

Peter von Rohr

Institute of Agricultural Sciences
D-USYS
ETH Zurich

751-6212-00L V
Exam
Applied Genetic Evaluation
SS 2019

Date: 27th May 2019

Name:

Legi-Nr:

Problem	Maximum Number of Points	Number of Points Reached
1	25	
2	26	
3	21	
4	39	
Total	111	

Questions in German are in italics

Problem 1: Model Selection

- a) Model selection can be done using **forward** selection or **backward** selection. Describe the four steps that are needed in both model selection procedures.

*Modellselektion kann mit **Forward** Selektion oder **Backward** Selektion gemacht werden. Beschreiben Sie die vier Schritte, welche es für beide Modellselektionsprozeduren braucht.*

16

Solution

Step	Forward	Backward
1		
2		
3		
4		

- b) We are given the following data set on body weight, breast circumference and shoulder height. Furthermore, the results of fitting the full regression model and a reduced model with only breast circumference as predictor. Compute the Mallows C_p value for both models and decide which of the two models is better based on C_p

Gegeben ist der folgende Datensatz zu Körpergewicht, Brustumfang und Widerristhöhe. Weiter sind auch die Resultate des Fits des vollen Modells und eines reduzierten Modells mit nur Brustumfang als beschreibende Variable gegeben. Berechnen Sie den Mallows C_p -Wert für beide Modelle und begründen Sie, welches der beiden Modelle besser ist.

6

Animal	BreastCircumference	BodyWeight	ShoulderHeight
1	176	471	151
2	177	463	140
3	178	481	155
4	179	470	151
5	179	496	150
6	180	491	151
7	181	518	146
8	182	511	143
9	183	510	144
10	184	541	142

The results of the full model

Die Resultate des vollen Modells

```
##
## Call:
## lm(formula = BodyWeight ~ BreastCircumference + ShoulderHeight,
##     data = tbl_reg_withsh)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -17.7937  -5.6473  -0.3162   9.2916  13.3003
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -1108.4357   393.3949  -2.818  0.02586 *
## BreastCircumference    8.8003    1.7287   5.091  0.00141 **
## ShoulderHeight    0.1389    0.9093   0.153  0.88289
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 11.83 on 7 degrees of freedom
## Multiple R-squared:  0.824, Adjusted R-squared:  0.7737
## F-statistic: 16.39 on 2 and 7 DF, p-value: 0.002287
```

The results of the reduced model

Die Resultate des reduzierten Modells

```
##
## Call:
## lm(formula = BodyWeight ~ BreastCircumference, data = tbl_reg_withsh)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -17.3941  -6.5525  -0.0673   9.3707  13.2594
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -1065.115    255.483  -4.169 0.003126 **
## BreastCircumference    8.673     1.420   6.108 0.000287 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 11.08 on 8 degrees of freedom
## Multiple R-squared:  0.8234, Adjusted R-squared:  0.8014
## F-statistic: 37.31 on 1 and 8 DF,  p-value: 0.000287
```

Solution

- c) What is the reason why we are doing model selection, why should we not take always the full model? Please explain the underlying phenomenon in your own words.

Was ist der Grund, dass wir Modellselektion machen? Weshalb können wir nicht einfach immer das volle Modell mit allen erklärenden Variablen zur Anpassung verwenden? Bitte erklären Sie das unterliegende Phänomen in Ihren eigenen Worten.

3

Solution

Problem 2: Breeding Programs

The breeding goal 2021 (https://homepage.braunvieh.ch/xml_1/internet/de/application/d2/d9/f464.cfm) for the **Brown Swiss** cattle breed contains the following information

Das Zuchtziel 2021 (https://homepage.braunvieh.ch/xml_1/internet/de/application/d2/d9/f464.cfm) für das Braunvieh enthält die folgenden Informationen

Merkmal	Zielwert für Braunvieh
Jährlicher Zuchtfortschritt Milch	+60 kg
Jährlicher Zuchtfortschritt Eiweiss	+2 kg
Milchleistung Ausgewachsene Kuh Talzone	8500 kg
Milchgehalt Eiweiss	3.5%
Laktationspersistenz	85% und höher
Nutzungsdauer (ZWS)	positiver genetischer Trend
Lebensleistung bei Abgang (Talzone)	35000 kg
Eutergesundheit (ZW Zellzahl)	positiver genetischer Trend
ZW Rastzeit	positiver genetischer Trend
ZW Non Return Rate	positiver genetischer Trend
Serviceperiode (phänotypisch)	unter 120 Tagen
Rahmen, Becken, Fundament, Euter, Zitzen	positiver genetischer Trend
Grösse ausgewachsene Kühe	140 - 152 cm

The aggregate genotype of the Swiss **Edelschein** dam lines has the following weights for the different trait groups **Conformation**, **Production** and **Reproduction**.

Im Gesamtzuchtwert der Edelschwein Mutterlinie werden die verschiedenen Merkmalsgruppen ‘Exterieur’, ‘Produktion’ und ‘Reproduktion’ wie folgt gewichtet.

Das Zuchtziel 2021 (https://homepage.braunvieh.ch/xml_1/internet/de/application/d2/d9/f464.cfm) für das Braunvieh enthält die folgenden Informationen

Merkmalsgruppe	Gewicht im Gesamtzuchtwert
Exterieur	19
Produktion	30
Reproduktion	51

- a) What are the names of the two different formulation of the breeding goals for 'Brown Swiss' cattle and for 'Edelschwein' pigs? Please complete the following table and in doing so, indicate both an advantage and a disadvantage of both formulations.

Wie heissen die beiden verschiedenen Arten der Formulierung eines Zuchtziels beim Braunvieh und beim Edelschwein? Bitte füllen Sie die nachfolgende Tabelle aus und geben Sie je einen Vorteil und einen Nachteil der beiden Formulierungen an.

12

Solution

	Braunvieh	Edelschwein
Formulation		
Advantage		
Disadvantage		

- b) Genomic Selection has the potential to increase the genetic gain per year significantly. Currently beef cattle breeders start to introduce genomic selection for carcass traits. Let us assume the following selection parameters for the trait average daily gain. Compute the selection response per year for both scenarios with and without genomic selection. Based on the selection response per year which scenario is better? The trait is assumed to be normally distributed.

Die genomische Selektion hat das Potential den Zuchtfortschritt pro Jahr signifikant zu steigern. Aktuell beginnen die Fleischrinderzüchter mit der Einführung der genomischen Selektion für Fleischleistungsmerkmale. Für das Merkmal Tageszunahme gehen wir von den folgenden Selektionsparametern aus. Berechnen Sie den Selektionserfolg pro Jahr für die beiden Szenarien mit und ohne genomische Selektion. Welches der beiden Szenarien ist besser basierend auf dem Selektionsfortschritt pro Jahr? Für die Verteilung des Merkmals nehmen wir eine Normalverteilung an.

8

Parameter	No Genomic Selection	With Genomic Selection
proportion of animals selected	0.05	0.05
accuracy of predicted breeding values	0.85	0.65
additive genetic standard deviation	10.00	10.00
generation interval	7.00	2.00

Solution

- c) In cattle and in pigs, the structure of the breeding programs are different. What is the name of the structure of the different breeding programs. Please, specify a reason for the association of the structure of the breeding program to the different species.

In der Rinderzucht und der Schweinezucht ist die Struktur der Zuchtprogramme verschieden. Wie heissen diese Strukturen der Zuchtprogramme? Geben Sie je einen Grund an weshalb die beiden Spezies die entsprechende Struktur des Zuchtprogramms gewählt haben.

6

Solution

	Cattle	Pigs
Structure of breeding program		
Reason		

Problem 3: Variance Components Estimation

The following data set with average amounts of methane gas emission per year for three offspring per bull are given in the following table.

Im folgenden Datensatz sind die mittleren Mengen an Methanemissionen pro Jahr für je drei Nachkommen für fünf verschiedene Bullen gegeben.

Offspring	Bull	Methane
1	1	85
2	1	104
3	1	102
4	2	93
5	2	104
6	2	104
7	3	100
8	3	89
9	3	98
10	4	108
11	4	112
12	4	105
13	5	108
14	5	101
15	5	83

- a) Given that we want to reduce the amount of methane gas emitted in cattle using the tools of livestock breeding, we first have to do a variance components estimation. Why is it important to have a certain variability in a given trait and which variance component is important when we want to improve a certain trait with livestock breeding tools?

Wir möchten die Methanemission beim Rind mit den Werkzeugen der Tierzucht senken. Weshalb brauchen wir für die züchterische Bearbeitung eines Merkmals ein gewisse Variabilität im Merkmal und welche Varianzkomponente interessiert uns besonders?

2

Solution

- b) What is the linear model that would be used to allow us a separation of the different variance components in the given dataset on methane gas emission?

Wie sieht das lineare Modell aus, welches eine Aufteilung der Varianz in ihre verschiedenen Komponenten im Methandatensatz erlaubt?

11

Solution

- c) Use an analysis of variance (ANOVA) to estimate the variance components given in the above data set.
Schätzen Sie mit einer Varianzanalyse (ANOVA) die Varianzkomponenten aus dem oben angegebenen Datensatz zu den Methanemissionen.

8

Solution

Problem 4: Prediction of Breeding Values

We are using the same dataset as in Problem 3. The results of the variance components estimation were convincing that it is possible to reduce methane gas emission via selection. The next step is to predict breeding values.

Wir verwenden nochmals den gleichen Datensatz wie in Aufgabe 3. Die Resultate der Varianzkomponenten waren positiv in dem Sinne, dass eine züchterische Bearbeitung des Merkmals Methanemission möglich erscheint.

Offspring	Bull	Methane
1	1	85
2	1	104
3	1	102
4	2	93
5	2	104
6	2	104
7	3	100
8	3	89
9	3	98
10	4	108
11	4	112
12	4	105
13	5	108
14	5	101
15	5	83

- a) Specify a sire model by writing down the model formula, describing the meaning of all model components and indicating the expected values and the variances of all random effects in the model. We are using a general mean μ as the only fixed effect. We assume that all bulls are unrelated. Input the numeric information from the dataset into the model components

Spezifizieren Sie für die Zuchtwertschätzung ein Vatermodell. Geben Sie die Modellformel, beschreiben Sie die Bedeutung aller Modellkomponenten und geben Sie für alle zufälligen Komponenten im Modell den Erwartungswert und die Varianz an. Als einziger fixer Effekt wird ein allgemeines Mittel μ angenommen. Die Väter sind nicht miteinander verwandt. Setzen Sie die numerische Information aus dem Datensatz in die Modellkomponenten ein.

19

Solution

- b) Set up the mixed model equations for the model specified in 4a) to estimate the fixed effect and to predict the sire effects. We assume the sire variance to be $\sigma_s^2 = 4$ and the error variance to be $\sigma_e^2 = 72$.

Stellen Sie die Mischmodellgleichungen für das Modell aus 4a) auf und berechnen Sie die Schätzung des fixen Effekts und der Vatereffekte.

20

Solution