

Peter von Rohr

Institute of Agricultural Sciences  
D-USYS  
ETH Zurich

751-6212-00L V  
Solutions to Exam  
Applied Genetic Evaluation  
SS 2020

Date: 25th May 2020

Name: Xenia Klaus

Legi-Nr: 13761044

| Problem | Maximum Number of Points | Number of Points Reached |
|---------|--------------------------|--------------------------|
| 1       | 24                       |                          |
| 2       | 5                        |                          |
| 3       | 16                       |                          |
| 4       | 17                       |                          |
| Total   | 62                       |                          |

*Questions in German are in italics*

## Problem 1: Breeding Program

- a) What are the constituent parts of a breeding program?

*Wie lauten die Bestandteile eines Zuchtprogramms?*

**6**

## Solution

1. Breeding goal
2. Performance Tests
3. Prediction of Breeding Values
4. Reproduction Technologies
5. Selection and Mating
6. Selection Response

- b) The following table lists the characteristic variables of two breeding programs. Which quantity do you use to compare the success of two breeding programs? Which of the two breeding programs shown below is better according to the quantity of comparison?

*Die nachfolgende Tabelle zeigt die charakteristischen Kennzahlen von zwei Zuchtprogrammen. Welche Grösse wird verwendet um den Erfolg von Zuchtprogrammen zu vergleichen? Welches der beiden unten beschriebenen Zuchtprogramme ist erfolgreicher gemessen an der verwendeten Vergleichsgrösse?*

4

| Variable                              | Breeding Program 1 | Breeding Program 2 |
|---------------------------------------|--------------------|--------------------|
| Selection intensity                   | 1.76               | 1.76               |
| Accuracy of predicted breeding values | 0.83               | 0.58               |
| Genetic standard deviation            | 261.00             | 261.00             |
| Generationinterval (in years)         | 9.00               | 3.00               |

### Solution

The comparison is done based on **selection response per year**  $R$  which is defined as

$$R = \frac{i * r_{u,\hat{u}} * \sigma_g}{L}$$

- Breeding Program 1:

```
sel_r1 <- sel_int * acc_bv1 * gen_sd / gen_int1
cat(" * Selection response for breeding program 1: ", sel_r1, "\n")
```

```
## * Selection response for breeding program 1: 42.3632
```

\*Breeding Program 2:

```
sel_r2 <- sel_int * acc_bv2 * gen_sd / gen_int2
cat(" * Selection response for breeding program 2: ", sel_r2, "\n")
```

```
## * Selection response for breeding program 2: 88.8096
```

- Comparison

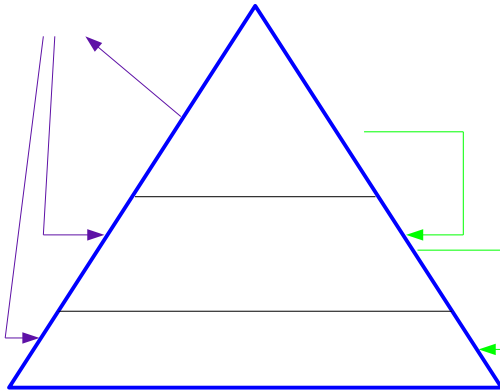
```
## * Breeding program 2 is better than breeding program 1.
```

- c) There are two basic schemes of a breeding program. These schemes are shown below. Please complete the diagrams shown below. In which species are the two schemes implemented?

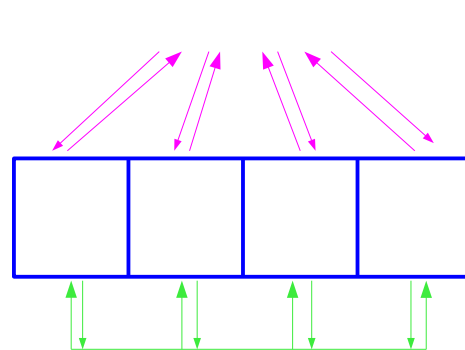
*Es gibt zwei grundsätzliche Schemas eines Zuchtprogramms. Diese Schemas sind unten gezeigt. Bitte vervollständigen Sie die nachfolgenden Darstellungen. In welchen Tierarten werden die beiden Schemas umgesetzt?*

14

Name of the scheme:

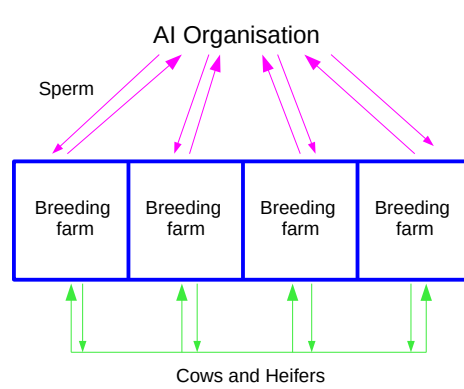
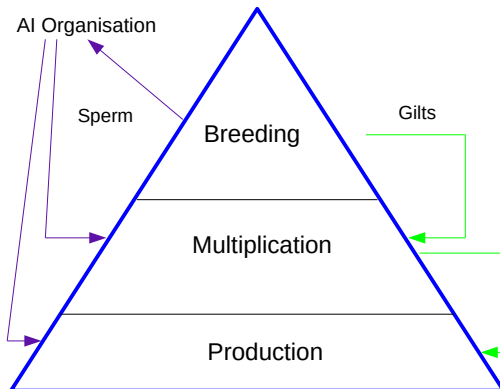


Name of the scheme:



## Solution

The solution



- Left: hierarchical and right: monolithic
- Left: pigs and right: cattle / sheep / goats / horse

## Problem 2: Economic Value

In a pig production farm that operates as a multiplier and has its own fattening operation, the revenues and the costs for producing a slaughter pig are listed below.

*In einem Schweineproduktionsbetrieb mit Ferkelvermehrung und eigener Ausmast, die Erlöse und die Kosten für die Produktion eines Schlachtschweins sind unten aufgelistet.*

| Variable                         | Amount in Swiss Francs |
|----------------------------------|------------------------|
| Revenue per kg slaughter weight  | 4.900                  |
| Net cost per kg slaughter weight | 4.221                  |

- a) Compute the profit per slaughter pig based on the information given in the above table. We assume that the slaughter weight per pig is 80kg per slaughter pig.

*Berechnen Sie den Gewinn pro Mastschwein basierend auf den oben angegebenen Informationen. Dabei betrage das Schlachtgewicht 80kg pro Schlachtschwein.*

**2**

## Solution

The profit per slaughter pig is computed as

```
sl_pig_profit <- slaughter_weight * (slaughter_price - net_cost)
cat(" * Profit per slaughter pig: ", round(sl_pig_profit, digits = 2), "\n")
```

```
## * Profit per slaughter pig: 54.32
```

- b) The number of piglets born alive is an important economic trait for the pig producing farm. Using the trait ‘number of piglets born alive’ in the breeding program allowed the farm to increase the average of trait by 0.1 piglets. This changed the costs as shown in the table below. Compute the economic value for the trait ‘number of piglets born alive’ using the profit computed in Problem 2a as base situation and the profit in the new situation with the incremented level of the trait ‘number of piglets born alive’.

*Die Anzahl lebend geborener Ferkel ist ein wirtschaftlich wichtiges Merkmal in der Schweineproduktion. Dieses Merkmal wird deshalb im Zuchtprogramm verwendet und hat dem Produktionsbetrieb erlaubt die mittlere Anzahl an lebend geborenen Ferkeln um 0.1 Ferkel zu erhöhen. Dadurch verändern sich die Kosten, wie in der unten gezeigten Tabelle aufgeführt. Berechnen Sie das wirtschaftliche Gewicht für das Merkmal Anzahl lebend geborener Ferkel. Verwenden Sie den Gewinn, den Sie in der Aufgabe 2a) berechnet haben als Basissituation. Der Gewinn in der Situation mit dem erhöhten Mittelwert des Merkmals Anzahl lebend geborener Ferkel kann aus den Zahlen aus der nachfolgenden Tabelle berechnet werden.*

3

| Variable                         | Amount in Swiss Francs |
|----------------------------------|------------------------|
| Revenu per kg slaughter weight   | 4.900                  |
| Net cost per kg slaughter weight | 4.207                  |

## Solution

The profit per slaughter pig with increased level of number of piglets born

```
sl_pig_profit_incr <- slaughter_weight * (slaughter_price - net_cost_incr)
cat(" * Profit per slaughter pig: ", round(sl_pig_profit_incr, digits = 2), "\n")

## * Profit per slaughter pig: 55.44
```

The economic value

```
ec_val <- (sl_pig_profit_incr - sl_pig_profit) / n_incr_npa
cat(" * Economic value per piglet: ", round(ec_val, digits = 2), "\n")

## * Economic value per piglet: 11.2
```

### Problem 3: Model Selection

In the search of a diagnostic tool to prevent Hyperketonemia (HYK), keton bodies in the milk (aceton and BHB) are measured as predictor variables. These predictor variables are to be used to quantify changes in the response variable which is Blood Serum BHB. The following table shows the measurements of 12 cows.

*Auf der Suche nach einem Diagnosewerkzeug zur Verhinderung von Hyperketonemia (HYK) wurden Ketonkörper (Aceton und BHB) in der Milch als unabhängige Variablen gemessen. Diese sollen verwendet werden um die Variation in der Zielgrösse BHB im Blut zu quantifizieren. Die folgende Tabelle zeigt die Messwerte von 12 Kühen.*

| Milk Aceton | Milk BHB | Blood Serum BHB |
|-------------|----------|-----------------|
| 0.151       | 0.925    | 1.635           |
| 0.158       | 0.895    | 1.613           |
| 0.181       | 0.875    | 1.644           |
| 0.157       | 0.886    | 1.638           |
| 0.154       | 0.887    | 1.625           |
| 0.147       | 0.901    | 1.565           |
| 0.170       | 0.897    | 1.664           |
| 0.160       | 0.883    | 1.626           |
| 0.136       | 0.894    | 1.580           |
| 0.157       | 0.927    | 1.616           |
| 0.143       | 0.871    | 1.588           |
| 0.161       | 0.921    | 1.595           |

- a) Specify the multiple linear regression model with 'Blood Serum BHB' as response variable. Quantify the effects of the two predictor variables 'Milk Aceton' and 'Milk BHB' on the response variable. What is the expected change of 'Blood Serum BHB' when the response variables 'Milk Aceton' and 'Milk BHB' change by one unit?

*Stellen Sie das multiple lineare Regressionsmodell mit Blut-BHB als Zielgrösse auf. Quantifizieren Sie die Effekte der unabhängigen Variablen Milch-Aceton und Milch-BHB auf die Zielgrösse. Welche Änderung erwarten Sie im Blut-BHB-Wert, wenn die unabhängigen Variablen Milch-Aceton und Milch-BHB um eine Einheit ändern?*

10

### Solution

The multiple linear regression model for Blood Serum BHB on Milk Aceton and Milk BHB can be specified as

$$y = X * b + e$$

where  $y$  is the vector of Blood Serum BHB,  $b$  is the vector consisting of the intercept and the effects of Milk Aceton and Milk BHB,  $e$  is the vector of random residuals. The design matrix  $X$  has dimension  $12 \times 3$  and consists of a column of ones. The second and the third column are the measurement values of Milk Aceton and Milk BHB.

The effects are quantified by the output of the following

```
lm_bbbh <- lm(`Blood Serum BHB` ~ `Milk BHB` + `Milk Aceton`, data = tbl_jer_reg)
summary(lm_bbbh)
```

##

```
## Call:
## lm(formula = `Blood Serum BHB` ~ `Milk BHB` + `Milk Aceton`,
##     data = tbl_jer_reg)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.034265 -0.008380 -0.000336  0.015130  0.028711
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   1.34122    0.33786   3.970  0.00326 **
## `Milk BHB`    -0.00406    0.35474  -0.011  0.99112
## `Milk Aceton`  1.78032    0.56090   3.174  0.01129 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02196 on 9 degrees of freedom
## Multiple R-squared:  0.5305, Adjusted R-squared:  0.4262
## F-statistic: 5.085 on 2 and 9 DF,  p-value: 0.03329
```

When Milk BHB changes by one unit, Blood Serum BHB is expected to change by -0.0040605 units. When Milk Aceton changes by one unit, then Blood Serum BHB is expected to change by 1.7803215 units.



- b) The full model where the response variable is described by all the predictor variables is not necessarily the best model. Please specify two model selection techniques to find the best model. Give a short description of the listed model selection techniques.

*Das volle Modell in welchem alle unabhängigen Variablen verwendet werden um die Zielgröße zu beschreiben ist nicht in allen Fällen das beste Modell. Nennen Sie zwei Modellselektionsstrategien und beschreiben sie diese kurz.*

4

### Solution

1. Forward selection: Start with an empty model. Include those predictors that decrease the residual sums of squares the most. Do this until there are no predictor variables available. From the sequence of models take that with the smallest  $C_p$  value as the best model.
2. Backwards elimination: Start with the full model. Remove the predictor that increases the residual sum of squares the least. Do this until there are no predictor variables to be excluded. From the sequence of models take that with the smallest  $C_p$  value as the best model.

c) Use one of the model selection strategies and determine the best model for the 'Blood-BHB' data.

*Verwenden Sie eine Modellselektionsstrategie und bestimmen Sie das beste Modell für den Blut-BHB Datensatz.*

2

## Solution

Using `olsrr` to do the model selection.

```
tbl_jer_reg_ms <- tibble::tibble(MilkAceton = vec_milk_ace_jer,
                                MilkBHB = vec_milk_bhb_jer,
                                BloodSerumBHB = vec_blood_bhb_jer)

lm_bhb_full <- lm(BloodSerumBHB ~ MilkBHB + MilkAceton, data = tbl_jer_reg_ms)
olsrr::ols_step_best_subset(lm_bhb_full)
```

```
##      Best Subsets Regression
## -----
## Model Index   Predictors
## -----
##      1        MilkAceton
##      2        MilkBHB MilkAceton
## -----
##
##                                     Subsets Regression Summary
## -----
## Model      R-Square    Adj.      Pred      C(p)      AIC      SBIC      SBC      MSEP
## -----
##      1      0.5305      0.4836      0.3206      1.0001     -55.0428    -87.9863    -53.5881    5e-04
##      2      0.5305      0.4262      0.1423      3.0000     -53.0430    -85.3197    -51.1034    7e-04
## -----
## AIC: Akaike Information Criteria
## SBIC: Sawa's Bayesian Information Criteria
## SBC: Schwarz Bayesian Criteria
## MSEP: Estimated error of prediction, assuming multivariate normality
## FPE: Final Prediction Error
## HSP: Hocking's Sp
## APC: Amemiya Prediction Criteria
```

## Problem 4: Genetic Evaluation

The trait **Blood Serum BHB** is used to do a genetic analysis. From the dataset in Problem 3, we selected 4 animals. For those animals pedigree information was added. The new dataset is shown below.

*Für das Merkmal Blut-BHB soll eine genetische Analyse durchgeführt werden. Vom Datensatz in Aufgabe 3 wurden 4 Tiere ausgewählt. Für diese Tiere wurden die Abstammungsdaten zum Datensatz hinzugefügt. Der resultierende Datensatz ist in der nachfolgenden Tabelle gezeigt.*

| Animal | Sire | Dam | Milk Aceton | Milk BHB | Blood Serum BHB |
|--------|------|-----|-------------|----------|-----------------|
| 3      | 1    | 2   | 0.181       | 0.875    | 1.644           |
| 4      | 1    | NA  | 0.136       | 0.894    | 1.580           |
| 5      | 3    | 4   | 0.154       | 0.887    | 1.625           |
| 6      | 3    | 5   | 0.158       | 0.895    | 1.613           |

- a) Use a BLUP animal model to predict breeding values for all the animals given in the dataset shown above. Specify all model components, insert all numeric information into the model components, and write down expected values and variance-covariance matrices for all random effects. As fixed effects you can take the components from the best model found in Problem 3c. If you did not solve Problem 3c, just use all covariables given in the dataset as fixed effects.

*Verwenden Sie ein BLUP-Tiermodell zur Schätzung der Zuchtwerte aller Tiere, welche im oben gezeigten Datensatz gegeben sind. Spezifizieren Sie alle Modellkomponenten, setzen Sie die numerischen Informationen vom Datensatz in die Modellkomponenten ein und notieren Sie die Erwartungswerte und Varianz-Kovarianzmatrizen für alle zufälligen Effekte im Modell. Als fixe Effekte können Sie die Komponenten des besten Modells aus Aufgabe 3c übernehmen. Falls Sie Aufgabe 3c nicht gelöst haben, dann verwenden Sie alle Kovariablen im Datensatz als fixe Effekte.*

15

## Solution

1. BLUP-animal model:

$$y = Xb + Zu + e$$

2. Model components

$$y =$$

$$b =$$

$$X =$$

$$u =$$

$$Z =$$

$$e =$$

### 3. Expected values

$$E \begin{bmatrix} y \\ u \\ e \end{bmatrix} = \begin{bmatrix} Xb \\ 0 \\ 0 \end{bmatrix}$$

### 4. Variance-Covariance Matrices

$$var \begin{bmatrix} y \\ u \\ e \end{bmatrix} = \begin{bmatrix} ZGZ^T + R & ZG & 0 \\ GZ^T & G & 0 \\ 0 & 0 & R \end{bmatrix}$$

- b) Animals 6 and 4 have two offspring which do not have any performance records yet. What type of information is required to determine which of the two fullsibs has the better genetic potential?

*Die Tiere 6 und 4 haben zwei Nachkommen, welche noch keine Leistungen haben. Welche Informationen sind notwendig, damit wir bestimmen können, welches der beiden Vollgeschwister das bessere genetische Potential hat?*

**2**

### **Solution**

Animals with records and the two fullsibs must be genotyped and based on this, genomic breeding values can be predicted. Because fullsibs do not have the same genotypes, their genomic breeding value is different and can be used to find the animal with the better genetic potential.