2) * SD(M1) * SD(M2)$ . Somit lautet diese Matrix

$$\mathbf{G} = \begin{bmatrix} 0.7 * 5 * 2 & 0.2 * 5 * 4 \\ 0.1 * 4 * 2 & 0.6 * 4 * 4 \end{bmatrix} = \begin{bmatrix} 7 & 4 \\ 0.8 & 9.6 \end{bmatrix}$$

Die Covarianz zwischen den Merkmalen im Selektionsindex

$$\mathbf{P} = \begin{bmatrix} Var(pH1) & Cov(pH1, ImF) \\ Cov(pH1, ImF) & Var(ImF) \end{bmatrix}$$

Setzen wir hier die Zahlen ein, dann folgt daraus

$$\mathbf{P} = \begin{bmatrix} 5 * 5 & 0 * 5 * 4 \\ 0 * 5 * 4 & 4 * 4 \end{bmatrix} = \begin{bmatrix} 25 & 0 \\ 0 & 16 \end{bmatrix}$$

- c) Berechnen Sie die Gewichte der Merkmale im Selektionsindex aufgrund der unter Teilaufgabe a) aufgestellten Indexgleichungen und den unter b) berechneten Kovarianzmatrizen

6

**Lösung:**

Die Gewichte berechnen sich als

$$\mathbf{b} = \mathbf{P}^{-1} * \mathbf{G} * \mathbf{v}$$

```
> (vecB = solve(matP) %*% matG %*% vEcoVal)
```

```
      [,1]
[1,] 0.60
[2,] 1.25
```

Checks

```
> matP %*% vecB - matG %*% vEcoVal
```

```
      [,1]
[1,] 1.776357e-15
[2,] 0.000000e+00
```

## Aufgabe 6: Varianzkomponentenschätzung (16)

Gegeben ist der folgende Datensatz.

Tochter	Herde	Vater	Leistung
1	1	C	110.00
2	1	A	127.00
3	2	B	110.00
4	2	A	101.00
5	2	C	200.00
6	3	C	170.00
7	3	C	110.00
8	3	A	100.00
9	3	B	150.00

- a) Um den Einfluss der Herde auf die Leistung der Töchter abschätzen zu können, verwenden wir das folgende fixe Modell.

8

$$y = Xb + e$$

mit  $y$  Vektor der Leistungen  
 $b$  Vektor der fixen Herdeneffekte  
 $X$  Inzidenzmatrix von  $b$   
 $e$  Vektor der Reste

Aus der folgenden Regressionsanalyse kennen wir die geschätzten Herdeneffekte

```
> lmRegHerde <- lm(Leistung ~ -1 + Herde, data = dfMlrData)
> coefficients(lmRegHerde)
```

```
Herde1 Herde2 Herde3
118.5 137.0 132.5
```

### Ihre Aufgaben

Wir nehmen an, dass die Resteffekte  $e$  unkorreliert sind und somit gilt, dass  $\text{var}(e) = I * \sigma_e^2$ , wobei  $I$  die Einheitsmatrix darstellt. Schätzen Sie die Restvarianz  $\sigma_e^2$  aufgrund der Residuen des oben gezeigten Regressionsmodells.

#### Lösung:

Die Residuen  $r_i$  erhalten wir durch Subtraktion der Beobachtungen  $y_i$  minus die gefitteten Werte  $\hat{y}_i = x_i^T \hat{b}$ . Für unser Zahlenbeispiel heisst das konkret, dass wir von jedem Beobachtungswert  $y_i$  einfach den entsprechenden Herdeneffekt abziehen müssen. Die

Herdeneffekte sind oben gegeben. Also ziehen wir von den Beobachtungen  $y_1$  und  $y_2$  den Effekt der Herde 1 ab, von den Beobachtungen  $y_3$  bis  $y_5$  den Effekt der Herde 2 usw. Als Resultat erhalten wir den Vektor der Residuen, der für unser Beispiel wie folgt aussieht.

```
> ### # Vektor der Residuen
> (vecRes <- as.vector(residuals(lmRegHerde)))

[1] -8.5  8.5 -27.0 -36.0  63.0  37.5 -22.5 -32.5  17.5
```

Daraus lässt sich der Vektor der quadrierten Residuen berechnen.

```
> ### # Vektor der quadrierten Residuen
> (vecRes2 <- vecRes^2)

[1]  72.25  72.25  729.00 1296.00 3969.00 1406.25  506.25 1056.25  306.25
```

Summieren wir über die quadrierten Residuen und teilen die Summe durch die Anzahl Freiheitsgrade, dann erhalten wir die gefragte Schätzung der Restvarianz. Die Anzahl Freiheitsgrade hier entspricht den Anzahl Beobachtungen minus die Anzahl Herden.

```
> ### # Restvarianz
> nResVarEst <- crossprod(vecRes) / (nrow(dfMlrData) - nNrHerde)
> cat("Restvarianz: ", nResVarEst, "\n")
```

```
Restvarianz: 1568.917
```

- b) Den Einfluss der Väter auf die Leistungen der Töchter analysieren wir mit einem Modell, in welchem die Vätereffekte als zufällig betrachtet werden.

8

Das Modell mit den zufälligen Vätereffekten sieht, wie folgt aus

$$y = 1\mu + Zu + e$$

mit  $y$  Vektor der Leistungen  
 $\mu$  allgemeiner Mittelwert  
 $1$  Vektor, deren Komponenten alle gleich 1  
 $u$  Vektor der zufälligen Vätereffekte  
 $Z$  Inzidenzmatrix für  $u$   
 $e$  Vektor der Reste

Die Varianzen für die zufälligen Effekte  $u$  und  $e$  sind bestimmt durch

$$\text{var}(u) = I * \sigma_u^2 \quad \text{und} \quad \text{var}(e) = I * \sigma_e^2$$

Die ANOVA-Tabelle des obigen Modells ist nachfolgend gegeben.

```
> tabAovVater <- aov(formula = Leistung ~ Vater, data = dfMlrData)
> summary(tabAovVater)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Vater	2	2499	1250	1.021	0.415
Residuals	6	7344	1224		

### Ihre Aufgaben

Schätzen Sie aufgrund der gegebenen ANOVA-Tabelle die Varianz ( $\sigma_u^2$ ) der zufälligen Vätereffekte und die Varianz ( $\sigma_e^2$ ) der Resteffekte.

#### Lösung:

Die Schätzung der Restvarianz  $\hat{\sigma}_e^2$  beträgt:

$$\hat{\sigma}_e^2 = 1224$$

Die Schätzung der Varianz  $\hat{\sigma}_u^2$  der Vätereffekte beträgt:

$$\hat{\sigma}_u^2 = \frac{1250 - 1224}{9} = 2.89$$