

Modeling the evolution of differences in variability between sexes

Theodore P. Hill
School of Mathematics
Georgia Institute of Technology
Atlanta
GA 30332
U.S.A.

Abstract

An elementary mathematical theory based on a “selectivity-variability” principle is proposed to address a question raised by Charles Darwin, namely, how one sex of a sexually dimorphic species might tend to evolve with greater variability than the other sex. Two mathematical models of the principle are presented: a discrete-time one-step probabilistic model of the short-term behavior of the subpopulations of a given sex, with an example using normally distributed perceived fitness values; and a continuous-time deterministic coupled ODE model for the long-term asymptotic behavior of the expected sizes of the subpopulations, with an example using exponentially distributed fitness levels.

Subject Classification: (2010) 92D25, 92D15, 91A22.

Keywords: *Variability hypothesis, Survival function, Evolutionary game theory, Selectivity-variability principle, and Population dynamics.*

1. Introduction

In his research on evolution in the 19th century Charles Darwin observed that throughout the animal kingdom males are generally more variable than females of the same species, and he raised the question of *why* this might occur:

“The cause of the greater general variability in the male sex, than in the female is unknown” [7, p. 224].

This question has persisted into the 21st century (e.g., see [12, p. 1-2]). For example, as statistician Howard Wainer phrased it “Why was

our genetic structure built to yield greater variation among males than females? And not just among humans, but virtually all mammals” [23, p. 255].

The past quarter century has produced much new research on this topic in different contexts, most of which refer to humans, and the great majority of which support Darwin’s observation of greater male variability (e.g., see Appendix A of [12]). The objective of this paper is to propose an elementary mathematical principle to help explain how one sex in a given species could naturally evolve toward greater or lesser variability depending on the preferences of the opposite sex. Together with two additional standard biological tenets, this principle might help provide an answer to Darwin’s question.

2. A Theory for Differences in Variability Between Sexes

In very general terms, the main principle of the theory introduced here is this:

A Selectivity-Variability Principle. *In a species with two sexes A and B, both of which are needed for reproduction, suppose that sex A is relatively **selective**, i.e., will mate only with a top tier (less than half) of B candidates. Then from one generation to the next, among subpopulations of B with comparable average attributes, those with **greater variability** will tend to prevail over those with lesser variability. Conversely, if A is relatively **non-selective**, accepting all but a bottom fraction (more than half) of the opposite sex, then subpopulations of B with **lesser variability** will tend to prevail over those with comparable means and greater variability.*

Note that this principle does not make any assumptions about inherent differences in means or other attributes between the sexes. For instance, it does not require that one sex is selective and the other non-selective, or even that one sex is more selective than the other, unlike Bateman’s principle [2], for example, or other related theories such as “the sex that experiences more intense...vetting by the other sex will tend to show greater within-sex variation on many traits” [10, p. 176]. If both sexes of a species happen to be selective, the selectivity-variability principle here predicts that the best evolutionary strategy for each is to tend toward greater variability.

It is also important to note that this principle alone says nothing about *a priori* or *a posteriori* comparisons of the variabilities across the sexes. For example, if sex A is not selective while sex B is selective, this principle

says that subpopulations of A with greater variability will prevail over subpopulations of A with lesser variability, and that subpopulations of B with lesser variability will prevail over subpopulations of sex B with greater variability. It says nothing about comparing the resulting variability of sex A with the variability of sex B. If all the subpopulations of B were initially more variable than all the subpopulations of A, for instance, then the next generation of sex B will still exhibit greater variability than the next

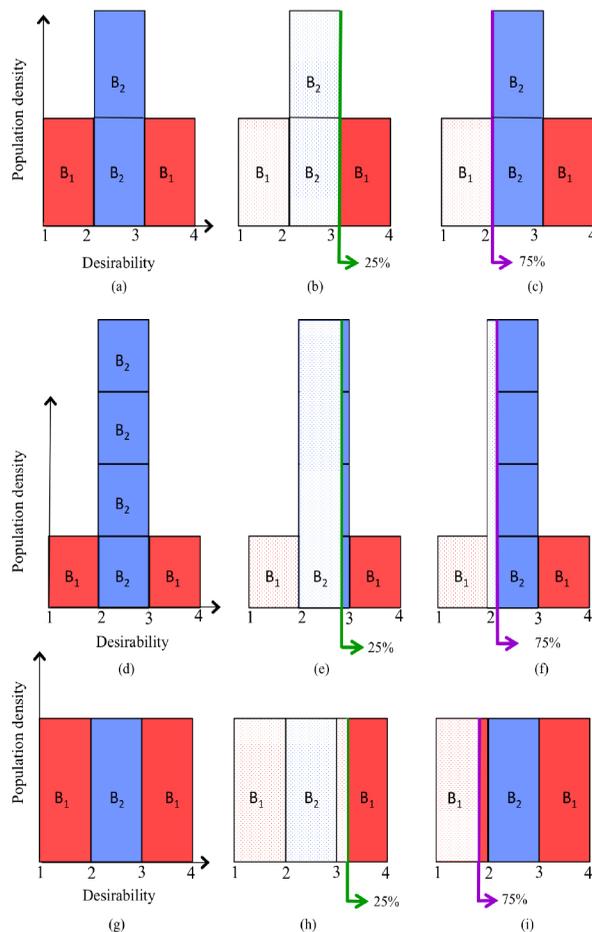


Figure 1

The three cases of Example 1: The red blocks represent the distribution of the more variable subpopulation B_1 of sex B, and the blue represent the less variable subpopulation B_2 .

generation of sex A whether either sex is selective or non-selective. Only under additional hypotheses, such as “both sexes began with comparable mid-range variability”, as will be done in an illustrative example below to address Darwin’s question, can this theory be useful to draw any conclusions about comparisons of variability between the two sexes.

In order to make this selectivity-variability theory more precise, of course, it is necessary to define formally what is meant by selectivity and variability in this context, and that will be done in the next section. First, the following simple informal hypothetical example may help convey the intuition behind this principle.

Example 1 : Sex B consists of two subpopulations B_1 and B_2 . Sex A considers half of the individuals in B_1 very desirable and the other half not very desirable, and it considers all of the individuals in subpopulation B_2 of mid-range desirability. Thus B_1 is more variable in desirability to sex A than B_2 , with B_1 and B_2 having comparable average desirability. (See Figure 1, where larger numbers reflect greater desirability.)

Special Case 1 : B_1 and B_2 are of equal size. Then sex A considers one quarter of sex B (the lower half of B_1) of relatively low desirability, half of B (all of B_2) of medium desirability, and one quarter of B (the upper half of B_1) of above-average desirability (see Figure 1a). If sex A is relatively selective and will mate only with the top most desirable quarter of sex B, then all of the next generation will be offspring of the more variable subpopulation B_1 (Figure 1b). On the other hand, if sex A is relatively non-selective and will mate with any but the lower quarter of B, then all of the less variable B_2 will mate, but only half of the more variable B_1 will mate (Figure 1c).

Similar conclusions follow if the initial subpopulations are not of equal size.

Special Case 2 : One third of sex B is the more variable B_1 and two thirds is the less variable B_2 (Figure 1d). If sex A only mates with the most desirable quarter of B, a short calculation shows that two thirds of the next generation will be offspring of B_1 and one third will be offspring of B_2 , so based on the initial distribution, the more variable subpopulation will be overrepresented (Figure 1e). If sex A will mate with any but the least desirable quarter of B, then only two ninths of the next generation will be offspring of B_1 and seven ninths will be offspring of B_2 , so the less variable subpopulation of sex B will be overrepresented (Figure 1f).

Special Case 3 : Two thirds of sex B is the more variable B_1 and only one third is the less variable B_2 (Figure 1g). If sex A only mates with the most

desirable quarter of B , then all of the next generation will be offspring of B_1 (Figure 1h), and if sex A will mate with any but the least desirable quarter of B , then only five ninths of the next generation will be offspring of B_1 and the rest will be offspring of B_2 , so again the less variable subpopulation of sex B will be overrepresented (Figure 1i).

Note the asymmetry in the mating probabilities in this example. Some intuition behind why this occurs may perhaps be gained from the observation that the most desirable individuals in the more variable population will always be able to mate, whether the opposite sex is selective or non-selective.

3. Desirability and Selectivity

In order to begin to try to interpret these ideas analytically, it is of course necessary to identify concrete definitions of “desirable”, “selective” and “more variable”. There are clearly many different candidates to capture the essence of each of these terms; the following assumptions and definitions are simply intended as a starting point to facilitate proposal of several models and analysis of the above selectivity-variability principle.

To begin with, the informal notion of desirability introduced in Example 1 above will be extended as follows.

Desirability Assumption : *Each individual (or phenotype) in each sex is assigned a real number which reflects its desirability to the opposite sex, with higher values indicating greater desirability.*

As a concrete example, one interpretation of the desirability value of an individual might be the opposite sex’s perception or estimation of its Darwinian fitness (e.g., [5]). The actual magnitudes of these desirability values are not assumed to have intrinsic significance in general, but are used only to make comparisons between individuals. Here and throughout, it will also be assumed that the same desirability value is assigned to each individual by every member of the opposite sex. In real life scenarios, of course, the desirability of an individual varies from one member of the opposite sex to another, and is not quantifiable in a single one-parameter value.

Next, the informal notion of selectivity introduced in Example 1 above will be formalized.

Selectivity Assumption : For each sex in a given sexually dimorphic species there is an upper proportion $p \in (0, 1)$ of the opposite sex that is acceptable for mating.

For example, if p_A is that proportion for sex A , then members of sex A will mate with individual b of the opposite sex B if and only if b is in the most desirable p_A fraction of individuals in B . If $p_A < 0.5$, then sex A is said to be (relatively) *selective*, and if $p_A > 0.5$, then A is said to be *non-selective*. For instance, if $p_A = 0.25$, then sex A is selective, since it will mate only with the most desirable quarter of sex B , and if $p_A = 0.75$, then sex A is non-selective, since it will mate with any but the least desirable quarter of sex B .

N.B. Of course these assumptions about desirability values and selectivity are clearly not satisfied in most real life scenarios, and are simply intended here as a starting point for discussion of the general ideas. For example, the acceptability fractions p_A may reflect not only desirability, but also availability or proximity. In this simple model it is therefore assumed that the populations are large and mobile so there are always available potential mates of the opposite sex above the threshold desirability cutoff. Similarly, for simplicity it will be assumed throughout that the offspring of any coupling consist of equal numbers of each sex.

4. Variability

The desirability of individuals in one sex by the opposite sex varies from individual to individual, and its normalized distribution is a probability distribution. Thus to address the notion of differences between two subpopulations of the same sex in the variability of their desirability to the opposite sex, the notion of one probability distribution being more (or less) variable than another must be specified. As will be seen in the next example, for instance, if by “more variable” is meant “larger standard deviation” (or statistical variance), then the above selectivity-variability principle is not true without additional assumptions on the underlying distributions.

Example 2 : Sex B consists of two subpopulations B_1 and B_2 , with six individuals each: B_1 has one individual of desirability value 1 (to sex A), one of desirability 5, and four individuals of desirability 3; B_2 has three individuals of desirability value 2 and three of desirability 4. Thus both B_1

and B_2 have mean desirability 3, the variance of B_1 is $4/3$ and the variance of B_2 is 1.

If sex A is *selective* with $p_A = 0.25$, then two of the three individuals that sex A selects from sex B will be from B_2 , the subpopulation with *smaller variance*. Conversely, if sex A is *non-selective* with $p_A = 0.75$, then five of the nine individuals that sex A selects from sex B will be from B_1 , the subpopulation with *larger variance*. Thus for these distributions and a standard deviation definition of variability, *both directions of the above selectivity-variability principle fail*.

There are many other possibilities for definitions of variability, such as comparisons of ranges or Gini mean differences, but those can be very misleading in this setting since a single outlier can dramatically alter the values of such statistics. On the other hand, basic comparisons of the tails of distributions leads to a natural notion of greater or lesser variability.

To that end, for a real Borel probability measure P let S_p denote the complementary cumulative distribution function of P , i.e., $S_p : \mathbb{R} \rightarrow [0, 1]$ is defined by $S_p(x) = P(x, \infty)$ for all $x \in \mathbb{R}$. That is, $S_p(x)$ is simply the proportion of a population with distribution P that is above the threshold x ; see Figure 2 for three examples. For brevity, the term *survival function* will be used here; in this context $S_p(x)$ may be thought of as the proportion of a given sex with desirability (by the opposite sex) distribution function P that “survives” the cut when the opposite sex only accepts individuals with desirability value x or larger.

Definition 3 : For two probability measures P_1 and P_2 on the real line \mathbb{R} with identical medians m , P_1 is more variable than P_2 , written $P_1 > P_2$, if for all x with $0 < S_{P_1}(x) < 1$,

$$S_{P_1}(x) > S_{P_2}(x) \text{ for all } x > m \text{ and } S_{P_1}(x) < S_{P_2}(x) \text{ for all } x < m.$$

In other words, P_1 is more variable than P_2 if the proportions of P_1 both above every upper (larger than median) threshold and the proportions below every lower threshold level are greater than those for P_2 . That is, both upper and lower tails of the P_1 distribution are heavier than those of the P_2 distribution, for all thresholds.

In Example 1 above, where the selectivity-variability principle was illustrated informally, the distribution of subpopulation B_1 is more variable than the distribution of subpopulation B_2 both in the sense of standard deviation and in the sense of Definition 3 (see Figure 2a), and it is this latter definition that will be seen below to lead to settings where

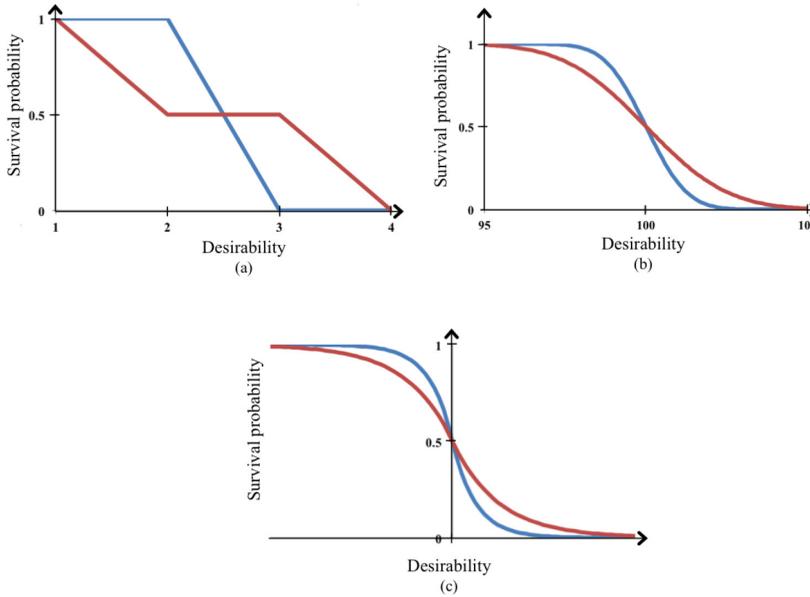


Figure 2

The survival functions and comparative variability of three pairs of distributions: (a) the uniform distributions in Example 1 above; (b) the normal distributions in Example 7 below; and (c) the Laplace distributions in Example 9. In each case, the red curves denote the more variable distribution.

the principle is valid. As was seen in Example 2, the selectivity-variability principle may fail for arbitrary distributions if variability is defined in terms of standard deviation, but the next proposition identifies several common and important classes of distributions where greater standard deviation or scale factor (when standard deviation is infinite) coincide with the notion of greater variability in Definition 3, and thus these distributions are applicable to the models below. The conclusions are perhaps well-known, but as no reference is known to the author, a short proof is included.

Proposition 4 : Let P_1 and P_2 be (real Borel) probability measures with identical medians.

- (i) If both P_1 and P_2 are uniform, symmetric triangular, Laplace, or Gaussian, then

$$P_1 \succ P_2 \quad \text{if and only if} \quad \text{variance}(P_1) > \text{variance}(P_2).$$

(ii) If both P_1 and P_2 are Cauchy, then

$$P_1 \succ P_2 \quad \text{if and only if} \quad \text{scale factor of } P_1 > \text{scale factor of } P_2.$$

Proof: Since the cumulative distribution functions for uniform, symmetric triangular, Laplace, and Cauchy distributions are known in closed form, the conclusions regarding those distributions follow from Definition 3 and routine calculations by comparisons of the respective piecewise linear, quadratic, exponential, and arcsin distribution functions.

To see the conclusion for Gaussian distributions, for which the distribution functions are not known in closed form, suppose $X_1 \sim N(\mu, \sigma_1^2)$ and $X_2 \sim N(\mu, \sigma_2^2)$, where $N(\mu, \sigma^2)$ denotes a normal distribution with mean μ and standard deviation σ . Without loss of generality suppose that $\sigma_1^2 > \sigma_2^2$. Then for all $c > \mu$,

$$\begin{aligned} P(X_1 > c) &= P(\sigma_2(X_1 - \mu) > \sigma_2(c - \mu)) \\ &= P(\sigma_1(X_2 - \mu) > \sigma_2(c - \mu)) \\ &> P(\sigma_1(X_2 - \mu) > \sigma_1(c - \mu)) \\ &= P(X_2 > c), \end{aligned}$$

where the second equality follows since, by the rescaling and translation properties of normal distributions,

$$\sigma_2(X_1 - \mu) \text{ and } \sigma_1(X_2 - \mu) \text{ are both } N(0, \sigma_1^2 \sigma_2^2).$$

The case $c < \mu$ follows similarly, and since the mean of every normal distribution is the same as the median, this completes the proof. \square

It should also be noted that distributions sufficiently close to the distributions in Proposition 4 will also obey the same variability conclusions. (E.g., no real-life data is ever exactly Gaussian, but in many applications Gaussian distributions are good approximations and very useful in practice.) Note that the above definition of greater variability does not require finite standard deviations or symmetry of the distributions, although the examples provided below have both. Some assumption on two distributions (of the same sex) having comparable average attributes is clearly necessary to be able to draw any useful conclusions in this selectivity context; the assumption of identical medians used here is one natural candidate. Similar conclusions may be drawn about weak-inequality versions of this definition and about one-sided variability, and these are left to the interested reader. For example, if both the median and

upper tails of one distribution are larger than those of another, then that distribution will also prevail if the opposite sex is selective.

Using the above definitions of variability and selectivity, the main objective of this paper is to present two mathematical models for the selectivity-variability principle above.

5. A Discrete-Time Probabilistic Model

Suppose that sex B of a given hypothetical species consists of two distinct subpopulations B_1 and B_2 , of which a proportion $\beta \in (0, 1)$ is of type B_1 (and $1 - \beta$ is of type B_2). Let P_1 and P_2 denote the desirability distributions of B_1 and B_2 , respectively, and assume that P_1 is more variable than P_2 , i.e., $P_1 \succ P_2$. It will now be shown that, for all β , if sex A is selective, then subpopulation B_1 will be overrepresented in the subsequent generation, and if sex A is non-selective, then subpopulation B_2 will be overrepresented in the subsequent generation. These are direct analogs and extensions of the informal observations in Example 1 above.

To that end, note that if β is the proportion of sex B that is from subpopulation B_1 , then letting S_1 and S_2 denote the survival functions of B_1 and B_2 , respectively, the number

$$\frac{\beta S_1(c)}{\beta S_1(c) + (1 - \beta) S_2(c)}$$

represents the proportion of sex B that is from subpopulation B_1 when A accepts only individuals in B with desirability value above cutoff level c . This motivates the following definition.

Definition 5 : If sex B consists of two subpopulations B_1 and B_2 and if $\beta \in (0, 1)$ is the proportion of sex B that is B_1 , then subpopulation B_1 will be overrepresented in the subsequent generation if and only if

$$\frac{\beta S_1(c^*)}{\beta S_1(c^*) + (1 - \beta) S_2(c^*)} > \beta,$$

where S_1 and S_2 are the survival functions of the desirability distributions of B_1 and B_2 , respectively, and c^* is the desirability cutoff of sex A for mating with individuals in sex B , i.e.,

$$\beta S_1(c^*) + (1 - \beta) S_2(c^*) = p_A.$$

Note that this definition does not assume that the offspring of B_1 will have desirability distributions identical to that of B_1 but simply that

a larger proportion of the subsequent generation will be offspring of B_1 than the proportion of B_1 in the original population. With this notion of overrepresentation, one elementary formalized version of the selectivity-variability principle for short-term behavior is as follows.

Theorem 6 : *Let sex B consist of two distinct subpopulations B_1 and B_2 with desirability distributions P_1 and P_2 , respectively, with identical medians m , and with desirability survival functions S_1 and S_2 which are continuous and strictly decreasing. Suppose subpopulation B_1 is more variable than B_2 , i.e., $P_1 \succ P_2$. Then*

- (i) *If $p_A < 0.5$, i.e., if sex A is selective, then the more variable subpopulation B_1 will be overrepresented in the subsequent generation.*
- (ii) *If $p_A > 0.5$, i.e., if sex A is non-selective, then the less variable subpopulation B_2 will be overrepresented in the subsequent generation.*

Proof : Let $\beta \in (0, 1)$ be the proportion of B that is B_1 , and let S_1 and S_2 denote the desirability survival functions for B_1 and B_2 , respectively. First, it will be shown that there exists a unique “threshold” desirability cutoff $c^* \in \mathbb{R}$ such that

$$\begin{aligned} \beta S_1(c^*) + (1 - \beta) S_2(c^*) &= p_A \\ \text{and} & \\ c^* > m \text{ if } p_A < 0.5 \text{ and } c^* < m \text{ if } p_A > 0.5. & \end{aligned} \quad (1)$$

To see (1), let $g : \mathbb{R} \rightarrow (0, 1)$ be given by $g(c) = \beta S_1(c) + (1 - \beta) S_2(c)$. Then g is continuous and strictly decreasing with $g(-\infty) = 1$, $g(m) = 0.5$, $g(\infty) = 0$, so c^* satisfying (1) exists and is unique, and since $S_1(m) = S_2(m) = 0.5$, $c^* > m$ if $p_A < 0.5$ and $c^* < m$ if $p_A > 0.5$.

To see (i), first note by (1) that $c^* > m$, so since $P_1 \succ P_2$,

$$S_1(c^*) > S_2(c^*).$$

Thus,

$$\beta(1 - \beta)S_1(c^*) > \beta(1 - \beta)S_2(c^*),$$

which implies

$$\beta^2 S_1(c^*) + \beta(1 - \beta)S_1(c^*) > \beta^2 S_1(c^*) + \beta(1 - \beta)S_2(c^*)$$

so

$$\frac{\beta S_1(c^*)}{\beta S_1(c^*) + (1 - \beta)S_2(c^*)} > \beta.$$

By Definition 5, this completes the proof of (i); the proof of (ii) follows similarly. \square

Thus, in this discrete-time setting, if one sex remains non-selective from each generation to the next, for example, then in each successive generation less variable subpopulations of the opposite sex will tend to prevail over more variable subpopulations of comparable average desirability. Although those successive generations and their desirability distributions are evolving over time, if less variable subpopulations in the opposite sex prevail over more variable subpopulations from each generation to the next, that suggests that over time the opposite sex will tend toward lesser variability. A key assumption here, of course, is the inheritability of variability itself.

That variability *per se* may be a heritable trait has recently been established in several different contexts (e.g., [8], [16], and [17]). For instance, theories of inherited variability have been developed and applied by animal husbandry scientists and geneticists who are interested in breeding livestock, not only for high averages of desirable traits, but also for uniformity (i.e., low variability; see [15], [18], [19]). Application of the above probabilistic model to the evolution of differences in variability only uses the premise that variability is inheritable; identification of the precise genetic, chromosomal, and epigenetic (including societal) mechanisms for how variability is inherited is beyond the scope of this paper, and the interested reader is referred, e.g., to [11], [14], and [20].

The next example is an application of Theorem 6 and Proposition 4 in a case where the desirability values have gaussian distributions. Note that the exact magnitudes of over- and underrepresentation, even in this standard setting, are not attainable in closed form, but must be estimated numerically.

Example 7 : Suppose that the desirability values (to sex A) of sex B are normally distributed, i.e., if X_1 and X_2 are the desirability values of two random individuals chosen from B_1 and B_2 , respectively, then X_1 has distribution $N(\mu, \sigma_1^2)$ and B_2 has distribution $N(\mu, \sigma_2^2)$. (The assumption of normality for the underlying distributions of desirability is not essential; this is merely an illustrative example, and chosen because of the ubiquity of the normal distribution in many population studies. Note the key assumption that the average values, i.e. the medians, are the same.) By Proposition 4, $N(\mu, \sigma_1^2)$ is more variable than $N(\mu, \sigma_2^2)$ if and only if $\sigma_1^2 > \sigma_2^2$.

In particular, suppose $X_1 \sim N(100, 4)$, $X_2 \sim N(100, 1)$, so B_1 is more variable than B_2 . Suppose that B_1 and B_2 are of equal size, and again consider the two illustrative cases above, namely, where sex A is selective with $p_A = 0.25$, and where sex A is non-selective with $p_A = 0.75$ (see Figure 3).

Suppose first that $p_A = 0.25$. Using a special function calculator (since the survival functions of normal distributions are not known in closed form), it can be determined numerically that sex A 's threshold desirability value cutoff for sex B is $c^* \cong 100.92$, $S_1(c^*) \cong 0.323$, and $S_2(c^*) \cong 0.179$. Thus a random individual from subpopulation B_1 has nearly twice the probability of mating than one from the less variable subpopulation B_2 , as is illustrated in Figure 3 with the areas to the right of the green desirability cutoff. Hence B_1 will be overrepresented in the subsequent generation.

Next suppose that $p_A = 0.75$. Then it can be determined that the threshold desirability value cutoff is $c^* \cong 99.08$, $S_1(c^*) \cong 0.677$, and $S_2(c^*) \cong 0.821$, i.e., a random individual from subpopulation B_2 is about one-fifth more likely to be able to mate than one from the more variable subpopulation B_1 . This is illustrated in Figure 3 with the areas to the right of the purple cutoff. Here again, note the asymmetry in that the selective case is more extreme than the non-selective case, as was seen in Example 1.

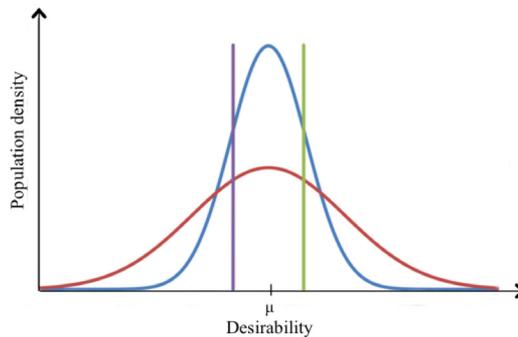


Figure 3

The red curve is the desirability distribution of the more variable normal subpopulation B_1 in Example 7 and the blue curve is the desirability distribution of the less variable subpopulation B_2 . The vertical green line is the threshold cutoff for the opposite sex A so that exactly 25% of the composite B population has desirability value above (to the right of) that point. The vertical purple line is the value so that exactly 75% of the B_2 population has desirability value above that point.

6. A Continuous-Time Deterministic Model

In this section, comparisons of the long-term asymptotic behavior of the sizes of competing subpopulations of the same sex and species are modeled using the general structure and logic of classical evolutionary game theory as applied to population dynamics (cf. [4], [24], [25]).

Here, sex B consists of two distinct subpopulations B_1 and B_2 , growing in time, whose sizes at time t are given by the continuous stochastic processes $X_1(t)$ and $X_2(t)$, respectively. Letting $x_1(t) = E[X_1(t)]$ and $x_2(t) = E[X_2(t)]$ denote the expected values of the subpopulation sizes at time t , the objective here will be to derive a coupled system of ODE's, directly analogous to the coupled systems of ODE's in classical evolutionary game theory, to model the growth rates of the expected values of the sizes of the two subpopulations (cf. [21]).

In contrast to the discrete-time model above, here there is no clear delineation between generations, and it will be assumed that the pace of evolution is negligible compared to the pace of reproduction, so the two subpopulations remain distinct, with offspring distributed the same way as the parent subpopulation. In this setting, it will now be seen that if one subpopulation is more variable than the other, then the more variable subpopulation will eventually eclipse the less variable subpopulation if the opposite sex is selective, and the less variable subpopulation will eclipse the more variable one if the opposite sex is non-selective.

Assume that the desirability distributions of B_1 and B_2 (to sex A) are given by probabilities P_1 and P_2 , respectively, that do not change with the sizes of the subpopulations, i.e., the survival and desirability distribution functions do not change with t . For further ease of analysis, assume that the expected population sizes $x_1(t)$ and $x_2(t)$ are strictly increasing and differentiable and that the survival functions S_1 and S_2 for P_1 and P_2 are both continuous and strictly decreasing, with identical (unique) medians m . In particular, exactly half of each subpopulation B_1 and B_2 has desirability value above m to sex A at all times $t > 0$, and exactly half of each has desirability values below m .

In this deterministic framework, the composite population of sex B is growing at an instantaneous rate that is proportional to the fraction p_A of its members that is acceptable to the opposite sex A . That is, with the constant of proportionality taken to be 1,

$$\frac{d(x_1 + x_2)}{dt} = p_A(x_1 + x_2). \quad (2)$$

Similarly, the expected values of the sizes of both subpopulations B_1 and B_2 are growing at rates proportional to the fractions $S_i(c^*)$ of each subpopulation that are acceptable to sex A at that time, i.e., which satisfy the coupled system of ordinary differential equations

$$\frac{dx_i}{dt} = x_i S_i(c^*), \quad i = 1, 2, \quad (3)$$

where $c^* = c^*(x_1, x_2)$ is the cutoff value so that the expected proportion of sex B that is above desirability level c^* at time t is exactly p_A . Since P_1 and P_2 , and hence S_1 and S_2 are assumed to be constant in time, c^* satisfies

$$\frac{x_1 S_1(c^*) + x_2 S_2(c^*)}{x_1 + x_2} = p_A. \quad (4)$$

The coupled system of ODE's (3) is very closely related to the classical replicator equation (cf.[3], [6]), which also captures the essence of selection via acceptability for mating but through rates proportional to deviation from the mean desirability (or fitness), rather than through rates proportional to fractions above selectivity cutoffs. Analogous to the discrete probabilistic model above, solutions of (3) are not generally available in closed form, and must be approximated numerically, as will be seen in Example 9 below.

Qualitative comparisons of the rates of growth of competing subpopulations satisfying (4) are possible, however, and the next theorem shows that the selectivity-variability principle above is also valid in this setting, in the following sense. If P_1 is more variable than P_2 , and if $p_A < 0.5$, i.e., if sex A is selective, then the relative instantaneous rate of growth of B_1 exceeds that of B_2 , and the proportion of sex B that is from B_1 approaches 1 in the limit as time goes to infinity. Conversely, if $p_A > 0.5$, i.e., if sex A is non-selective, then the relative instantaneous rate of growth of B_2 exceeds that of B_1 , and the less variable subpopulation B_2 prevails in the limit. This same conclusion can be extended to more general settings, such as time-dependent acceptability fractions $p_A(t)$, and these generalizations are left to the interested reader.

Recall that P_1 and P_2 , respectively, are the desirability distributions (to sex A) of subpopulations B_1 and B_2 of sex B .

Theorem 8 : *Suppose subpopulation B_1 is more variable than B_2 , i.e., $P_1 \succ P_2$.*

- (i) *If $p_A < 0.5$, i.e., if sex A is selective, then the relative rate of growth of B_1 exceeds that of B_2 ,*

$$\frac{1}{x_1} \frac{dx_1}{dt} > \frac{1}{x_2} \frac{dx_2}{dt}. \quad (5)$$

Moreover, $\frac{x_1}{x_1+x_2} \rightarrow 1$ as $t \rightarrow \infty$.

(ii) If $p_A > 0.5$, i.e., if sex A is non-selective, then the relative rate of growth of B_2 exceeds that of B_1 ,

$$\frac{1}{x_2} \frac{dx_2}{dt} > \frac{1}{x_1} \frac{dx_1}{dt}. \quad (6)$$

Moreover, $\frac{x_1}{x_1+x_2} \rightarrow 0$ as $t \rightarrow \infty$.

Proof of (i) : Analogous to the argument for Theorem 6, define $g : \mathbb{R} \rightarrow (0, 1)$ by

$$g(c) = \frac{x_1 S_1(c) + x_2 S_2(c)}{x_1 + x_2},$$

where S_1 and S_2 are the desirability survival functions for P_1 and P_2 , respectively. Recall that S_1 and S_2 are both continuous and strictly decreasing with identical medians $m > 0$, and fix $t > 0$. Since g is continuous and strictly decreasing with $g(-\infty) = 1$, $g(m) = 0.5$, and $g(\infty) = 0$, there exists a unique threshold desirability cutoff $c^* = c^*(t)$ satisfying (4), where, as before, p_A is the most desirable fraction of sex B that is acceptable to sex A , and $c^* = c^*(t)$ is the threshold desirability cutoff for sex A for the combined populations of sex $B = B_1 \cup B_2$ at time t .

Note that $S_1(m) = S_2(m) = 0.5$, so since $p_A < 0.5$, $c^* > m$. Since P_1 is more variable than B_2 this implies that $S_1(c^*) > S_2(c^*)$. Since $S_1(c^*)$ and $S_2(c^*)$ are the proportions of B_1 and B_2 , respectively, that are above the threshold cutoff at time $t > 0$, (3) implies (5).

To see that $\frac{x_1}{x_1+x_2} \rightarrow 1$ as $t \rightarrow \infty$, note that since P_1 is more variable than P_2 , $m < S_2^{-1}(p_A) < S_1^{-1}(p_A)$ for $p_A < 0.5$. Clearly $c^* \in [S_2^{-1}(p_A), S_1^{-1}(p_A)]$ for all $t > 0$, so since $S_2(x) < S_1(x)$ for all $x > m$, the continuity of S_1 and S_2 implies the existence of $\delta > 0$ so that

$$S_1(c^*) > S_2(c^*) + \delta \text{ for all } c^* \in [S_2^{-1}(p_A), S_1^{-1}(p_A)] \text{ and for all } t > 0.$$

Thus by (3),

$$\frac{1}{x_1} \frac{dx_1}{dt} > \frac{1}{x_2} \frac{dx_2}{dt} + \delta \text{ for all } t > 0,$$

so $\ln x_1 - \ln x_2 \geq \delta t + \alpha$, which implies that $\frac{x_1}{x_1 + x_2} \rightarrow 1$ as $t \rightarrow \infty$, completing the proof of (i).

The proof of (ii) is analogous. \square

The next example is a numerical application of Theorem 8 and Proposition 4 where the desirability values decrease exponentially from the common median. Note that, as was the case in Example 7 above, the solution, in this case of the underlying coupled system of ODE's, is not available in closed form, but must be approximated numerically.

Example 9: Let the survival functions S_1 and S_2 for subpopulations B_1 and B_2 be Laplace distributions with $S_1(x) = e^{-x}/2$ for $x \geq 0$ and $S_2(x) = e^{-2x}/2$ for $x \geq 0$ (see Figure 4). (For ease of exposition, the common median desirability of both subpopulations here is taken to be 0; translation to a different median value is trivial.) By Proposition 4, $P_1 \succ P_2$, so subpopulation B_1 is more variable than B_2 .

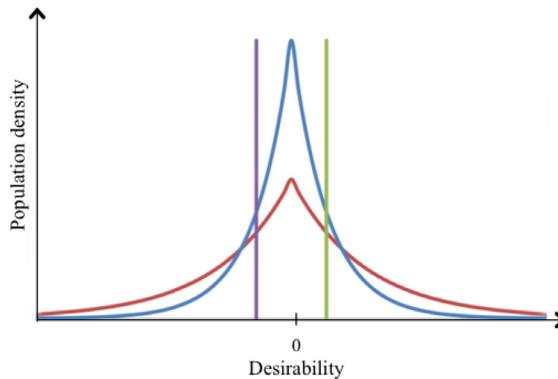


Figure 4

The red curve is the density of the desirability value of the more variable subpopulation B_1 in Example 9, and the blue curve is the density of the less variable subpopulation B_2 . If B_1 and B_2 are of equal size, then the vertical green line is the threshold cutoff for the opposite sex A so that exactly 25% of the composite B population has desirability value above that point. The vertical purple line is the value so that exactly 75% of the B population has desirability value above that point. Note that the desirability values of both drop off exponentially fast from the mean in both directions.

Suppose first that sex A is selective and accepts only the most desirable quarter of individuals in sex B , i.e., $p_A = 0.25$. Using (2) and (3), and noting that $S_2(x) = 2S_1^2(x)$ for $x \geq 0$ yields the following coupled system of ordinary differential equations:

$$\begin{aligned} \frac{dx_1}{dt} &= x_1 \left(\frac{\sqrt{x_1^2 + 2x_1x_2 + 2x_2^2} - x_1}{4x_2} \right) \\ \frac{dx_2}{dt} &= \left(\frac{x_1 + x_2}{4} \right) - x_1 \left(\frac{\sqrt{x_1^2 + 2x_1x_2 + 2x_2^2} - x_1}{4x_2} \right). \end{aligned} \quad (7)$$

No closed-form solution of (7) is known, and Figure 5 illustrates a numerical solution with the initial condition $x_1(0) = x_2(0) = 1$. Note that in this case where sex A is selective, the more variable subpopulation B_1 eventually eclipses the less variable B_2 .

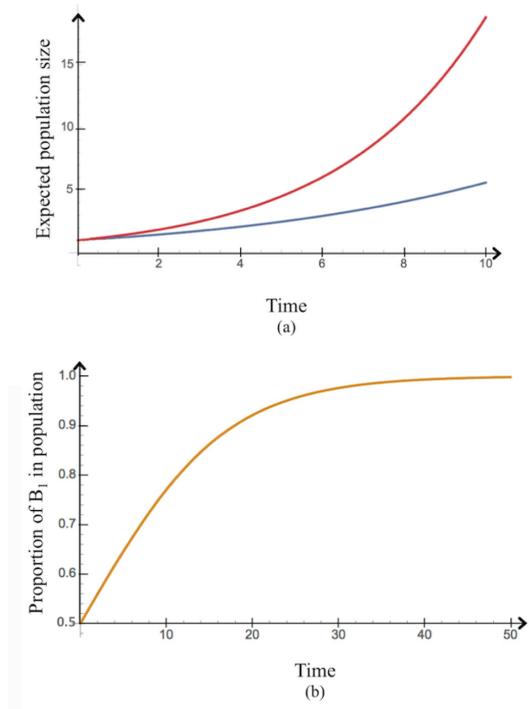


Figure 5

Selective case—Population sizes and ratio. The graphs (a) of the more variable $x_1(t)$ in red and $x_2(t)$ in blue, and (b) the ratio $x_1(t)/(x_1(t) + x_2(t))$ satisfying (7).

Suppose next that sex A is non-selective and accepts only individuals in the most desirable three-quarters of sex B , i.e., $p_A = 0.75$. Using (2) and (3) again, and noting that $S_2(x) = 4S_1(x) - 2S_1^2(x) - 1$ for $x \leq 0$ yields the following system:

$$\begin{aligned} \frac{dx_1}{dt} &= x_1 \left(\frac{x_1 + 4x_2 - \sqrt{x_1^2 + 2x_1x_2 + 2x_2^2}}{4x_2} \right) \\ \frac{dx_2}{dt} &= \left(\frac{3x_1 + 3x_2}{4} \right) - x_1 \left(\frac{x_1 + 4x_2 - \sqrt{x_1^2 + 2x_1x_2 + 2x_2^2}}{4x_2} \right). \end{aligned} \quad (9)$$

Figure 6 illustrates a numerical solution of the coupled system of ODE's (8) with the same initial condition $x_1(0) = x_2(0) = 1$. Note that in this situation where sex A is non-selective, the less variable subpopulation B_2 eventually eclipses the more variable B_1 .

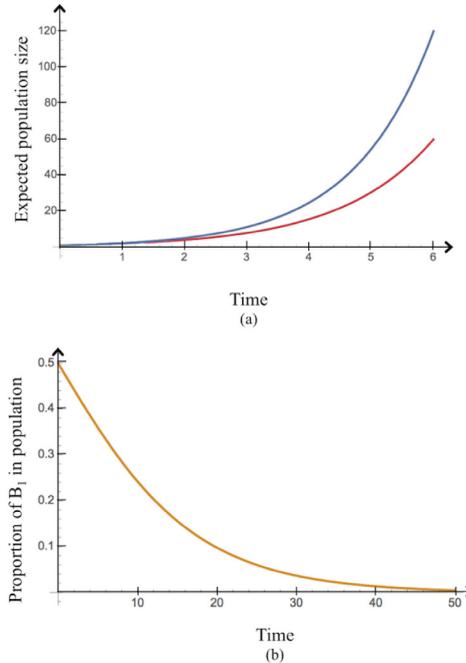


Figure 6

Non-selective case—Population sizes and ratio. The graphs (a) of the more variable subpopulation size $x_1(t)$ in red and less variable $x_2(t)$ in blue, and (b) the ratio $x_1(t)/(x_1(t) + x_2(t))$ satisfying (8).

Note also that the birth process model above also implicitly includes simple *birth-death* processes, via the simple observation that a population growing, for example, at a rate of eight per cent and dying at a rate of three percent, can be viewed as a pure birth process growing at a rate of five per cent.

7. Darwin's Question, Selectivity and Parenting

The selectivity-variability principle introduced above is neutral with respect to the two sexes, and by itself does not explain any differences in variability between them - either that there should be a difference in variability between the sexes, or which sex might be expected to be more variable. But together with two other basic biological tenets, the selectivity-variability principle can perhaps help provide a theory for Darwin's observation and the empirical evidence of greater male variability reported in many subsequent studies.

One of these two additional biological tenets is parenting-selectivity, which posits that a "basic cross-species pattern is that the sex with the slower potential rate of reproduction invests more in parenting, [and] is selective in mate choices [and the] sex with the faster potential rate of reproduction invests less in parenting, [and] is less selective in mate choices" [9, p. 273]. Although the genetic mechanisms of pre- and postcopulatory sexual selection are still far from being fully understood, molecular genetic and genomic tools now enable their detailed experimental testing [1].

The second additional biological tenet is gender-parenting, which says that typical species have less parental investment by males than females, which occurs, for example, in more than ninety-five percent of mammalian species [10]. Combining these two biological maxims with the selectivity-variability principle suggests an answer to Darwin's question. By the gender-parenting tenet, females in typical species invest more in parenting than males, so by the parenting-selectivity tenet females will typically be relatively selective and males relatively non-selective. Then the selectivity-variability principle implies that females in such species will tend toward less variability and the males toward greater variability.

If both sexes in a certain species began with comparable mid-range variability, for example, and if either its females were generally selective ($p_F < 0.5$) or its males were generally non-selective ($p_M < 0.5$), or both, this would have led to the relatively greater male variability observed by Darwin. In the constraints of this cross-species model, therefore, this would offer *two independent explanations* for the appearance of greater

male variability in many species. Unlike other species, of course, in humans, cultural factors may also play a role in the perceived differences in variability between the sexes.

8. Conclusions

The goal here has been neither to challenge nor to confirm Darwin's and other researchers' observations of greater male variability for any given species or any given trait, but rather to propose an elementary mathematical theory based on biological/evolutionary mechanisms that might serve as a starting point to help explain how one sex of a species might tend to evolve with greater variability than the other sex. As such, the contribution here is primarily intended to open the discussion and stimulate further mathematical and statistical modeling and analysis.

9. Acknowledgments

The author is grateful to several referees and to Professors Ron Fox, Igor Rivin, Erika Rogers, Marjorie Senechal, Sergei Tabachnikov, and Julius Van Der Werf for very helpful comments.

References

- [1] M. Andersson and L. W. Simmons, Sexual selection and mate choice, *Trends in Ecology and Evolution* 21 (2006), no. 6, 296–302. doi: 10.1016/j.tree.2006.03.015
- [2] A. J. Bateman, Intra-sexual selection in *Drosophila*, *Heredity* 2 (1948), no. 2, 349-368. doi: 10.1038/hdy.1948.21
- [3] I. M. Bomze, Lotka-Volterra equation and replicator dynamics: new issues in classification, *Biological Cybernetics* 72 (1995), no. 5, 447–453. doi: 10.1007/BF00201420
- [4] J. S. Brown, Why Darwin would have loved evolutionary game theory, *Proceedings of the Royal Society B* 283 (2016), 20160847. doi: 10.1098/rspb.2016.0847
- [5] E. L. Charnov, The genetical evolution of patterns of sexuality: Darwinian fitness, *The American Naturalist* 2 (1979), no. 4, 465-480. doi: 10.1086/283407

- [6] R. Cressman and Y. Tao, The replicator equation and other game dynamics, *Proceedings of the National Academy of Sciences* 111 (2014), Supplement 3, 10810-10817. doi: 10.1073/pnas.1400823111
- [7] C. Darwin, *The Descent of Man and Selection in Relation to Sex*, D. Appleton and Company, New York, 1875.
- [8] G. De Benedictis, G. Carrieri, O. Varcasia, et al., Inherited variability of the mitochondrial genome and successful aging in humans, *Annals of the New York Academy of Sciences* 908 (2006), no. 1, 208-218. doi: 10.1111/j.1749-6632.2000.tb06648.x
- [9] D. C. Geary, Sex differences in social behavior and cognition: Utility of sexual selection for hypothesis generation, *Hormones and Behavior* 49 (2006), 273-275. doi: 10.1016/j.yhbeh.2005.07.014
- [10] D. C. Geary, An evolutionary perspective on sex differences in mathematics and the sciences, *Why aren't more women in science?* (S. Ceci, W. Williams, eds.) American Psychological Association, Washington, DC, 2007, 173-188.
- [11] W. Hill and H. Mulder, Genetic analysis of environmental variation, *Genetics Research* 92 (2010), nos. 5-6, 381-395. doi: 10.1017/S0016672310000546
- [12] T. P. Hill, An evolutionary theory for the variability hypothesis, Available online: <https://arxiv.org/abs/1703.04184>, 2019, accessed on 3 May 2020.
- [13] J. Hofbauer and K. Sigmund, Evolutionary game dynamics, *Bulletin of the American Mathematical Society* 40 (2003), no. 4, 479-519. doi: 10.1090/S0273-0979-03-00988-1
- [14] R. Lande, Models of speciation by sexual selection on polygenic traits, *Proceedings of the National Academy of Sciences*, 78 (1981), no. 6, 3721-3725. doi: 10.1073/pnas.78.6.3721
- [15] J. Marjanovič, Genetics of inherited variability: increasing uniformity by reducing competition. Thesis (Ph.D.), Swedish University of Agricultural Sciences and Wageningen Institute of Animal Sciences of Wageningen University & Research, 2018. doi: 10.18174/442792
- [16] J. Marjanovič, H. A. Mulder, L. Rönnegård, and P. Bijma, Modelling the co-evolution of indirect genetic effects and inherited variability, *Heredity* 121 (2018), 631-647. doi: 10.1038/s41437-018-0068-z

- [17] T. Moriyama, M. V. Relling, and J. J. Yang, Inherited genetic variation in childhood acute lymphoblastic leukemia, *Blood* 125 (2015), no. 26, 3988–3995. doi: 10.1182/blood-2014-12-580001
- [18] H. Mulder, P. Bijma, and W. Hill, Prediction of breeding values and selection responses with genetic heterogeneity of environmental variance, *Genetics* 175 (2007), no. 4, 1895–1910. doi: 10.1534/genetics.106.063743
- [19] H. Mulder, P. Bijma, and W. Hill, Selection for uniformity in livestock by exploiting genetic heterogeneity of residual variance, *Genetics Selection Evolution* 40 (2008), no.37. doi: 10.1186/1297-9686-40-1-37
- [20] K. Reinhold and L. Engqvist, The variability is in the sex chromosomes, *Evolution* 67 (2013), no. 12, 3662–3668. doi: 10.1111/evo.12224
- [21] W. Sandholm, Evolutionary game theory, *Encyclopedia of Complexity and Systems Science* (R. Meyers, ed.), Springer, Heidelberg, 2009, 3176–3205 . doi: 10.1007/978-0-387-30440-3 188
- [22] P. D. Taylor and L. Jonker, Evolutionarily stable strategies and game dynamics, *Mathematical Biosciences* 40 (1978), no. 1, 145–156. doi: 10.1016/0025-5564(78)90077-9
- [23] H. Wainer, The most dangerous equation, *American Scientist* 95 (2007), 249–256.
- [24] J. W. Weibull, *Evolutionary Game Theory*, MIT Press, Cambridge, MA, 1997.
- [25] E. C. Zeeman, Population dynamics from game theory, *Global Theory of Dynamical Systems: Lecture Notes in Mathematics*, vol. 819 (Z. Nitecki, C. Robinson, eds.), Springer, Berlin, 1980, 471–497.