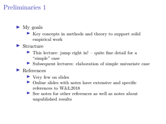
Analysis of univariate phenotypic selection

Michael Morrissey
February 3, 2020

Preliminaries 1

- My goals
  - Key concepts in methods and theory to support solid empirical work
- Structure
  - This lecture: jump right in! – quite fine detail for a “simple” case
  - Subsequent lectures: elaboration of simple univariate case
- References
  - Very few on slides
  - Online slides with notes have extensive and specific references to W&L2018
  - See notes for other references as well as notes about unpublished results



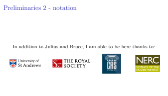
Preliminaries 1

- My goals
  - Key concepts in methods and theory to support solid empirical work
- Structure
  - This lecture: jump right in! – quite fine detail for a “simple” case
  - Subsequent lectures: elaboration of simple univariate case
- References
  - Very few on slides
  - Online slides with notes have extensive and specific references to W&L2018
  - See notes for other references as well as notes about unpublished results

Preliminaries 2 - notation

In addition to Julius and Bruce, I am able to be here thanks to:

University of St Andrews | THE ROYAL SOCIETY | CAS OSLO | NERC SCIENCE OF THE ENVIRONMENT



Preliminaries 2 - notation

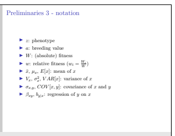
In addition to Julius and Bruce, I am able to be here thanks to:

Preliminaries 3 - notation

- z: phenotype
- a: breeding value
- W: (absolute) fitness
- w: relative fitness ($w_i = \frac{W_i}{\bar{W}}$)
- $\bar{x}$, μ_x, $E[x]$: mean of x
- V_x, σ^2_x, $VAR[x]$: variance of x
- $\sigma_{x,y}$, $COV[x,y]$: covariance of x and y
- β_{xy}, $b_{y|x}$: regression of y on x

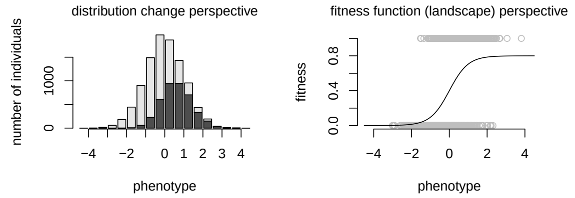
Analysis of univariate phenotypic selection

└Preliminaries 3 - notation



Changes and slopes

Two complimentary ways of thinking about natural selection:

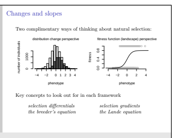


Key concepts to look out for in each framework

selection differentials	*selection gradients*
the breeder's equation	*the Lande equation*

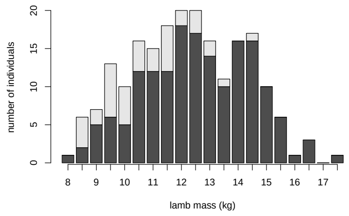
Analysis of univariate phenotypic selection

└Changes and slopes



The change in the mean, within a generation

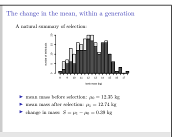
A natural summary of selection:



- mean mass before selection: $\mu_0 = 12.35$ kg
- mean mass after selection: $\mu_1 = 12.74$ kg
- change in mass: $S = \mu_1 - \mu_0 = 0.39$ kg

Analysis of univariate phenotypic selection

└The change in the mean, within a generation



W&L2018 pp. 482, 1107, and onwards

The selection differential's justification

Justification comes from the mechanics of evolution

$$evolution = f(genetics,\ selection)$$

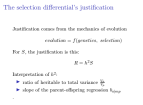
For S, the justification is this:

$$R = h^2 S$$

Interpretation of h^2:

- ratio of heritable to total variance $\frac{V_a}{V_p}$
- slope of the parent-offspring regression $b_{o|mp}$

.



To see how, and under what conditions, these are equal, note that the covariance of (single) parent and offspring is $\frac{1}{2}V_a$. The covariance of mid-parrents and offsrping is thus V_a. The variance in parents is V_p, so from the definition of a regression (we shall focus on it soon when it becomes immediatley relevant to selection coefficeints, the slope of the offspring on mid-parent regression is $\frac{V_a}{V_p}$

Derivation of the Breeder's equation

By construction, the regression of offspring phenotype on mid-parent phenotype is a function that predicts offspring phenotype according to

$$z_o = \mu + b_{o|mp}(z_{mp} - \mu) + e$$

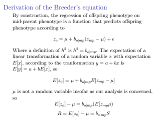
Where a definition of h^2 is $h^2 = b_{o|mp}$. The expectation of a linear transformation of a random variable x with expectation $E[x]$, according to the tranformation $y = a + bx$ is $E[y] = a + bE[x]$, so

$$E[z_o] = \mu + b_{o|mp}E[z_{mp} - \mu]$$

μ is not a random variable insofar as our analysis is concerned, so

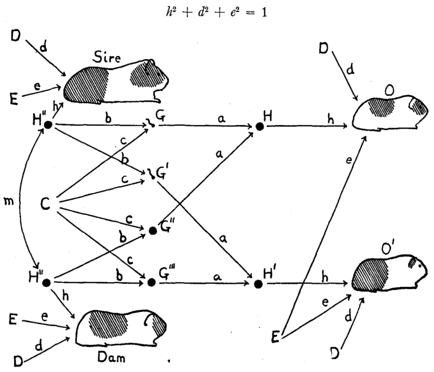
$$E[z_o] - \mu = b_{o|mp}(E[z_{mp}\mu)$$

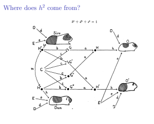
$$R = E[z_o] - \mu = b_{o|mp}S$$



I'm not going to justify all other relationships between selection, genetics, and evolution to the same extent, but I want to do this one because this justification is the ultimate basis the entire selection coefficient concept. Without the grounding in evolutionary QG, we might as well just measure correlations of everything.This derivation of the breeder's equation is given in W&L2018 pp. 482

Where does h^2 come from?





Where does h^2 come from, and why is it squared (i.e., why is the symbol squared, when h the square root of h^2 doesn't seem to ever come up in quantitative genetics? The answer is kinda cool!S. Wright. 1921. Systems of mating. I. The biometric relations between parent and offspring. Genetics 6:111-123.

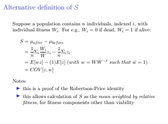
Alternative definition of S

Suppose a population contains n individuals, indexed i, with individual fitness W_i. For e.g., $W_i = 0$ if dead, $W_i = 1$ if alive:

$$\begin{aligned} S &= \mu_{after} - \mu_{before} \\ &= \frac{1}{n}\Sigma_i \frac{W_i}{\bar{W}} z_i - \frac{1}{n}\Sigma_i z_i \\ &= E[wz] - (1)E[z] \ (with\ w = W\bar{W}^{-1}\ such\ that\ \bar{w} = 1) \\ &= COV[z, w] \end{aligned}$$

Notes:

- this is a proof of the Robertson-Price identity
- this allows calculation of S as the *mean weighted by relative fitness*, for fitness components other than viability



Not just an alternative, one of the most important and useful relationships in evolutionary biology.

Applied to phenotypes, it is another way of calculating the selection differential. Applied to breeding values, it says what the mean breeding value is, among individuals that pass on their genes. This directly tells us how much evolution to expect (if only we could easily directly measure breeding values!), and is the secondary theorem of selection, which we will encounter later.

Background on OLS regression

If the covariance of A and B is σ_{AB} and the variance of A is σ^2_A then the regression of B on A is given by

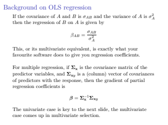
$$\beta_{AB} = \frac{\sigma_{AB}}{\sigma^2_A}$$

This, or its multivariate equivalent, is exactly what your favourite software does to give you regression coefficients.

For multiple regression, if $\mathbf{\Sigma_x}$ is the covariance matrix of the predictor variables, and $\mathbf{\Sigma_{x}}_y$ is a (column) vector of covariances of predictors with the response, then the gradient of partial regression coefficients is

$$\boldsymbol{\beta} = \mathbf{\Sigma_x^{-1}\Sigma_{x}}_y$$

The univariate case is key to the next slide, the multivariate case comes up in multivariate selection.



A general stats interlude: not ignorable though!

The selection gradient: another selection coefficient

Recall that

$$R = h^2 S$$

and that

$$h^2 = \frac{V_a}{V_p}$$

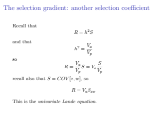
so

$$R = \frac{V_a}{V_p} S = V_a \frac{S}{V_p}$$

recall also that $S = COV[z, w]$, so

$$R = V_a \beta_{zw}$$

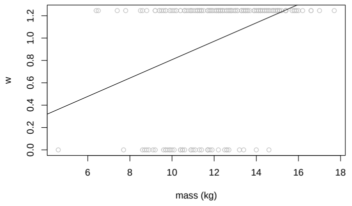
This is the *univariate Lande equation.*



The key thing to remember from the previous slide is that $\beta_{AB} = \frac{\sigma_{AB}}{\sigma^2_A}$.

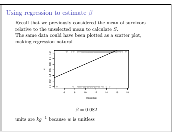
Using regression to estimate β

Recall that we previously considered the mean of survivors relative to the unselected mean to calculate S.
The same data could have been plotted as a scatter plot, making regression natural.



$$\beta = 0.082$$

units are kg^{-1} because w is unitless



forget everything you've ever been told about OLS assuming normal residuals, it isn't true.

The relation between β and S

Since

$$\beta = \frac{S}{V_z}$$

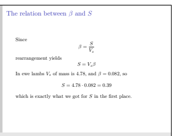
rearrangement yields

$$S = V_z\beta$$

In ewe lambs V_z of mass is 4.78, and $\beta = 0.082$, so

$$S = 4.78 \cdot 0.082 = 0.39$$

which is exactly what we got for S in the first place.



The relation between S and β may seem trivial, and the alternative coefficients redundant. Soon their different uses will become very interesting.

Standardisations of differentials and gradients 1: σ

Is selection of $S = 0.5$ kg of lamb mass stronger or weaker than (also positive directional) selection of $S = 50$ mm of oak tree sapling height? *Some kind of standardisation is required for most comparisons of selection coefficients.*

- Standardising to unit variance is by far the most common in empirical studies.
- variance-standardising S:

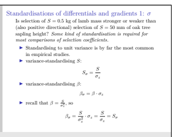
$$S_\sigma = \frac{S}{\sigma_z}$$

- variance-standardising β:

$$\beta_\sigma = \beta \cdot \sigma_z$$

- recall that $\beta = \frac{S}{\sigma_z^2}$, so

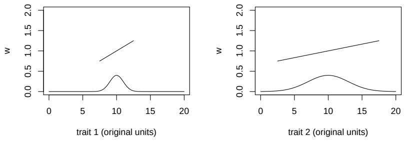
$$\beta_\sigma = \frac{S}{\sigma_z^2} \cdot \sigma_z = \frac{S}{\sigma_z} = S_\sigma$$



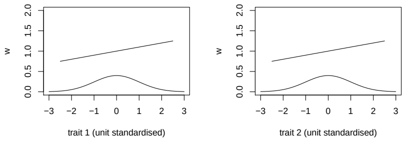
In practice, S_σ and β_σ are estimated by the simple regression of w on z; we'll come to a justification of this presently.

Properties of S_σ and β_σ

Consider these two associations between a trait and relative fitness:



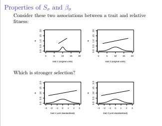
Which is stronger selection?



2020-02-03

Analysis of univariate phenotypic selection

└─Properties of S_σ and β_σ



Properties of S_σ and β_σ

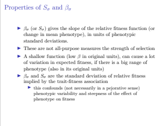
- β_σ (or S_σ) gives the slope of the relative fitness function (or change in mean phenotype), in units of phenotypic standard deviations.
- These are not all-purpose measures the strength of selection
- A shallow function (low β in original units), can cause a lot of variation in expected fitness, if there is a big range of phenotype (also in its original units)
- β_σ and S_σ are the standard deviation of relative fitness implied by the trait-fitness association
  - this confounds (not necessarily in a pejorative sense) phenotypic variability and steepness of the effect of phenotype on fitness

---

2020-02-03

Analysis of univariate phenotypic selection

└─Properties of S_σ and β_σ



Standardisations of differentials and gradients 2: μ

mean-standardising S:

$$S_\mu = \frac{S}{\mu_z}$$

Q: By what percent are survivors larger (smaller) than the initial average? A: $S_\mu(\cdot 100)$.

mean-standardising β:

$$\beta_\mu = \beta \cdot \mu_z$$

Q: By what percent does a 1% change in phenotype change relative fitness? A: β_μ.

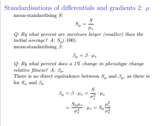
There is no direct equivalence between S_μ and β_μ, as there is for S_σ and β_σ

$$\beta_\mu = \beta \cdot \mu_z = \frac{S}{\sigma_z^2} \cdot \mu_z$$
$$= \frac{S_\mu \mu_z}{\sigma_z^2} \cdot \mu_z = S_\mu \frac{\mu_z^2}{\sigma_z^2}$$

2020-02-03

Analysis of univariate phenotypic selection

└─Standardisations of differentials and gradients 2: μ

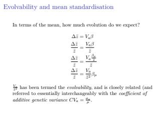


Evolvability and mean standardisation

In terms of the mean, how much evolution do we expect?

$$\Delta\bar{z} = V_a\beta$$
$$\frac{\Delta\bar{z}}{\bar{z}} = \frac{V_a\beta}{\bar{z}}$$
$$\frac{\Delta\bar{z}}{\bar{z}} = \frac{V_a\frac{\beta_\mu}{\bar{z}}}{\bar{z}}$$
$$\frac{\Delta\bar{z}}{\bar{z}} = \frac{V_a}{\bar{z}^2}\beta_\mu$$

$\frac{V_a}{\bar{z}^2}$ has been termed the *evolvability*, and is closely related (and referred to essentially interchangeably with the *coefficient of additive genetic variance* $CV_a = \frac{\sigma_a}{\mu}$.



T.F. Hansen and D. Houle. 2008. Measuring and comparing evolvability and constraint in multivariate characters. Journal of Evolutionary Biology 21: 1201-1219.

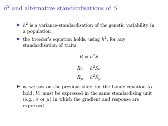
---

h^2 and alternative standardisations of S

- h^2 is a variance-standardisation of the genetic variability in a population
- the breeder's equation holds, using h^2, for any standardisation of traits

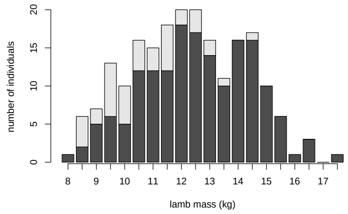
$$R = h^2S$$
$$R_\sigma = h^2S_\sigma$$
$$R_\mu = h^2S_\mu$$

- as we saw on the previous slide, for the Lande equation to hold, V_a must be expressed in the same standardising unit (e.g., σ or μ) in which the gradient and response are expressed.



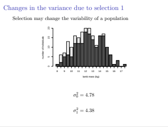
Changes in the variance due to selection 1

Selection may change the variability of a population



$$\sigma_0^2 = 4.78$$

$$\sigma_1^2 = 4.38$$



Changes in the variance due to selection 2

Care is needed: purely directional selection changes the variance too:

$$\Delta\sigma^2_z(directional) = -S^2$$

So, the change in the variance, over and above the effect of purely directional selection to reduce the variance, could be defined as

$$C = \Delta\sigma^2_z + S^2$$

In ewe lambs:

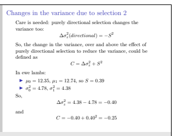
- $\mu_0 = 12.35$, $\mu_1 = 12.74$, so $S = 0.39$
- $\sigma^2_0 = 4.78$, $\sigma^2_1 = 4.38$

So,

$$\Delta\sigma^2_z = 4.38 - 4.78 = -0.40$$

and

$$C = -0.40 + 0.40^2 = -0.25$$



The *Lande-Arnold regression*

Just like the change in the mean is related to a linear regression, the change in the variance is related to a quadratic regression coefficient.
Lande and Arnold (1983) showed that

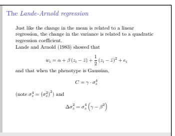
$$w_i = \alpha + \beta\left(z_i - \bar{z}\right) + \frac{1}{2}\left(z_i - \bar{z}\right)^2 + e_i$$

and that when the phenotype is Gaussian,

$$C = \gamma \cdot \sigma^4_z$$

(note $\sigma^4_z = \left(\sigma^2_z\right)^2$) and

$$\Delta\sigma^2_z = \sigma^4_z\left(\gamma - \beta^2\right)$$

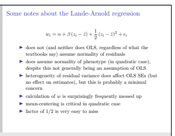


R. Lande and S.J. Arnold. 1983. The measurement of selection on correlated characters. Evolution 37: 1210–1226.

Some notes about the Lande-Arnold regression

$$w_i = \alpha + \beta\left(z_i - \bar{z}\right) + \frac{1}{2}\left(z_i - \bar{z}\right)^2 + e_i$$

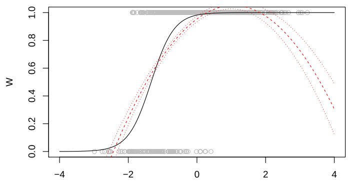
- does not (and neither does OLS, regardless of what the textbooks say) assume normality of residuals
- does assume normality of phenotype (in quadratic case), despite this not generally being an assumption of OLS
- heterogeneity of residual variance does affect OLS SEs (but no effect on estimates), but this is probably a minimal concern
- calculation of w is surprisingly frequently messed up
- mean-centering is critical in quadratic case
- factor of 1/2 is very easy to miss



for more on the factor of a half for quadratic selection gradients: Stinchcombe, J.R., Agrawal, A.F., Hohenlohe, P.A., Arnold, S.J. and Blows, M.W. 2008. Estimating nonlinear selection gradients using quadratic regression coefficients: Double or nothing? Evolution 62: 2435–2440.

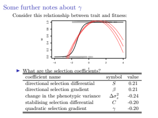
Some further notes about γ

Consider this relationship between trait and fitness:



- What are the selection coefficients?

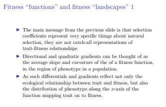
coefficient name	symbol	value
directional selection differential	S	0.21
directional selection gradient	β	0.21
change in the phenotypic variance	$\Delta\sigma_z^2$	-0.24
stabilising selection differential	C	-0.20
quadratic selection gradient	γ	-0.20



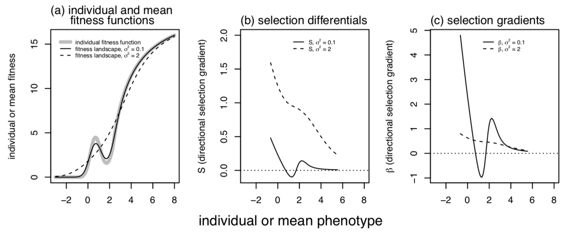
Schluter, D. 1988. Estimating the form of natural selection on a quantitative trait. Evolution 42: 849–861.

Fitness “functions” and fitness “landscapes” 1

- The main message from the previous slide is that selection coefficients represent very specific things about natural selection, they are not catch-all representations of trait-fitness relationships
- Directional and quadratic gradients can be thought of as the average slope and curvature of the of a fitness function, in the region of phenotype in a population.
- As such differentials and gradients reflect not only the ecological relationship between trait and fitness, but also the distribution of phenotype along the x-axis of the function mapping trait on to fitness.



Fitness “functions” and fitness “landscapes” 2



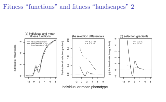


Figure is from B. Walsh and M.B. Morrissey. 2019. *Evolutionary Quantitative Genetics*, in: Handbook of Statistical Genetics, 4th ed.

A bit more formality about the average slope and curvature

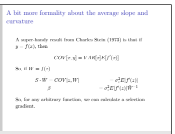
A super-handy result from Charles Stein (1973) is that if $y = f(x)$, then

$$COV[x, y] = VAR[x]E[f'(x)]$$

So, if $W = f(z)$

$$\begin{aligned} S \cdot \bar{W} = COV[z, W] &= \sigma_z^2 E[f'(z)] \\ \beta &= \sigma_z^2 E[f'(z)]\bar{W}^{-1} \end{aligned}$$

So, for any arbitrary function, we can calculate a selection gradient.



Lande and Arnold 1983's justification for OLS analysis of selection gradients is (among a great many things) an independent discovery of Stein's lemma, and an extension to the quadratic case.

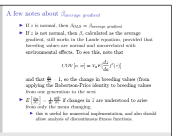
C.M. Stein. 1974. Estimation of the mean of a multivariate normal distribution. Proceedings of the Prague Symposium on Asymptotic Statistics 345–381. Univ. Karlova, Prague.

A few notes about $\beta_{average\ gradient}$

- If z is normal, then $\beta_{OLS} = \beta_{average\ gradient}$
- If z is not normal, then β, calculated as the average gradient, still works in the Lande equation, provided that breeding values are normal and uncorrelated with environmental effects. To see this, note that

  $$COV[a, w] = V_a E[\frac{dz}{da} f'(z)]$$

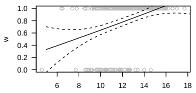
  and that $\frac{dz}{da} = 1$, so the change in breeding values (from applying the Robertson-Price identity to breeding values from one generation to the next
- $E\left[\frac{\partial w}{\partial z}\right] = \frac{1}{\bar{W}}\frac{\partial \bar{W}}{\partial \bar{z}}$ if changes in $\bar{z}$ are understood to arise from only the mean changing.
  - this is useful for numerical implementation, and also should allow analysis of discontinuous fitness functions.



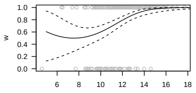
By only the mean changing, I mean that all other non-central moments stay the same. This corresponds to the thought experiment of shifting the distribution to the right, or adding a tiny bit to each individual's phenotype, just like a single generation of evolution would do.

Inference of selection gradients from arbitrary functions

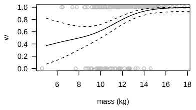
The relationship between the average partial derivatives of the fitness function and selection gradients suggests a numerical scheme applicable to any fitness function shape



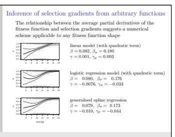
linear model (with quadratic term)
$\beta = 0.082,\ \beta_\sigma = 0.180$
$\gamma = 0.001,\ \gamma_\sigma = 0.003$



logistic regression model (with quadratic term)
$\beta = \ 0.080,\ \beta_\sigma = \ 0.176$
$\gamma = -0.0076,\ \gamma_\sigma = -0.033$



generalised spline regression
$\beta = \ 0.079,\ \beta_\sigma = \ 0.173$
$\gamma = -0.010,\ \gamma_\sigma = -0.044$



Selection coefficients and other fitness functions

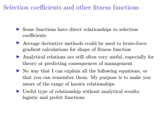
- Some functions have direct relationships to selection coefficients
- Average derivative methods could be used to brute-force gradient calculations for shape of fitness function
- Analytical relations are still often very useful, especially for theory or predicting consequences of management
- No way that I can explain all the following equations, or that you can remember them. My purpose is to make you aware of the range of known relationships
- Useful type of relationship without analytical results: logistic and probit functions

---

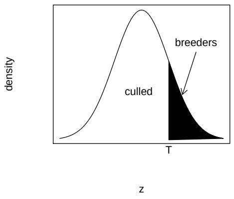
2020-02-03

Analysis of univariate phenotypic selection

└Selection coefficients and other fitness functions



Selection coefficients and other fitness functions - truncation



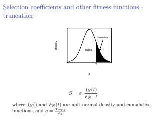
$$S = \sigma_z \frac{f_N(t)}{F_N{-}t}$$

where $f_N()$ and $F_N(t)$ are unit normal density and cumulative functions, and $g = \frac{T-\mu_z}{\sigma_z}$

2020-02-03

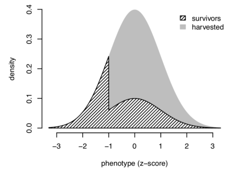
Analysis of univariate phenotypic selection

└Selection coefficients and other fitness functions - truncation



see beginning of chapter 14 in W&L2018

Selection coefficients and other fitness functions - partial truncation



- like truncation, but individuals above or below the critical trait value are culled with probability α

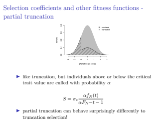
$$S = \sigma_z \frac{\alpha f_N(t)}{\alpha F_N{-}t - 1}$$

- partial truncation can behave surprisingly differently to truncation selection!

2020-02-03

Analysis of univariate phenotypic selection

└Selection coefficients and other fitness functions - partial truncation

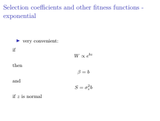


This result, and others related to functions that might arise in management of exploited populations are unpublished. Contact me if this seems useful to you!

Selection coefficients and other fitness functions - exponential

- very convenient:

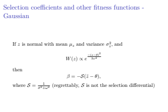
if
$$W \propto e^{bz}$$
then
$$\beta = b$$
and
$$S = \sigma_z^2 b$$
if z is normal



For discussions of how GLMs, GLMMs, and special cases that arise in mark-recapture, survival analysis, and molecular parentage analysis relate to exponential functions, see: M.B. Morrissey, and I.B.J. Goudie. 2016. Analytical results for directional and quadratic selection gradients for log-linear models of fitness functions. biorxiv.org preprint: 10.1101/040618

Selection coefficients and other fitness functions - Gaussian

If z is normal with mean μ_z and variance σ_z^2, and
$$W(z) \propto e^{\frac{-(z-\theta)^2}{2\omega^2}}$$
then
$$\beta = -\mathcal{S}(\bar{z} - \theta),$$
where $\mathcal{S} = \frac{1}{\sigma^2+\omega^2}$ (regrettably, $\mathcal{S}$ is not the selection differential)

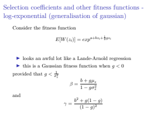


Selection coefficients and other fitness functions - log-exponential (generalisation of gaussian)

Consider the fitness function
$$E[W(z_i)] = exp^{a+bz_i+\frac{1}{2}gz_i}$$

- looks an awful lot like a Lande-Arnold regression
- this is a Gaussian fitness function when $g < 0$

provided that $g < \frac{1}{\sigma_z^2}$
$$\beta = \frac{b + g\mu_z}{1 - g\sigma_z^2}$$
and
$$\gamma = \frac{b^2 + g(1-g)}{(1-g)^2}$$



M.B. Morrissey, and I.B.J. Goudie. 2016. Analytical results for directional and quadratic selection gradients for log-linear models of fitness functions. biorxiv.org preprint: 10.1101/040618