

ASMNW - Lösung 2

Peter von Rohr

03/12/2017

Aufgabe 1 - Begründung für Deregression

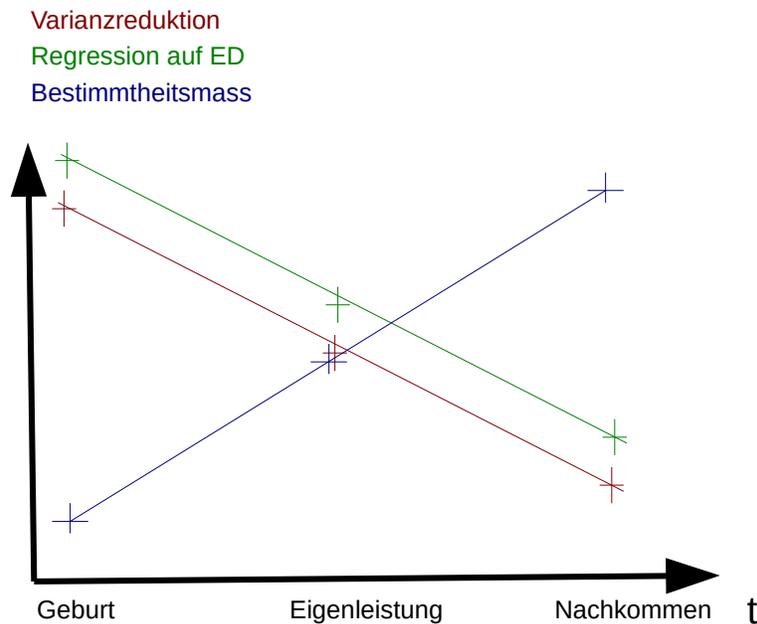
Nennen Sie die beiden Probleme, weshalb BLUP-Zuchtwerte korrigiert werden müssen, damit sie in der genomischen Selektion als Beobachtungen verwendet werden können.

Lösung

- Varianzreduktion, d.h. die Varianz der geschätzten Zuchtwerte \hat{g}_i ist kleiner oder gleich als die Varianz der wahren Zuchtwerte g_i .
- Regression des geschätzten Zuchtwertes \hat{g}_i auf den Durchschnitt der geschätzten Zuchtwerte der Eltern (ED)

Aufgabe 2 - Bestimmtheitsmass

Varianzreduktion und Regression des geschätzten Zuchtwertes auf den ED wird erst zum Problem, da das Ausmass dieser Effekte sich im Laufe des Lebens eines Tieres verändert. Zeichnen Sie im nachfolgenden Diagramm den Verlauf der Varianzreduktion, der Regression des geschätzten Zuchtwertes auf den ED und des Bestimmtheitsmasses ein.



Aufgabe 3 - Einfaches Beispiel

Für ein bestimmtes Merkmal sollen genomische Zuchtwerte geschätzt werden. Aus früheren Auswertungen wurde die genetische Varianz auf 25 geschätzt. Wir nehmen diesen Wert als eine Approximation für die wahre genetische Varianz $var(g)$. In der folgenden Tabelle sind für 6 die geschätzten Zuchtwerte (\hat{g}) und die Bestimmtheitsmasse (B) angegeben.

Tier	geschätzter Zuchtwert	Bestimmtheitsmasse
1	0.090	0.25
2	0.870	0.42
3	0.372	0.12
4	-0.668	0.19
5	1.527	0.49
6	-0.341	0.17

Ihre Aufgabe

Verwenden Sie die folgende vereinfachte Gleichung zur Deregression der geschätzten Zuchtwerte.

$$\hat{g}_D = \lambda \times \frac{1}{B} \times \hat{g}$$

Berechnen Sie die deregressierten Zuchtwerte \hat{g}_D und den Faktor λ . Bestimmen Sie λ so, dass keine Varianzreduktion mehr auftritt. Es muss also gelten, dass

$$var(\hat{g}_D) = var(g)$$

Lösung

Die geschätzten Zuchtwerte und die Bestimmtheitsmasse wurden vorgängig im Dataframe `dfBvRelTable` abgelegt. Der Einfachheit halber speichern wir diese in separaten Vektoren.

```
# Vektor der geschätzten Zuchtwerte
vecZw <- dfBvRelTable[,2]
# Vektor der Bestimmtheitsmasse
vecBe <- dfBvRelTable[,3]
```

Als erstes können wir für jedes Tier i den Quotienten aus geschätztem Zuchtwert \hat{g}_i geteilt durch das Bestimmtheitsmass B_i berechnen. Das Resultat davon ist dann wieder ein Vektor

```
(vecQuo <- vecZw / vecBe)
```

```
## [1] 0.360000 2.071429 3.100000 -3.515789 3.116327 -2.005882
```

Der Faktor λ wird über die Bedingung der konstanten Varianzen bestimmt. Es muss also gelten, dass

$$var(g) = var(\hat{g}_D) = var\left(\lambda \times \frac{1}{B} \times \hat{g}\right) = \lambda^2 \times var\left(\frac{1}{B} \times \hat{g}\right)$$

Die Varianz $var\left(\frac{1}{B} \times \hat{g}\right)$ können wir empirisch aus dem Vektor `vecQuo` berechnen. Der Wert für $var(g)$ ist in der Aufgabe gegeben. Somit können wir λ bestimmen als

$$\lambda = \sqrt{\frac{\text{var}(g)}{\text{var}(\frac{1}{B} \times \hat{g})}}$$

Da wir die Vorzeichen der geschätzten Zuchtwerte durch die Deregression nicht ändern dürfen, verwenden wir nur die positive Lösung für λ .

Die Berechnung des Zahlenwertes für λ lautet

```
(lambda <- sqrt(nGenVar/var(vecQuo)))
```

```
## [1] 1.801933
```

Somit können wir für jedes Tier den deregressierten geschätzten Zuchtwert angeben. Wir legen diese im Vektor `vecDereg` ab und fügen diese zur obigen Tabelle dazu.

```
(vecDereg <- lambda * vecQuo)
```

```
## [1] 0.6486959 3.7325757 5.5859926 -6.3352174 5.6154119 -3.6144658
```

Die Tabelle mit den Resultaten sieht dann wie folgt aus.

Tier	geschätzter Zuchtwert	Bestimmtheitsmass	Deregressierte Zuchtwerte
1	0.090	0.25	0.649
2	0.870	0.42	3.733
3	0.372	0.12	5.586
4	-0.668	0.19	-6.335
5	1.527	0.49	5.615
6	-0.341	0.17	-3.614

Zur Kontrolle überprüfen wir noch, dass es wirklich keine Varianzreduktion gegeben hat.

```
var(vecDereg)
```

```
## [1] 25
```

Durch Rundungsfehler bei der Berechnung hat es eine Abweichung zwischen den Varianzen gegeben. Diese beträgt

```
(abs(var(vecDereg)-nGenVar)/nGenVar)
```

```
## [1] 1.421085e-16
```