Selection in structured population

Beyond well-mixed and homogeneous populations

We saw yesterday how invasion analyses could be used to model trait evolution in well-mixed and homogeneous populations. But most natural populations are neither (different sexes, physiologies, habitats or hosts, genetic backgrounds). For the next two days, we'll go over how invasion analyses can be performed for more complex populations that show different types of structure :

- Temporal heterogeneity (briefly)
- Within-generation heterogeneity (spending a bit more time on age-structure)
- Group structure (tomorrow)

We'll look at the expressions for directional selection $S(x)$ and disruptive selection $H(x^*)$ in terms of individual fitness. This is to facilitate cross-talk among models with individual fitness as common vocabulary and so aid with the interpretation of selection.

Between-generation structure: when environmental fluctuations affect the whole population

Temporal fluctuations are common

In nature, conditions often vary from one generation to the next.

- Resource availability may follow seasonal cycles
- Climate variables such as temperature or rainfall fluctuate year to year
- Epidemics can wax and wane
- Host or habitat availability may change unpredictably

These fluctuations affect the entire population. A rare mutant will also experience these changes and, to determine whether it can invade, we need to ask whether it is on average doing better than the resident. But what kind of average should we take?

Invasion fitness in temporally fluctuating populations

Suppose the environment switches between states $c \in \{1, \ldots, M\}$ (e.g. hot and cold, wet and dry). The probability of experiencing state c is $q_c(x)$, which may depend on the resident trait x (e.g. if the population modifies its habitat). Let $w_c(y, x)$ be the expected number of offspring of a mutant with trait y in environment c, when the resident is fixed for x.

Invasion fitness is given by the geometric mean of mutant fitness across environments

$$\rho(y, x) = \prod_{c=1}^{M} w_c(y, x)^{q_c(x)}. \tag{1}$$

Directional selection: favouring traits exposed most often

The selection gradient in a fluctuating environment is

$$S(x) = \sum_{c=1}^{M} \left. \frac{\partial}{\partial y} \log w_c(y, x) \right|_{y=x} q_c(x). \tag{2}$$

This expression shows that:

- the contribution of each environment comes from how the trait affects fitness there,
- and each contribution is weighted by how often that environment occurs.

As a result, selection tends to favour trait values that are best suited to the most common environments.

An example of adaptation to temporal heterogeneity

Consider a simple case where the environment alternates between two states each year (e.g. wet vs dry, hot vs cold).

- The species is *semelparous*: individuals reproduce once and then die.
- Fecundity in environment $i \in \{1, 2\}$:

$$\exp\left(-\frac{(z - \theta_i)^2}{2\sigma_i^2}\right) \tag{3}$$

- Thus fecundity is maximised by trait value θ_1 in environment 1 and θ_2 in environment 2, with the strength of selection set by the width parameters σ_1 and σ_2.
- The environment switches such that it is in state 1 with probability q_1 and in state 2 with probability $q_2 = 1 - q_1$.

We will now solve this example together on the whiteboard.

Disruptive selection: not so easy due to bet-hedging

The Hessian is

$$H(x^*) = \sum_{c=1}^{M} \frac{\partial^2}{\partial y^2} \log w_c(y, x^*) \bigg|_{y=x^*} q_c(x^*). \tag{4}$$

As before, the contribution of each environment is weighted by how often that environment occurs. Note that it involves the derivative of *log* fitness making it more difficult for $H(x^*) > 0$ and thus for polymorphism to emerge where different morphs are better adapted to different environments.

One way to interpret this is in terms of **bet-hedging**: because of the geometric mean form of invasion fitness, strategies that avoid very low fitness (or worse, zero!) in bad years are favoured. This tends to favour generalists who are never too bad in any particular environment.

Temporal heterogeneity cannot sustain polymorphism in semel-parous species

In the simple example we just analysed, we always find

$$H(x^*) < 0.$$

This is not an artefact of the Gaussian functions but reflects a more general result: when reproduction occurs once per generation and is synchronised with environmental fluctuations (e.g. annual reproduction in an annually fluctuating environment), temporal heterogeneity cannot generate disruptive selection and therefore cannot sustain polymorphism. One of the group projects will be to investigate conditions where temporal heterogeneity *can* lead to polymorphism.

Within-generation structure:
when individuals vary in state

Why within-generation structure matters

So far we have assumed that within a generation, populations are homogeneous; every individual is the same. But in reality individuals often differ :

- **Sex differences** — males and females contribute differently to reproduction.
- **Life stages** — juveniles survive and grow, adults reproduce.
- **Social status** — dominant and subordinate individuals have unequal access to resources.
- **Physiological states** — healthy vs. infected, well-fed vs. starved.
- **Habitats or hosts** — individuals live in different patch types or infect different host types.
- **Genetic background** — the effect of a trait may depend on other loci.

These differences matter because they change how traits affect fitness. To model them, we need to keep track of multiple classes of individuals simultaneously.

Matrix population model

Now consider a population divided into M discrete classes (e.g. males vs. females, juveniles vs. adults, dominant vs. subordinate, or individuals in different hosts).

To track the number of mutants across classes, we use a matrix recursion:

$$E[\mathbf{n}_{t+1} \mid \mathbf{n}_t] = \mathbf{W}(y, x) \cdot \mathbf{n}_t \tag{5}$$

where

- $\mathbf{n}_t = (n_{1,t}, \ldots, n_{M,t})$: vector of mutant numbers in each class at time t.
- $w_{ij}(y, x)$: expected number of class i mutants produced by one class j mutant.
- $\mathbf{W}(y, x)$: the $M \times M$ *mean matrix* of the multi-type branching process (each mutant lineage produces offspring that may be in different classes).

Invasion fitness: an eigenvalue

> **From multi-type branching process theory**
>
> $$\rho(y, x) = \text{leading eigenvalue of } \mathbf{W}(y, x). \tag{6}$$

i.e., $\rho(y, x)$ satisfies

$$\rho(y, x)\, \mathbf{q}(y, x) \;=\; \mathbf{W}(y, x) \cdot \mathbf{q}(y, x), \tag{7}$$

where $\mathbf{q}(y, x)$ is the right eigenvector of $\mathbf{W}(y, x)$, which we'll scale such that its entries sum to one:

$$\sum_{j=1}^{M} q_j(y, x) = 1. \tag{8}$$

- Asymptotically, the mutant population grows at rate $\rho(y, x)$.

- Its distribution across classes converges to $\mathbf{q}(y, x)$.

- In other words, if we sample a mutant at random, $q_j(y, x)$ is the probability that it is in class j.

Think of it this way: a new mutant arises and through reproduction its descendants spread across the classes in a stochastic way (since it's rare). While still rare, the mutant distribution across classes is going to stabilise to $\mathbf{q}(y, x)$, while the lineage grows (or declines) at rate $\rho(y, x)$.

From lineage- to individual-based fitness

It may look like we are done: we have a clear definition of invasion fitness as the leading eigenvalue of the mean matrix. But using invasion fitness directly has some drawbacks:

1. **Conceptual:** It is a *gene- or lineage-centered* measure of reproductive success. In behavioural and evolutionary ecology, however, we usually want to reason at the level of the *individual*.

2. **Computational/Practical:** The mean matrix can be large and complex. Its dominant eigenvalue is often hard to analyse mathematically, which can obscure biological insight.

Goal: rewrite invasion fitness and selection in terms of *individual fitness*, making the biology more transparent and the maths more accessible.

Reproductive value: demographic contribution of each class to long term mutant growth

A key tool for this translation is the concept of **reproductive value**. Let $\mathbf{W}^\circ = \mathbf{W}(x, x)$ be the mean matrix under neutrality. Define $\mathbf{v}^\circ$ as its left eigenvector:

$$\mathbf{v}^\circ \mathbf{W}^\circ = \mathbf{v}^\circ \qquad (\text{since } \rho(y, x) = 1). \qquad (9)$$

For practical reasons, we normalise $\mathbf{v}^\circ$ such that $\mathbf{v}^\circ \cdot \mathbf{q}^\circ = 1$, where $\mathbf{q}^\circ$ is the right eigenvector of $\mathbf{W}^\circ$. **Interpretation:** the ith entry of $\mathbf{v}^\circ$ is the *reproductive value* of class i: its relative asymptotic contribution to the long-term growth of the mutant lineage, normalised so that the average reproductive value is 1.

Reproductive value: class contribution to long-term growth

The intuition for $\mathbf{v}^\circ$ can be found by looking at the deterministic solution of the neutral process

$$\mathbf{n}_t = (\mathbf{v}^\circ \cdot \mathbf{n}_0)\rho(x,x)^t\mathbf{q}^\circ \qquad \text{as } t \to \infty. \qquad (10)$$

Imagine we introduce neutral mutants ($y = x$) in numbers $\mathbf{n}_0$ across different classes. Because the mutant is neutral, the growth rate is one ($\rho(x,x) = 1$). But some classes may contribute more strongly to the future than others e.g. mutants in high-value classes tend to leave more descendants. As generations pass, the lineage settles into the stable class distribution $\mathbf{q}^\circ$ and what matters from the start is not just how many mutants were placed in each class $\mathbf{n}_0$, but how much those classes count for the future — their *reproductive value* $\mathbf{v}^\circ$. The true founder number of the lineage is thus $\mathbf{v}^\circ \cdot \mathbf{n}_0$.

14

Directional selection

A series of algebraic steps on the leading eigenvalue yield:

$$S(x) = \sum_{i=1}^{M} \sum_{j=1}^{M} v_i^{\circ} \left. \frac{\partial w_{ij}(y, x)}{\partial y} \right|_{y=x} q_j^{\circ}. \tag{11}$$

How to read this (right to left):

- q_j°: probability a randomly sampled resident individual is in class j.
- $\dfrac{\partial w_{ij}}{\partial y}$: effect of a trait change in such a class-j individual on the expected number of offspring in class i (including survival and transitions).
- v_i°: weight of each offspring in class i according to its *reproductive value*.

Interpretation: Directional selection is the expected effect of a trait change, averaged over class frequencies and weighted by reproductive value.

Implications of the selection gradient

Eq. (11) says: selection is strongest on trait effects that **increase the production of high-value offspring from common classes**.

Example: high vs. low condition individuals

- Condition is set at birth; low condition is more common.
- High-condition individuals are more fecund $\Rightarrow$ higher reproductive value.
- Evolution therefore tends to favour traits that make *low-condition parents produce more high-condition offspring*.
- Conversely, traits that make high-condition parents produce more low-condition offspring are under weak selection.

Two sources of disruptive selection in heterogeneous populations

At a singular strategy x^*, selection can be stabilising or disruptive depending on whether $H(x^*) < 0$ or $H(x^*) > 0$ and we find that it can be decomposed into

$$H(x^*) = H_{\mathrm{w}}(x^*) + 2 \times H_{\mathrm{q}}(x^*) \tag{12}$$

where

$$
\begin{aligned}
H_{\mathrm{w}}(x^*) &= \sum_{i=1}^{M} \sum_{j=1}^{M} v_i^{\circ} \left. \frac{\partial^2 w_{ij}(y, x^*)}{\partial y^2} \right|_{y=x^*} q_j^{\circ} \\
H_{\mathrm{q}}(x^*) &= \sum_{i=1}^{M} \sum_{j=1}^{M} v_i^{\circ} \left. \frac{\partial w_{ij}(y, x^*)}{\partial y} \right|_{y=x^*} \times \left. \frac{\partial q_j(y, x^*)}{\partial y} \right|_{y=x^*}.
\end{aligned}
\tag{13}
$$

where

- $H_{\mathrm{w}}(x^*)$: disruptive selection generated by *non-linear trait effects on fitness*.
- $H_{\mathrm{q}}(x^*)$: disruptive selection generated by *feedback between traits and class*.

Disruptive selection due to non linear fitness effects

$$H_{\mathrm{w}}(x^*) = \sum_{i=1}^{M} \sum_{j=1}^{M} v_i^{\circ} \left. \frac{\partial^2 w_{ij}(y, x^*)}{\partial y^2} \right|_{y=x^*} q_j^{\circ} \tag{14}$$

This term is conceptually the same as disruptive selection in a homogeneous population:

- selection is disruptive when fitness changes *more than linearly* with the trait,
- but under class-structure these effects are weighted by how common class-j individuals are (q_j°) and the reproductive value of their offspring (v_i°).

Disruptive selection due to context-by-fitness synergy

$$H_{\mathrm{q}}(x^*) = \sum_{i=1}^{M} \sum_{j=1}^{M} v_i^{\circ} \; \left.\frac{\partial w_{ij}(y, x^*)}{\partial y}\right|_{y=x^*} \times \left.\frac{\partial q_j(y, x^*)}{\partial y}\right|_{y=x^*} \qquad (15)$$

This term is unique to class-structured populations. It captures synergy between:

1. the effect of the trait on fitness in class j,
2. and the effect of the trait on the probability of being in class j.

Intuition: Disruptive selection arises when carrying the mutant both *increases the chance of being in a certain class and increases fitness within that class*.

This can generate polymorphism with morphs specialising in different classes.

Many potential paths for context-by-fitness synergy

Disruptive selection via $H_q(x^*)$ can arise whenever a trait affects both the *context an individual finds itself in* and its *fitness within that context*.

- **Habitat choice:** oviposition preference in insects (where eggs are laid) affects both the probability of offspring developing in wet vs. dry sites, and their survival once in those sites.
- **Parasite transmission:** a surface protein or timing trait biases infection towards juvenile vs. adult hosts, while also determining parasite replication within each host type.
- **Sex determination:** a modifier gene biases the probability of developing as male or female, and also shapes mating success or fecundity within each sex.
- **Condition dependence:** a metabolic or foraging trait influences whether an individual develops in high vs. low condition, and also its fecundity given that condition.

Still complicated

The decompositions of directional and disruptive selection $S(x)$ and $H(x^*)$ offer intuitive interpretations of how selection acts in heterogeneous populations.

But in practice:

- Calculations can be cumbersome.
- Require eigenvectors and their perturbations, e.g. $\left.\dfrac{\partial q_j(y, x^*)}{\partial y}\right|_{y=x^*}$.
- Tricks exist, but they are model-specific.
- Often we must resort to numerical analysis.

For some forms of heterogeneity, however, the picture becomes much clearer. In age-structured populations in particular, the full analysis of $S(x)$ and $H(x^*)$ can be carried out explicitly. That will be our next step.

Selection in age-structured populations

Age: a major axis of variation

Age is a fundamental sources of heterogeneity, especially in endotherms. It is associated with differences in physiology, behaviour, and morphology.

To model evolution in age-structured populations (in discrete time), we can use the same matrix formalism introduced in the previous section, with classes now corresponding to age classes.

Life history can be represented as a sequence of transitions between ages via survival, combined with reproduction at each stage.

Leslie matrix

In age-structured populations, all offspring are born into the same age class, and individuals move linearly through ages. As a result, the mean matrix $\mathbf{W}(y, x)$ takes the special form of a **Leslie matrix**:

$$\mathbf{W}(y, x) = \begin{pmatrix} b_1(y, x) & b_2(y, x) & \ldots & b_{M-1}(y, x) & b_M(y, x) \\ s_1(y, x) & 0 & \ldots & 0 & 0 \\ 0 & s_2(y, x) & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & s_{M-1}(y, x) & 0 \end{pmatrix} \qquad (16)$$

where

- $b_j(y, x)$: expected number of newborn offspring produced by a mutant of age j (all offspring enter age class 1).
- $s_j(y, x)$: probability that a mutant of age j survives to age $j + 1$.

Decomposition of survival and reproduction

We can decompose the Leslie matrix as

$$\mathbf{W}(y, x) = \mathbf{S}(y, x) + \mathbf{R}(y, x) \tag{17}$$

where $\mathbf{S}(y, x)$ contains only survival terms and $\mathbf{R}(y, x)$ only reproduction:

$$\mathbf{S}(y, x) = \begin{pmatrix} 0 & 0 & \ldots & 0 \\ s_1(y, x) & 0 & \ldots & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & s_{M-1}(y, x) & 0 \end{pmatrix}, \quad \mathbf{R}(y, x) = \begin{pmatrix} b_1(y, x) & \ldots & b_M(y, x) \\ 0 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & 0 \end{pmatrix}.$$

Next generation theorem

For $\mathbf{W}(y, x) = \mathbf{S}(y, x) + \mathbf{R}(y, x)$ (with $\mathbf{S}(y, x) \geq 0$, $\mathbf{R}(y, x) \geq 0$, and $[\mathbf{I} - \mathbf{S}(y, x)]$ invertible), the leading eigenvalue of $\mathbf{W}(y, x)$ is less or equal to 1 if and only if the leading eigenvalue of

$$[\mathbf{I} - \mathbf{S}(y, x)]^{-1} \mathbf{R}(y, x)$$

is less or equal to 1, where $\mathbf{I}$ is the identity matrix.

Next generation theorem : Biological interpretation

The matrix

$$[\mathbf{I} - \mathbf{S}(y, x)]^{-1} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ s_1(y, x) & 1 & 0 & \dots & 0 \\ s_1(y, x)s_2(y, x) & s_2(y, x) & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \prod_{k=1}^{M-1} s_k(y, x) & \prod_{k=2}^{M-1} s_k(y, x) & \prod_{k=3}^{M-1} s_k(y, x) & \dots & 1 \end{pmatrix}$$

gives the expected time spent in each age (via survival): its (i, j) entry gives the expected number of time steps that an individual *starting life at age j* spends in age *i*. Meanwhile, $\mathbf{R}(y, x)$ gives the expected number of offspring produced in each age per time step.

R_0 : the lifetime reproductive number

> Multiplying $(\mathbf{I} - \mathbf{S}(y, x))^{-1}$ and $\mathbf{R}(y, x)$ shows that the leading eigenvalue is
>
> $$R_0(y, x) = \sum_{j=1}^{M} l_j(y, x)\, b_j(y, x) \qquad (18)$$
>
> where $l_j(y, x) = \prod_{k=1}^{j-1} s_k(y, x)$ is the probability a mutant survives at least to age j.

$R_0(y, x)$ is the **lifetime reproductive number**: the expected number of offspring produced over a lifetime. Therefore for having a chance to invade – $R_0(y, x) > 1$ – a mutant carrier must on average produce more that on offspring in its lifetime.

Directional selection on age-specific effects

The selection gradient can then be written as

$$\tilde{S}(x) \;=\; \left.\frac{\partial R_0(y,x)}{\partial y}\right|_{y=x} \;=\; \sum_{a=1}^{M}\left(\left.\frac{\partial b_a(y,x)}{\partial y}\right|_{y=x} + \tilde{v}_{a+1}^{\circ}\left.\frac{\partial s_a(y,x)}{\partial y}\right|_{y=x}\right) l_a^{\circ}\,,$$
(19)

where

$$\tilde{v}_j^{\circ} \;=\; \sum_{k=j}^{M}\frac{l_k^{\circ}\, b_k^{\circ}}{l_j^{\circ}}\,,$$
(20)

i.e. the **expected number of offspring an individual will produce for the rest of its lifetime, given it has already survived to age** j. This is sometimes called the *current reproductive value*. This is like the reproductive value we saw earlier, but normalised so that $\tilde{v}_1^{\circ} = 1$, so a newborn (age class 1) is the unit of measure, following Fisher's seminal work.

Directional selection is weaker on later effects

Selection on a mutant effect at age a is proportional to l_a°: the probability that a mutant survives to age a and therefore for its effect at that age to be expressed and exposed to selection. Since l_a° always declines with age, selection generally acts more strongly on early-acting than on late-acting mutations.

Age a	Survival s_a	Fecundity b_a
1	0.46	0.28
2	0.77	0.51
3	0.65	0.51
4	0.67	0.51
5	0.64	0.51
6	0.88	0.51
7	–	0.51
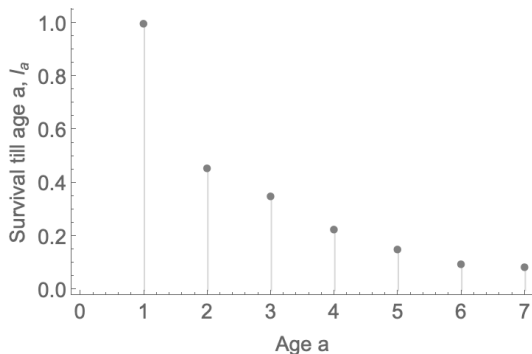
Directional selection: survival till reproduction

For mutants that affect survival from age a to $a + 1$, the strength of selection also depends on the current reproductive value $\tilde{v}_{a+1}^{\circ}$, the expected number of offspring an individual will still produce if it survives to age $a + 1$.

This means that selection on survival is strongest when individuals still have a large amount of reproductive potential ahead of them. Selection against mortality is therefore strongest near the age of first reproduction or greatest fecundity; once reproduction has begun, reproductive value typically declines so that selection against death becomes weaker.
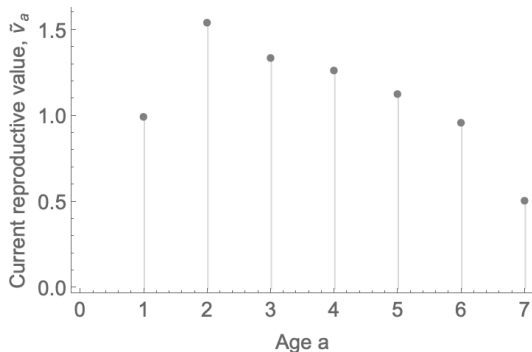
Disruptive selection in age-structured populations

Once the population expresses a singular trait value x^*, the two terms of disruptive selection

$$\tilde{H}(x^*) \;=\; \left.\frac{\partial^2 R_0(y, x)}{\partial y^2}\right|_{y=x^*} \;=\; \tilde{H}_{\mathrm{w}}(x^*) \,+\, 2\,\tilde{H}_{\mathrm{q}}(x^*)$$

are given by

$$\tilde{H}_{\mathrm{w}}(x^*) = \sum_{a=1}^{M} \left[\left.\frac{\partial^2 b_a(y, x^*)}{\partial y^2}\right|_{y=x^*} \;+\; \tilde{v}^{\circ}_{a+1} \left.\frac{\partial^2 s_a(y, x^*)}{\partial y^2}\right|_{y=x^*} \right] l^{\circ}_a$$

$$\tilde{H}_{\mathrm{q}}(x^*) = \sum_{a=1}^{M} \left[\left.\frac{\partial b_a(y, x^*)}{\partial y}\right|_{y=x^*} \;+\; \tilde{v}^{\circ}_{a+1} \left.\frac{\partial s_a(y, x^*)}{\partial y}\right|_{y=x^*} \right] \left.\frac{\partial l_a(y, x^*)}{\partial y}\right|_{y=x^*},$$

(21)

where $\tilde{H}_{\mathrm{w}}(x^*)$ is selection from *non-linear effects* of traits on fecundity or survival and $\tilde{H}_{\mathrm{q}}(x^*)$ is selection from *context-by-fitness synergy*.

The role of context-by-fitness synergy

The second term $\tilde{H}_{q}(x^*)$ captures cases where a trait affects both:

- the probability of surviving to age a $(\partial l_a(y, x^*)/\partial y)$, and
- the fecundity or survival experienced at that age (the bracketed term).

If both factors change in the same direction, their product is positive and $\tilde{H}_{q}(x^*)$ increases, which makes disruptive selection more likely.

There'll be an exercise that exemplifies this effect later this afternoon.

Selection strength is inversely proportional to generation time

Careful: $R_0(y, x)$ is sign-equivalent around 1 to invasion fitness $\rho(y, x)$ (the leading eigenvalue of $\mathbf{W}(y, x)$) but it's not equal to it! Therefore $R_0(y, x)$ is good at giving the *nature* of selection but not its overall *strength*. In fact,

$$S(x) = \frac{1}{T^\circ} \tilde{S}(x) \qquad \text{and} \qquad H(x^*) = \frac{1}{T^\circ} \tilde{H}(x^*)$$

where T° is the **generation time** in a resident population, i.e. the expected age of a parent. Selection is thus weaker in species with long generation times. This is because such species spread reproduction across many ages so the effect of a change at one age makes a smaller difference to reproduction over a whole generation. Remember: $R_0(y, x)$ is convenient for analysing the sign/structure of selection, but be careful if you wanna make statements about the absolute strength of selection (e.g. versus mutation or drift).

Summary

- Many populations are heterogeneous so that a mutant allele finds itself in individuals in different classes.
- Invasion fitness then is a weighted mean of individual fitness over the distribution of classes that a carrier of the mutant can be in.
- Under between-generation heterogeneity, the mean is geometric, favouring bet-hedging. Under within-generation heterogeneity, the mean is arithmetic and the weights are the reproductive values.
- Selection tends to favour fitness effects in classes that are most common to produce offspring with highest reproductive value.
- This means prioritizing reproduction at early age and survival till fecund ages
- Within-generation heterogeneity favours polymorphism and class specialisation when trait expression simultaneously increases the probability that an individual is in a class and fitness in that class.