



LOUISIANA STATE UNIVERSITY
College of Agriculture
School of Plant, Environmental, and Soil Sciences
AGRO 7075 Prediction-based Breeding



Response to Selection

Prof. Roberto Fritsche-Neto

rfneto@agcenter.lsu.edu

Baton Rouge, Feb 13th, 2023

Breeding value

- The breeding value (a) of a progeny i is the mean of its parents' plus the mendelian sampling

$$E(a_i) = \frac{1}{2}a_{p1} + \frac{1}{2}a_{p2} + m_i \qquad E(\bar{a}) = \frac{1}{2}a_{p1} + \frac{1}{2}a_{p2}$$

- It is key for deciding between possible breeding schemes
- Therefore, the expected response from the current to the next generation is entirely based on the genetic superiority of the selected parents

$$E(RS) = \frac{1}{2}(S_m + S_f)$$

- Based on the standard regression theory we have

$$b_{yx} = \frac{\sigma_{xy}}{V_x} = r_{xy} \frac{\sigma_y}{\sigma_x}$$

- Thus, we can predict the breeding value given an estimate which we will call the index value P

- P is the criteria for selection

$$\bar{a}^* = \bar{a} + b_{aP}(P^* - \bar{P})$$

$$RS = \bar{a}^* - \bar{a} = b_{aP}(P - \bar{P})$$

The Breeder's equation

- Using the deviation of the index values of selected values from the mean index value of all individuals in the population in standard deviations units

$$i = (P^* - \bar{P}) / \sigma_P$$

$$(P^* - \bar{P}) = i \cdot \sigma_P$$

$$b_{aP} = r_{aP} \frac{\sigma_a}{\sigma_P}$$

$$r_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} = \frac{\sigma_a^2}{\sigma_a \sigma_P} = \frac{\sigma_a}{\sigma_P} = h$$

$$RS = b_{aP} (P^* - \bar{P})$$

$$RS = b_{aP} \cdot i \cdot \sigma_P$$

$$RS = r_{aP} \frac{\sigma_a}{\sigma_P} i \cdot \sigma_P$$

$$RS = r_{aP} \cdot \sigma_a \cdot i$$

$$RS = i \cdot h \cdot \sigma_a$$

$$i = z / p \quad z = \frac{e^{-1/2x^2}}{\sqrt{2\pi}} \quad p^* = \frac{s + 1/2}{n + \frac{s}{2n}}$$

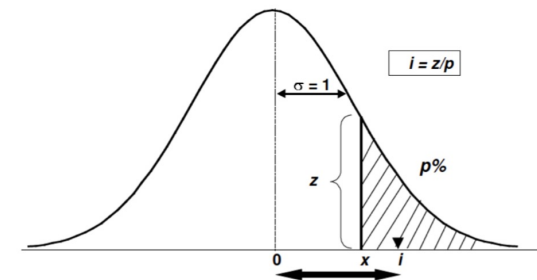
- p** for small populations

- where s is the number selected and n is the population size

- On R, $i = dnorm(qnorm(p))/p$

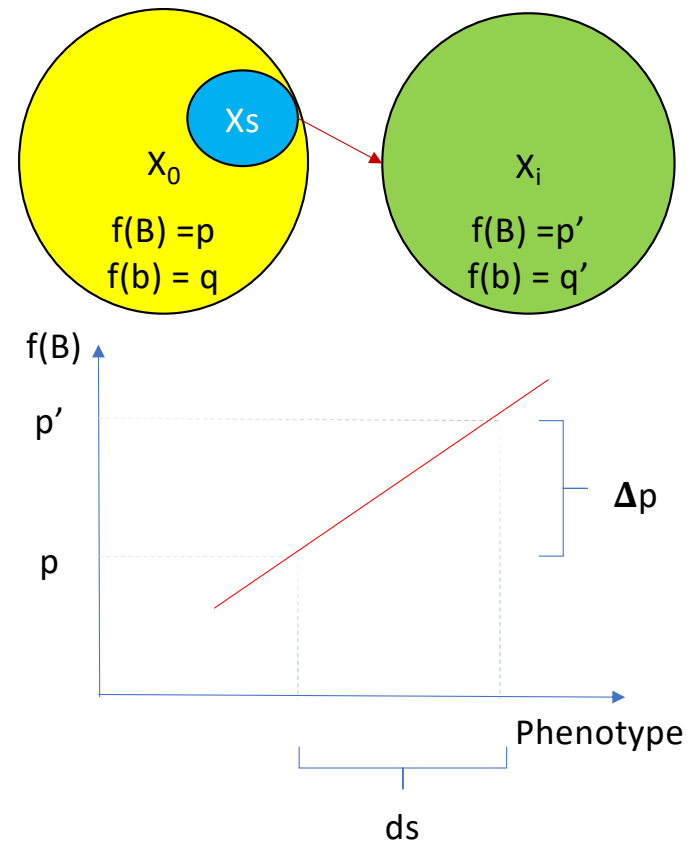
- Response per unit time**

$$RS = \frac{i \cdot r_{aP} \cdot \sigma_a}{T}$$



Effect of selection on allele frequencies

- $p + q = 1$
- $p' + q' = 1$
- $p' = p + \Delta p$
- $q' = q - \Delta p$
- $\Delta p = \Delta q$
- $(p' - p) = (u_s - u_0)b_{Pf(B)}$
- $\Delta p = ds \cdot b_{Pf(B)}$
- $b_{Pf(B)} = \text{COV}[P, f(B)] / V_P$
- $\text{COV}[P, f(B)] = \text{COV}[G + E, f(B)]$
- $= \text{COV}[G + E, f(B)]$
- $= \text{COV}[G, f(B)] + \text{COV}[E, f(B)]$
- $= \text{COV}[G, f(B)]$



Effect of selection on allele frequencies

- $\text{COV}(x,y) = \sum_i f_i x_i y_i - (\sum_i f_i x_i)(\sum_i f_i y_i)$
- $\sum_i f_i y_i = (p - q)a + 2pqd$
- $\sum_i f_i x_i = p^2(1) + 2pq(1/2) + q^2(0)$
- $= p^2 + pq$
- $= p^2 + p(1 - p)$
- $= p^2 + p - p^2 = p$
- $\sum_i f_i x_i y_i = p^2(1)(a) + 2pq(1/2)d + q^2(0)0$
- $= p^2a + pqd$
- $\text{COV}(x,y) = p^2a + pqd - p[(p - q)a + 2pqd]$
- $= p^2a + pqd - p^2a - pqa + 2p^2d$
- $= pqd - pqa + 2p^2d$
- $= pq[a + (1 - 2p)d]$
- $= pq[a + (q - p)d]$
- $= pq\alpha$

| Genotype | f | VG | $f(B)$ |
|----------|-------|------|--------|
| BB | p^2 | a | 1 |
| Bb | $2pq$ | d | 1/2 |
| bb | q^2 | $-a$ | 0 |

- $\Delta p = ds \cdot b_{Pf(B)}$
- $b_{Pf(B)} = \text{COV}[P, f(B)] / V_P$

$$\Delta p = \frac{ds \cdot p \cdot q \cdot \alpha}{\sigma_p^2}$$

$$\Delta p = \frac{i \cdot p \cdot q \cdot \alpha}{\sigma_p}$$

$$(P^* - \bar{P}) = i \cdot \sigma_p$$

Effect of selection on population mean

- $u_{c_0} = u + (p - q)a + 2pqd$
- $u_{c_1} = u + (p' - q')a + 2p'q'd$

- $p' = p + \Delta p$
- $q' = q - \Delta p$
- $\Delta p^2 \approx 0$

- $u_{c_1} = u + (p' - q')a + 2p'q'd$
- $= u + ((p + \Delta p) - (q - \Delta p))a + 2d(p + \Delta p)(q - \Delta p)$
- $= u + (p + \Delta p - q)a + 2pqd + 2qd\Delta p - 2p\Delta p d - 2\Delta^2 pd$
- $= u + (p - q)a + 2\Delta pa + 2pqd + 2d\Delta p(q - p)$
- $= u + (p - q)a + 2pqd + 2\Delta pa + 2d\Delta p(q - p)$
- $= u_0 + 2\Delta pa + 2d\Delta p(q - p)$
- $= u_0 + 2\Delta p[a + d(q - p)]$
- $= u_0 + 2\Delta p\alpha$
- **Improvement in the population mean**

| Genotype | C_0 | C_1 | VG |
|----------|-------|---------|------|
| BB | p^2 | p'^2 | a |
| Bb | $2pq$ | $2p'q'$ | d |
| bb | q^2 | q'^2 | $-a$ |

- $\Delta u = u_{c_1} - u_{c_0} = 2\Delta p\alpha$

$$\Delta p = \frac{i \cdot p \cdot q \cdot \alpha}{\sigma_P}$$

$$\Delta u = 2 \frac{i \cdot p \cdot q \cdot \alpha}{\sigma_P} \alpha$$

$$\Delta u = i \frac{2pq\alpha^2}{\sigma_P} \quad RS = i \frac{\sigma_a^2}{\sigma_P}$$

$$RS = i \cdot h \cdot \sigma_a$$

Response to Selection – parents to offspring

- $V_x = V_p$

- $RS = ds \cdot b_{xy} \quad i \cdot \sigma_P = (P^* - \bar{P})$

$$RS = i \cdot \sigma_P \cdot \frac{COV(po)}{\sigma_P^2}$$

$$b_{po} = \frac{COV(po)}{\sigma_P^2}$$

$$RS = \frac{i}{\sigma_P} cov_{po}$$

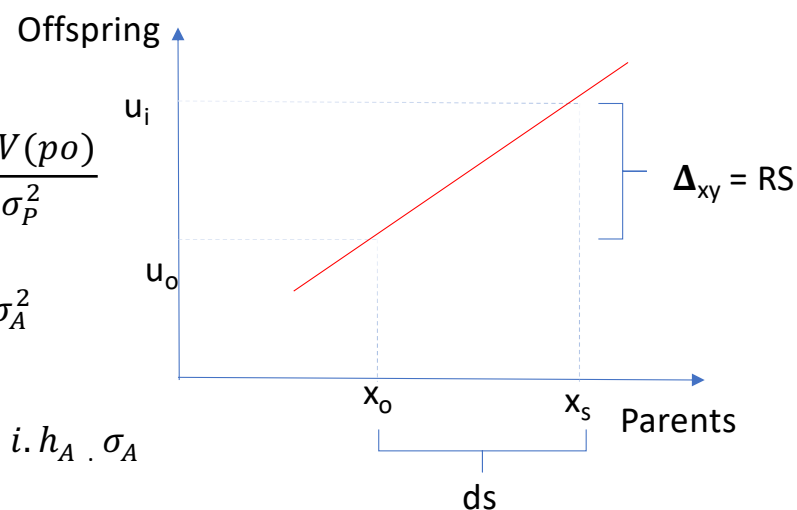
$$RS = \frac{i}{\sigma_P} \cdot \frac{1}{2} \cdot \sigma_A^2$$

$$RS = \frac{i}{\sigma_P} \cdot \frac{1}{2} \cdot \sigma_A^2 + \frac{i}{\sigma_P} \cdot \frac{1}{2} \cdot \sigma_A^2$$

$$RS = \frac{i}{\sigma_P} \cdot \sigma_A^2$$

$$RS = \frac{i}{\sigma_P} \cdot \sigma_A \cdot \sigma_A$$

$$RS = i \cdot h_A \cdot \sigma_A$$



- $F = 0$

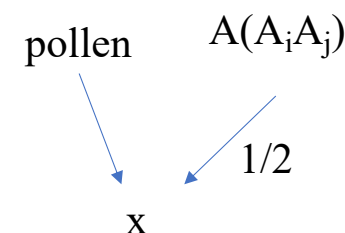
- $COV(p, o) = 2 \cdot f_{xy} \cdot V_a + u_{xy} \cdot V_d$

- $f_{xy} = \frac{1}{4} [P(x \equiv A \equiv A_i) + P(x \equiv A \equiv A_j) + P(x \equiv A_i \equiv A \equiv A_j) + P(x \equiv A_j \equiv A \equiv A_i)]$

- $f_{po} = \frac{1}{4} [1/2 + 1/2 + 0 + 0] = \frac{1}{4}$

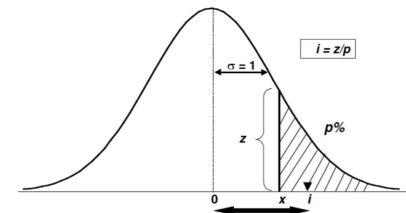
- $COV(p, o) = 2 \cdot \frac{1}{4} V_a + 0 \cdot V_d$

- $COV(p, o) = \frac{1}{2} \cdot V_a$



Drift and Selection working together

- Frequencies after selection
- $f(B)' = p' = p + \Delta p$
- $f(b)' = q' = q - \Delta p$
- Frequencies after selection + drift (**too small sample**)
- $E(p') = p + \Delta p + Sp$
- $E(q') = q - \Delta p - Sp$
- $E(Sp) = E(Sq) = 0$



- Sp is the variance among samples $Sp = \frac{p'q'}{2Ne}$
- $= (p + \Delta p) \cdot (q - \Delta p) / 2Ne$
- $= (pq + \Delta p^2 - p\Delta p + q\Delta p) / 2Ne$
- $= (pq + \Delta p(q - p)) / 2Ne$
- $= pq/2Ne + \Delta p(q - p)/2Ne$
- **drift per se** + **drift due to selection**

- $uc_0 = u + (p - q)a + 2pqd$
- $uc_1 = u + (p' - q')a + 2p'q'd$

$$\Delta p = i \frac{pq\alpha}{\sigma_p}$$

$$\sigma_A^2 = 2pq\alpha^2$$

$$ID = -2pqd$$

$$D1 = 2pq(q - p)\alpha d$$

$$\Delta F = \frac{1}{2Ne}$$

Drift and Selection working together

- $uc_1 = u + (p' - q')a + 2p'q'd$
- $= u + ((p + \Delta p + Sp) - (q - \Delta p - Sp))a + 2(p + \Delta p + Sp)(q - \Delta p + Sp)d$
- $= u + p + \Delta p + Sp - q + \Delta p + Sp)a + 2d(pq + p\Delta p + pSp + q\Delta p - \Delta^2 p - \Delta pSp + qSp - \Delta pSp - Sp^2)$
- $= u + (p - q)a + 2\Delta pa + Spa + 2pqd + 2dp\Delta p + 2dpSp + 2dq\Delta p - 2d\Delta pSp + 2dqSp - 2d\Delta pSp - 2dSp^2)$
- Since $E(Sp) = 0$
- $= u + (p - q)a + 2\Delta pa + 2pqd + 2dp\Delta p + 2dq\Delta p - 2dSp^2$
- $= 2\Delta p[a + d(q - p)] - 2dSp^2$
- $= u_0 + 2\Delta p\alpha - 2\Delta p(q - p)/2Ne - 2pqd/2Ne$

$$\Delta x = u_1 - u_0 = i \frac{2pq\alpha^2}{\sigma_p} + \frac{i}{\sigma_p} \cdot \frac{2pq(q - p)\alpha d}{2Ne} - \frac{2pqd}{2Ne}$$

$$\Delta x = \frac{i\sigma_A^2}{\sigma_p} + \frac{iD1}{\sigma_p 2Ne} - \frac{ID}{2Ne}$$

$$\Delta x = \frac{i}{\sigma_p} \left(\sigma_A^2 + \frac{D1}{2Ne} \right) - \frac{ID}{2Ne}$$

$$RS = \frac{i}{\sigma_p} (\sigma_A^2 + \Delta F D1) - \Delta F ID$$

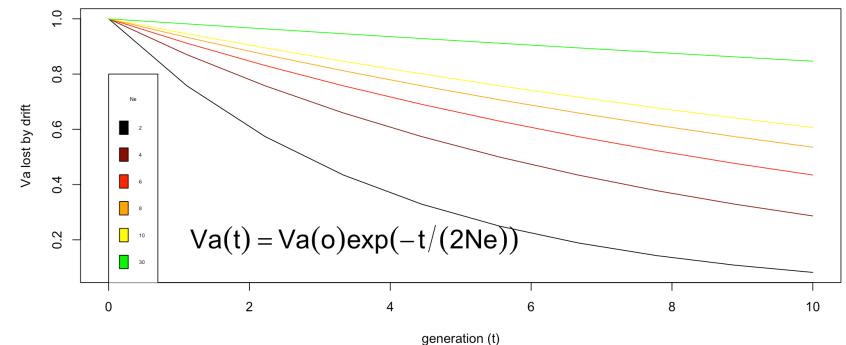
$$\Delta p = i \frac{pq\alpha}{\sigma_p}$$

$$\sigma_A^2 = 2pq\alpha^2$$

$$ID = -2pqd$$

$$D1 = 2pq(q - p)\alpha d$$

$$\Delta F = \frac{1}{2Ne}$$



Indirect response to selection

- Pleiotropy and linked genes
- Hardly or costly to evaluate
- The second (t) trait should be high heritable and correlated with the first (z)
- $Vg_z = Va_z + Vd_z$
- $Vg_t = Va_t + Vd_t$
- $Va = 2pq[a + (q - p)d]^2$
- $Vd = (2pqd)^2$
- $COVg_{(z,t)} = COVa_{(z,t)} + COVd_{(z,t)}$

• **F = 0**

- $COVg_{(z,t)} = 2f_{xy}COVa_{(z,t)} + u_{xy}COVd_{(z,t)}$
- $COVa = 2pq[a_z + (q - p)d_z][a_t + (q - p)d_t]$
- $COVd = 4p^2q^2d_t^2d_z^2$

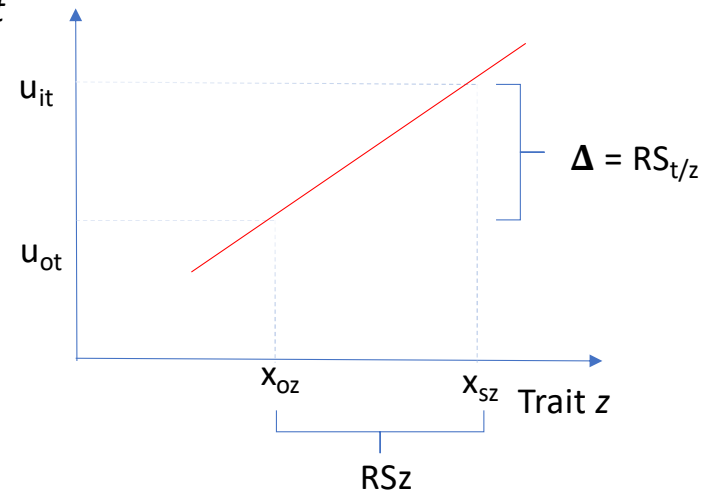
$$RS(t, z) = RS_z \cdot b_{zt}$$

$$RS_z = i \frac{\sigma_{Gz}^2}{\sigma_{Pz}}$$

$$b_{zt} = \frac{COV_{G(z,t)}}{\sigma_{Gz}^2}$$

$$r_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

Trait t



$$RS(t, z) = i \frac{\sigma_{Gz}^2}{\sigma_{Pz}} \frac{COV_{G(z,t)}}{\sigma_{Gz}^2}$$

$$RS(t, z) = i \frac{COV_{G(z,t)}}{\sigma_{Pz}}$$

$$RS(t, z) = i \frac{r_{g(z,t)} \sigma_{gz} \sigma_{gt}}{\sigma_{Pz}}$$

$$RS(t, z) = i r_{g(z,t)} h_z \sigma_{gt}$$

Genomic Selection in(direct) response to selection

- The Breeder's equation

$$RS = \frac{i \cdot r_a \cdot \sigma_a}{T}$$

- Predictive Ability vs. Accuracy

$$RS(p, m) = RSm \cdot b_{pm}$$

$$COV_{(\hat{y}, y)} = COV_{(g_m, g+e)} = COV_{(g_m, g)} + COV_{(g_m, e)} = \sigma_g^2$$

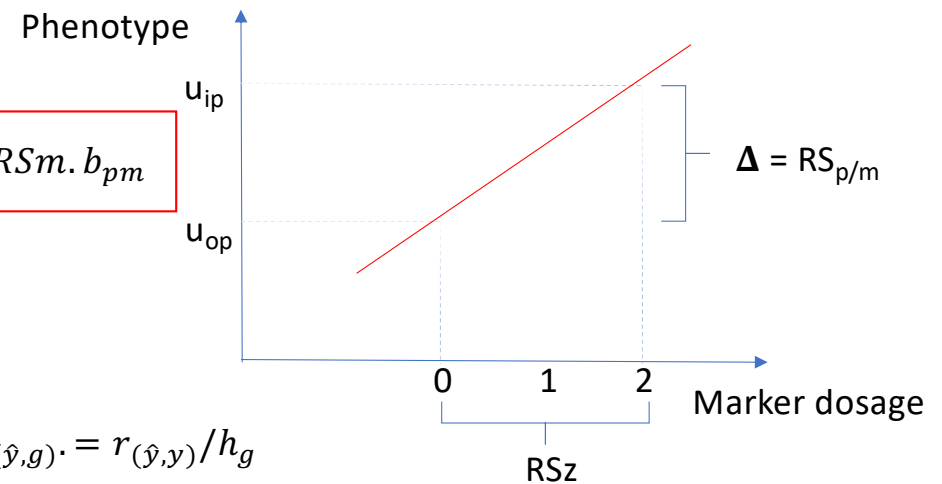
$$V_{(\hat{y})} = V_{(g_m)} = \sigma_{g_m}^2 \quad V_{(y)} = Vg + Ve = \sigma_g^2 / h_g^2$$

$$r_{(\hat{y}, y)} = \frac{COV_{(\hat{y}, y)}}{\sigma_{\hat{y}} \sigma_y} = \frac{\sigma_g^2}{\sigma_{\hat{y}} \cdot \sigma_g / h_g} = \frac{\sigma_g^2}{\sigma_{\hat{y}} \sigma_g} \cdot h_g = r_{(\hat{y}, g)} \cdot h_g \quad r_{(\hat{y}, g)} = r_{(\hat{y}, y)} / h_g$$

- Response to selection in GS

$$RS(p, m) = i \frac{\sigma_{Gm}^2}{\sigma_{Pm}} \cdot \frac{COV_{G(p, m)}}{\sigma_{Gm}^2} \quad RS(p, m) = i \frac{COV_{G(p, m)}}{\sigma_{Pm}} \quad RS(p, m) = i \frac{r_{g(p, m)} \sigma_{gp} \sigma_{gm}}{\sigma_{Pm}} \quad RS(p, m) = i r_{g(p, m)} h_m \sigma_{gP}$$

$$b_{yx} = \frac{COV_{(x, y)}}{\sigma_x^2} \quad r_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \quad RS(p, m) = i r_{g(p, m)} \sigma_{gP}$$



Genomic Selection in(direct) response to selection

- Let's see why

$$V_{(y)} = Vg + Ve = \sigma_g^2 / h_g^2$$

- $E(\text{EBV}) = b_{g,y} \cdot y$
- $V(\text{cX}) = c^2 V(X)$

$$\sigma_{EBV}^2 = b_{g,y}^2 \sigma_y^2$$

$$\sigma_y^2 = \frac{\sigma_{EBV}^2}{b_{g,y}^2}$$

$$b_{g,y} = \frac{COV_{(g,y)}}{\sigma_y^2}$$

$$b_{g,y} = r_{g,y} = \frac{\sigma_{g,y}}{\sigma_g \sigma_y} = \frac{\sigma_g^2}{\sigma_g \cdot \sigma_y} = \frac{\sigma_g}{\sigma_y} = h_g$$

$$\sigma_y^2 = \frac{\sigma_g^2}{h_g^2}$$

Is there a limit for RS?

- What is the RS over many cycles of selection?
- It is the sum of all cycles

- What is the maximum of RS?

$$RS = 2 \cdot Ne \cdot i \cdot r_{aP} \cdot \sigma_a$$

- Part of the observed variation in RS is due to drift

$$RS = t \cdot \frac{\sigma_a^2}{Ne}$$

- **Real Limit and asymmetric of RS**

- In real experiments the the RS limit is further that the theory
- Epistasis, GE, linkage phase, major genes, drift, ID, i, biological limits...

- **Illinois long-term corn selection experiments, started in 1896**

- Stopped for low-oil after 87 generations due to nearly zero V_a (RS = 22sdA)
- However, the high-oil still showing V_a (RS > 26sdA)

