

Tierzucht und Selektionsindex

Peter von Rohr

2017-10-13

Fundamentale Fragen

1. Welches Tier ist das beste?
2. Wie sollen Tiere ausgewählt werden, dass ihre Nachkommen besser sind als die aktuelle Population

Bestes Tier

- ▶ “Bestes Tier”: relativer Begriff, abhängig von der Umwelt
- ▶ Beschreibung von Tieren basiert auf **Merkmalen**
- ▶ Beispiele für beobachtete Merkmale sind
 - ▶ Fellfarbe
 - ▶ Grösse
 - ▶ Bemuskelung
 - ▶ Fundament
 - ▶ Euter
- ▶ Beispiele für messbare Merkmale sind
 - ▶ Körpergewicht
 - ▶ Milchproduktion
 - ▶ Wachstum
 - ▶ Futteraufnahme

Genetische Veränderung

- ▶ Tierzucht beschäftigt sich mit längerfristigen genetischen Veränderungen von Populationen
- ▶ Genetisches Potential wird von einer Generation an die nächste weitergegeben
- ▶ Beobachtbar sind phänotypische Merkmale
- ▶ Zusammenhang zwischen Phänotyp y und Genotyp g

$$y = \mu + g + \epsilon$$

Zerlegung des Genotypischen Wertes

- ▶ Genotypischer Wert in Zuchtwert a , Dominanzabweichung d und Epistasie i

$$y = \mu + a + d + i + \epsilon$$

- ▶ Neugruppierung von d , i und ϵ in e

$$y = \mu + a + e$$

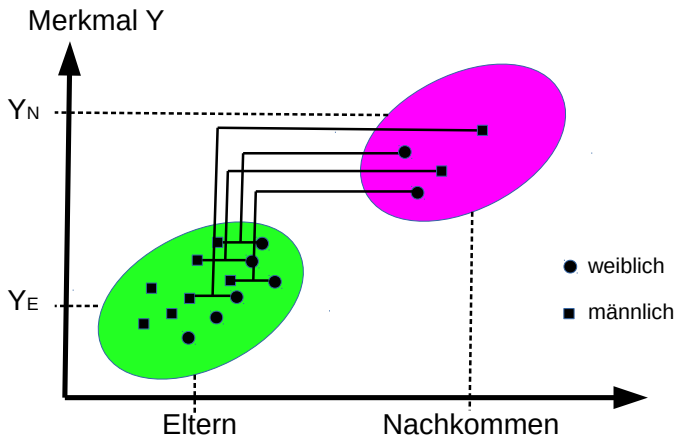
- ▶ additiver Effekt ist wichtig bei züchterischer Veränderung einer Population

Auswahl der Elterntiere

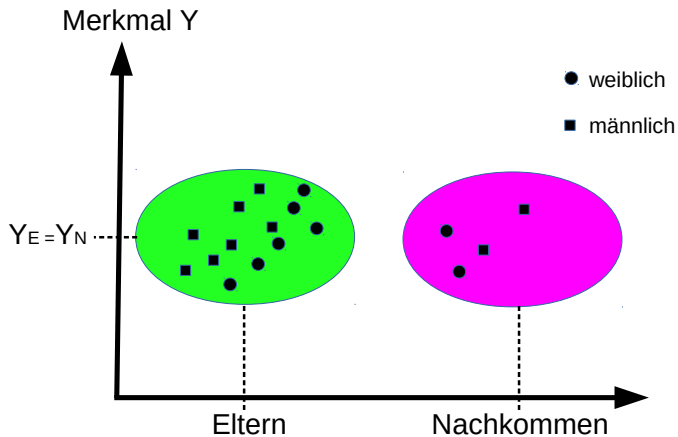
Zwei Werkzeuge

1. Selektion
2. Gezielte Anpaarung

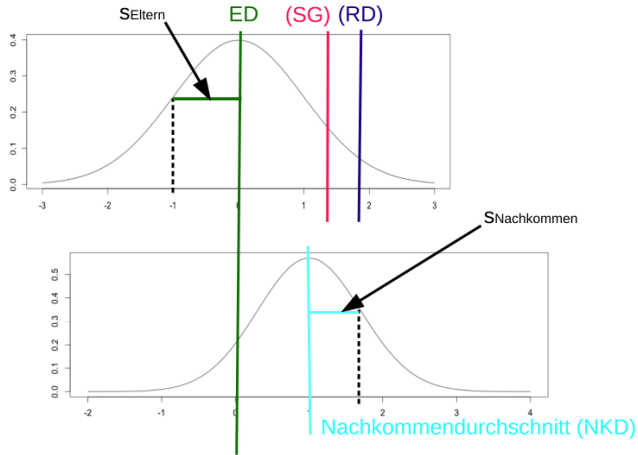
Selektion



Vermehrung



Resultat der gerichteten Selektion



ED: Elterndurchschnitt, SG: Selektionsgrenze, RD:
Remontendurchschnitt

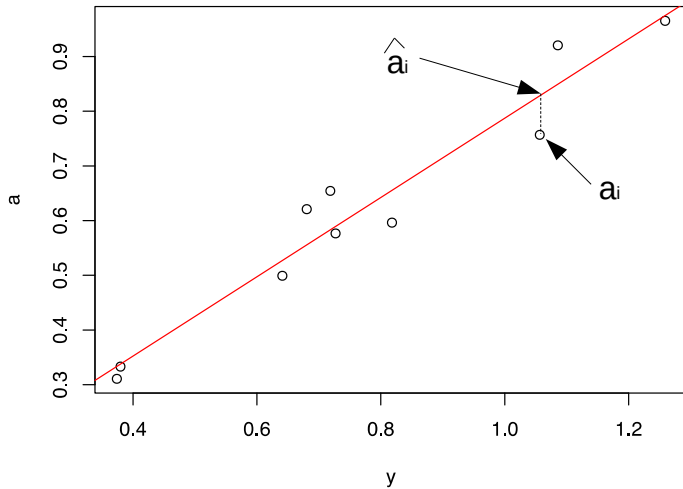
Gezielte Paarung

- ▶ verschieden von Selektion
- ▶ Zuordnung der Paarungspartner
- ▶ verschiedene Kriterien
 - ▶ Maximierung des Zuchtwertes
 - ▶ Ausnützung gegensätzlicher Eigenschaften
 - ▶ Erzielung von Heterosis-Effekten
- ▶ Paarungen innerhalb Rasse → **Reinzucht**
- ▶ Paarungen zwischen Rassen → **Kreuzungszucht**

Zuchtwertschätzung mit verschiedenen Informationsquellen

- ▶ Gegensatz: Quantitative Genetik - Berechnung der wahren Zuchtwerte aufgrund
 - ▶ Allelfrequenzen
 - ▶ Genotypischer Werte
- ▶ Hier: Schätzung der Zuchtwerte aufgrund von phänotypischen Beobachtungen
 - ▶ unterliegendes Modell
 - ▶ Schätzung basiert auf Regression der wahren Zuchtwerte auf phänotypische Beobachtungen
 - ▶ Schätzfehler

Regression



Einzelne Eigenleistung

- ▶ Modell:

$$y_{ij} = \mu_i + a_i + e_{ij}$$

- ▶ geschätzter Zuchtwert:

$$\hat{a}_i = b(y_i - \mu)$$

- ▶ Regressionskoeffizient:

$$b = \frac{\text{cov}(a, y)}{\text{var}(y)} = \frac{\text{var}(a)}{\text{var}(y)} = \frac{\sigma_a^2}{\sigma_y^2} = h^2$$

$$\rightarrow \hat{a}_i = h^2(y_i - \mu)$$

Einzelne Eigenleistung II

- ▶ Genauigkeit: Korrelation zwischen Selektionskriterium und wahren Zuchtwert

$$r_{a,y} = \frac{\text{cov}(a, y)}{\sigma_a \sigma_y} = \frac{\sigma_a^2}{\sigma_a \sigma_y} = h$$

- ▶ Erwarteter Selektionserfolg pro Generation: Verbesserung der Nachkommen im Vergleich zu den Eltern

$$R = i r_{a,y}^2 \sigma_y = i h^2 \sigma_y$$

wobei i die Selektionsintensität ist. Diese entspricht der Überlegenheit der selektierten Individuen ausgedrückt in Einheiten einer phänotypischen Standardabweichung.

Wiederholte Messungen

- ▶ mehrere wiederholte Messungen des gleichen Merkmals
- ▶ Modell:

$$\text{var}(y) = \text{var}(a) + \text{var}(pe) + \text{var}(te)$$

- ▶ Wiederholbarkeit:

$$t = \frac{\text{var}(a) + \text{var}(pe)}{\text{var}(y)}$$

- ▶ Zuchtwert \hat{a}_i aufgrund des Mittelwertes \bar{y}

$$\hat{a}_i = b(\bar{y}_i - \mu)$$

Wiederholte Messungen II

- ▶ Regressionskoeffizient b

$$b = \frac{\text{cov}(a, \bar{y})}{\text{var}(\bar{y})}$$

- ▶ Zähler Covarianz

$$\text{cov}(a, \bar{y}) = \text{cov}(a, a + pe + \frac{1}{n} \sum te) = \sigma_a^2$$

- ▶ Nenner Varianz

$$\text{var}(\bar{y}) = \text{var}(a) + \text{var}(pe) + \frac{1}{n} \text{var}(te) = [t + (1 - t)/n] \sigma_y^2$$

Wiederholte Messungen III

- ▶ Einsetzen in Regressionskoeffizienten

$$b = \frac{\text{cov}(a, \bar{y})}{\text{var}(\bar{y})} = \frac{\sigma_a^2}{[t + (1 - t)/n] \sigma_y^2} = \frac{nh^2}{1 + (n - 1)t}$$

- ▶ Einsetzen in geschätzten Zuchtwert

$$\hat{a}_i = b(\bar{y}_i - \mu) = \frac{nh^2}{1 + (n - 1)t} (\bar{y}_i - \mu)$$

Wiederholte Messungen IV

- ▶ Genauigkeit

$$\begin{aligned}r_{a,\bar{y}} &= \frac{\text{cov}(a, \bar{y})}{\sigma_a \sigma_{\bar{y}}} = \frac{\sigma_a^2}{\sigma_a \sigma_y \sqrt{[t + (1-t)/n]}} \\ &= \frac{h\sqrt{n}}{\sqrt{(1 + (n-1)t)}} = \sqrt{\frac{nh^2}{1 + (n-1)t}} = \sqrt{b}\end{aligned}$$

- ▶ Zuwachs an Genauigkeit

$$\frac{r_{a,\bar{y}}}{r_{a,y}} = \frac{\sqrt{\frac{nh^2}{1+(n-1)t}}}{h} = \sqrt{\frac{n}{1 + (n-1)t}}$$

Nachkommen

- ▶ \tilde{y}_s : Mittelwert von n Nachkommen von Stier s
- ▶ Zuchtwert für Stier s :

$$\hat{a}_s = b(\tilde{y}_s - \mu)$$

- ▶ Regressionskoeffizient:

$$b = \frac{\text{cov}(a, \tilde{y})}{\text{var}(\tilde{y})}$$

Nachkommen II

- ▶ Zerlegung:

$$\tilde{y}_s = \mu + \frac{1}{n} \sum_i a_i + \frac{1}{n} \sum_i e_i$$

wobei: $a_i = \frac{1}{2}(a_s + a_{d(i)}) + m_i$

- ▶ s und d nicht verwandt, a_s ist konstant und $\sum_i m_i = 0$

$$\tilde{y}_s = \mu + \frac{1}{2}a_s + \frac{1}{n} \sum_i a_{d(i)}/2 + \frac{1}{n} \sum_i e_i$$

Nachkommen III

- ▶ Kovarianz:

$$\begin{aligned} \text{cov}(a, \tilde{y}) &= \text{cov}\left(a, \mu + \frac{1}{2}a_s + \frac{1}{n} \sum_i a_{d(i)}/2 + \frac{1}{n} \sum_i e_i\right) \\ &= \text{cov}\left(a, \frac{1}{2}a_s\right) = \frac{1}{2} \text{cov}(a, a_s) = \frac{1}{2} \text{var}(a) \end{aligned}$$

- ▶ Varianz:

$$\text{var}(\tilde{y}) = [t + (1 - t)/n] \sigma_y^2$$

$$\text{mit } t = \frac{0.25\sigma_a^2}{\sigma_y^2} = \frac{1}{4}h^2$$

Nachkommen IV: Regressionskoeffizient

$$\begin{aligned} b &= \frac{\frac{1}{2}\sigma_a^2}{[t + (1 - t)/n] \sigma_y^2} \\ &= \frac{\frac{1}{2}h^2\sigma_y^2}{\left[\frac{1}{4}h^2 + (1 - \frac{1}{4}h^2)/n\right] \sigma_y^2} \\ &= \frac{\frac{1}{2}nh^2}{\frac{1}{4}nh^2 + (1 - \frac{1}{4}h^2)} \\ &= \frac{2nh^2}{nh^2 + 4 - h^2} \\ &= \frac{2n}{n + \frac{(4-h^2)}{h^2}} \\ &= \frac{2n}{n + k} \end{aligned}$$

wobei $k = \frac{(4-h^2)}{h^2}$

Nachkommen V: Zuchtwert und Genauigkeit

$$\hat{a}_s = \frac{2n}{n+k} (\tilde{y}_s - \mu)$$

$$\begin{aligned} r_{a,\tilde{y}} &= \frac{\text{cov}(a, \tilde{y})}{\sqrt{\text{var}(a) \text{var}(\tilde{y})}} \\ &= \frac{\frac{1}{2} \text{var}(a)}{\sqrt{\text{var}(a) \left[\frac{1}{4} h^2 + \left(1 - \frac{1}{4} h^2\right) / n \right] \sigma_y^2}} \\ &= \frac{\frac{1}{2} h^2 \sigma_y^2}{\sqrt{h^2 \sigma_y^2 \left[\frac{1}{4} h^2 + \left(1 - \frac{1}{4} h^2\right) / n \right] \sigma_y^2}} \\ &= \sqrt{\frac{nh^2}{nh^2 + (4 - h^2)}} \\ &= \sqrt{\frac{n}{n+k}} \end{aligned}$$

Selektionsindex

- ▶ Früher: Zuchtwerte eines Merkmals aufgrund verschiedener phänotypischer Informationsquellen

$$I_i = \hat{a}_i = b_1(y_1 - \mu_1) + b_2(y_2 - \mu_2) + b_3(y_3 - \mu_3) + \dots$$

- ▶ Heute: Schätzung des Gesamtzuchtwertes

Gesamtzuchtwert

- ▶ mathematische Formulierung des Zuchtziels

$$H = w_1 * a_1 + w_2 * a_2 + w_3 * a_3 + \dots$$

- ▶ Schätzung mit

$$I = \hat{H} = b_1 * x_1 + b_2 * x_2 + \dots$$

wobei x_i verfügbare Informationen, meist mit BLUP geschätzte Zuchtwerte.

Indexgleichung

- ▶ Aufgrund der Anforderung, dass I den Gesamtzuchtwert H möglichst gut schätzen soll, folgt die **Indexgleichung**

$$Pb = Gw$$

wobei P die Co-Varianzmatrix der Informationen im Index und G die genetische Co-Varianzmatrix zwischen Informationen im Index und den Merkmalen im Gesamtzuchtwert

- ▶ Bestimmung der Gewichte b als

$$b = P^{-1}Gw$$