The Expected Value of a Ratio

Finding the expected value of a ratio of random variables that are correlated with one another is much more challenging that finding the expected value of a product. The trick is to rewrite the ratio, by expanding it as a series, so that all random variables are in the numerator. There are (at least) two ways to do this: we can use orthogonal polynomials or Taylor's theorem. Here, I explain both methods and compare them.

Note that $\mathrm{E}(\frac{a}{b})$ is undefined if there is any nonzero probability that $b = 0$. Thus, we will calculate $\mathrm{E}(\frac{a}{b} \mid b \neq 0)$ - the expected value of the ratio conditional on b not equaling zero. This means that the distributions of both a and b that we are working with are conditional distributions, and the probabilities of different values may have to be recalculated accordingly. In evolutionary theory, when we are dealing with $\frac{w}{\overline{w}}$, this condition makes complete sense; since $\overline{w} = 0$ if and only if the population goes extinct - in which case the result *should* be undefined (for the case of migration, the condition is different. See Rice & Papadopoulos (2009).

The orthogonal polynomial approach yields a series that converges faster than the Taylor series and converges for all values of $b > 0$ (which the Taylor series does not). It is also unbiased in the sense that going out to the i^{th} order term yields a polynomial of degree i for which the mean of the polynomial is the same as the mean of $\frac{1}{b}$. The downside of this approach is that the coefficients quickly become rather complicated as we go to higher order terms.

The Taylor series approach yields much simpler coefficients. However, it converges more slowly and – critically – does not converge at all for some values of b that are biologically reasonable. It also gives us a series for which partial series (going out only to some finite term) are biased estimators of $\frac{1}{b}$.

Orthogonal Polynomial Approach

For a ratio $\frac{a}{b}$, we first project $\frac{1}{b}$ onto the orthogonal polynomials for b, then multiply the resulting series by a.

Expanding a reciprocal

We need to write $\frac{1}{b}$ as a series with the random variable, b, always in the numerator. To do this, we construct orthogonal polynomials in b and project $\frac{1}{b}$ into these. Defining $\mathbf{B}^{[i]}$ as the i^{th} order orthogonal polynomial in b, we can write $\frac{1}{b}$ as:

$$\frac{1}{b} = \frac{1}{{}^{h}\widehat{b}} + \sum_{i=1}^{n-1} \left\langle\!\!\left\langle \frac{\frac{1}{b}}{\mathbf{B}^{[i]}} \right\rangle\!\!\right\rangle \mathbf{B}^{[i]} \tag{1}$$

where ${}^{h}\widehat{b}$ is the probability harmonic mean of b. Because $\frac{1}{b}$ is a function of b, there is no residual variation in $\frac{1}{b}$ once b is specified. This means that, given n distinct values of b, the series will converge exactly by the $(n-1)^{th}$ order term.

Evaluating the series in Equation 1 requires finding the regressions of $\frac{1}{b}$ on the $\mathbf{B}$ polynomials. We can calculate this from the covariance:

$$\left\langle\!\!\left\langle \tfrac{1}{b}, \mathbf{B}^{[i]} \right\rangle\!\!\right\rangle = k_i \left(\frac{\widehat{b} - {}^{h}\widehat{b}}{{}^{h}\widehat{b}\,\widehat{b}} + \sum_{j=1}^{i-1} \frac{\langle\!\langle {}^{2}\mathbf{B}^{[j]} \rangle\!\rangle}{k_{j+1}k_j} \right) \tag{2}$$

The k terms in Equation 2 are the constant terms in the orthogonal polynomials $\mathbf{B}$. Specifically, k_i contains all constant terms (not multiples of b) in the polynomial $\mathbf{B}^{[i]}$. The first three k terms are:

$$\begin{aligned} k_0 &= 1 \\ k_1 &= -\widehat{b} \\ k_2 &= \widehat{b}^2 + \widehat{b}\frac{\langle\!\langle {}^{3}b \rangle\!\rangle}{\langle\!\langle {}^{2}b \rangle\!\rangle} - \langle\!\langle {}^{2}b \rangle\!\rangle \end{aligned} \tag{3}$$

Subsequent k terms can be constructed from lower order ones using the following relation:

$$k_n = (-\widehat{b})^n - \sum_{i=1}^{n-1} \left\langle\!\!\left\langle \frac{\mathbf{B}^{[1]n}}{\mathbf{B}^{[i]}} \right\rangle\!\!\right\rangle k_i - \langle\!\langle {}^{n}b \rangle\!\rangle \tag{4}$$

Reintroducing a

With this series expression for $\frac{1}{b}$, we can find the expected value of $\frac{a}{b}$ by multiplying the series by a and taking the expected value. The result can be written in either of two ways:

$$\mathrm{E}\left(\frac{a}{b} \middle| b \neq 0 \right) = \frac{\widehat{a}}{{}^{h}\widehat{b}} + \sum_{i=1}^{n-1} \left\langle\!\!\left\langle \frac{\frac{1}{b}}{\mathbf{B}^{[i]}} \right\rangle\!\!\right\rangle \left\langle\!\!\left\langle a, \mathbf{B}^{[i]} \right\rangle\!\!\right\rangle \tag{5}$$

or, grouping the $\frac{1}{\langle\!\langle {}^{2}\mathbf{B} \rangle\!\rangle}$ terms with the $\langle\!\langle a, \mathbf{B} \rangle\!\rangle$ terms:

$$\mathrm{E}\left(\frac{a}{b}\middle| b \neq 0\right) = \frac{\widehat{a}}{{}^{h}\widehat{b}} + \sum_{i=1}^{n-1} \left\langle\!\!\left\langle \frac{1}{b}, \mathbf{B}^{[i]} \right\rangle\!\!\right\rangle \left\langle\!\!\left\langle \frac{a}{\mathbf{B}^{[i]}} \right\rangle\!\!\right\rangle \tag{6}$$

The first three terms (0^{th}, 1^{st}, and 2^{nd} order) of this series give the approximation:

$$\mathrm{E}\left(\frac{a}{b}\middle| b \neq 0\right) \approx \frac{\widehat{a}}{{}^{h}\widehat{b}} - \frac{\widehat{b} - {}^{h}\widehat{b}}{\langle\!\langle {}^{2}b \rangle\!\rangle\, {}^{h}\widehat{b}} \langle\!\langle a, b \rangle\!\rangle + \frac{1}{\langle\!\langle {}^{2}\mathbf{B}^{[2]} \rangle\!\rangle} \left(\frac{k_2}{{}^{h}\widehat{b}} - \widehat{b} - \frac{\langle\!\langle {}^{3}b \rangle\!\rangle}{\langle\!\langle {}^{2}b \rangle\!\rangle} \right) \langle\!\langle a, \mathbf{B}^{[2]} \rangle\!\rangle \tag{7}$$

or, equivalently:

$$\mathrm{E}\left(\frac{a}{b}\middle| b \neq 0\right) \approx \frac{\widehat{a}}{{}^{h}\widehat{b}} - \frac{\widehat{b} - {}^{h}\widehat{b}}{{}^{h}\widehat{b}} \left\langle\!\!\left\langle \frac{a}{b} \right\rangle\!\!\right\rangle + \left(\frac{k_2}{{}^{h}\widehat{b}} - \widehat{b} - \frac{\langle\!\langle {}^{3}b \rangle\!\rangle}{\langle\!\langle {}^{2}b \rangle\!\rangle} \right) \left\langle\!\!\left\langle \frac{a}{\mathbf{B}^{[2]}} \right\rangle\!\!\right\rangle \tag{8}$$

Taylor Series Approach

Here, we use Taylor's theorem to expand the ratio $\frac{a}{b}$ around the mean of b. We will make use of the "delta method", which involves replacing random variables a and b with new variables, a^* and b^*, that that are 0 at the mean values of $a$and b. So:

Consider random variables a and b. We can write these as:

$$\begin{aligned} a &= \widehat{a} + a^* \\ b &= \widehat{b} + b^* \end{aligned} \tag{9}$$

We are still measuring the same things, we just shift the axes so that 0 is the expected value (*e.g.* if the expected number of descendants is 2, then we measure the actual number by how much it differs from 2; if the individual ends up leaving just 1 descendant, then $a^* = -1$).

Using the definitions in Equation 9. we can now write:

$$\mathrm{E}\left(\frac{a}{b}\middle| b \neq 0\right) = \mathrm{E}\left(\frac{\widehat{a} + a^*}{\widehat{b} + b^*}\right) = \mathrm{E}\left(\frac{\widehat{a}}{\widehat{b}} \frac{1 + \frac{a^*}{\widehat{a}}}{1 + \frac{b^*}{\widehat{b}}}\right) \tag{10}$$

This approach is sometimes called the "delta method", since what we are calling a^* is often represented as δa. I use a different notation here since δ has another meaning in our papers.

The key here is to note that *the expected values, $\widehat{a}$ and $\widehat{b}$, are not random variables* - they thus can come outside the expectation on the righthand side of Equation 10, yielding:

$$\mathrm{E}\left(\frac{a}{b}\middle|b \neq 0\right) = \frac{\widehat{a}}{\widehat{b}}\,\mathrm{E}\left(\frac{1+\frac{a^*}{\widehat{a}}}{1+\frac{b^*}{\widehat{b}}}\right) = \frac{\widehat{a}}{\widehat{b}}\,\mathrm{E}\left[\left(1+\frac{a^*}{\widehat{a}}\right)\left(1+\frac{b^*}{\widehat{b}}\right)^{-1}\right] \tag{11}$$

Multiplying out the term in the square brackets yields:

$$\mathrm{E}\left(\frac{a}{b}\middle|b \neq 0\right) = \frac{\widehat{a}}{\widehat{b}}\,\mathrm{E}\left[\left(1+\frac{b^*}{\widehat{b}}\right)^{-1}\right] + \frac{1}{\widehat{b}}\,\mathrm{E}\left[a^*\left(1+\frac{b^*}{\widehat{b}}\right)^{-1}\right] \tag{12}$$

Now: Note that, by the definition of the harmonic mean, $\mathrm{E}(\frac{1}{b}) = \frac{1}{{}^h\widehat{b}}$, where ${}^h\widehat{b}$ is the harmonic mean of b. We can use Equation 11 to find $\mathrm{E}(\frac{1}{b})$ by setting $a = 1$ (so $\widehat{a} = 1$ and $a^* = 0$). Doing so, we find:

$$\mathrm{E}\left(\frac{1}{b}\middle|b \neq 0\right) \equiv \frac{1}{{}^h\widehat{b}} = \frac{1}{\widehat{b}}\,\mathrm{E}\left[\left(1+\frac{b^*}{\widehat{b}}\right)^{-1}\right] \tag{13}$$

We can now rewrite the first term on the righthand side of Equation 12 by using Equation 13:

$$\mathrm{E}\left(\frac{a}{b}\middle|b \neq 0\right) = \frac{\widehat{a}}{{}^h\widehat{b}} + \frac{1}{\widehat{b}}\,\mathrm{E}\left[a^*\left(1+\frac{b^*}{\widehat{b}}\right)^{-1}\right] \tag{14}$$

We now have to deal with the term $\mathrm{E}[a^*(1+\frac{b^*}{\widehat{b}})^{-1}]$. So long as $b^* < \widehat{b}$ (*i.e.* $b < 2\cdot\widehat{b}$, we can expand $(1+\frac{b^*}{\widehat{b}})^{-1}$ as a Taylor series in b^*. Defining:

$$f_{b^*} = (1+\frac{b^*}{\widehat{b}})^{-1} \tag{15}$$

Taylor's theorem yields:

$$f_{b^*} = 1 + \sum_{i=1}^{\infty}(-1)^i\frac{{b^*}^i}{\widehat{b}^i} \tag{16}$$

It is important to note that the use of Taylor's theorem here is not applicable in all cases. Specifically, Equation 16 does not converge to Equation 15 if $b^* \geqslant \widehat{b}$ (in other words, if $b \geqslant 2\widehat{b}$).

In such cases, we can use the calculus of Finite Differences, described in the last section below.

When we can use the Taylor expansion in Equation 16, then we get:

$$\mathrm{E}\left[a^*\left(1+\frac{b^*}{\widehat{b}}\right)^{-1}\right]=\mathrm{E}\left[a^*+\sum_{i=1}^{\infty}(-1)^i\frac{a^*{b^*}^i}{\widehat{b}^i}\right] \tag{17}$$

Given the definitions of a^* and b^* from Equation 9, we know that $\mathrm{E}(a^*)=0$, $\mathrm{E}(a^*b^*)=\mathrm{cov}ab$, and, in general, $\mathrm{E}(a^*{b^*}^i)$ is the the mixed central moment defined as $\mathrm{E}\{[a-\widehat{a}][b-\widehat{b}]^i\}$. In the notation of our papers (from 2009 on), this last moment is written: $\langle\!\langle a,{}^i b\rangle\!\rangle$. We can now rewrite Equation 17 as:

$$\mathrm{E}\left[a^*\left(1+\frac{b^*}{\widehat{b}}\right)^{-1}\right]=\sum_{i=1}^{\infty}(-1)^i\frac{\langle\!\langle a,{}^i b\rangle\!\rangle}{\widehat{b}^i} \tag{18}$$

Substituting Equation 18 into Equation 14 gives the equation for the expected value of the ratio:

$$\mathrm{E}\left(\frac{a}{b}\middle|b\neq 0\right)=\frac{\widehat{a}}{\widehat{b}}+\sum_{i=1}^{\infty}(-1)^i\frac{\langle\!\langle a,{}^i b\rangle\!\rangle}{\widehat{b}^{i+1}} \tag{19}$$

An alternate representation

Equation 19 is the form that we use in our papers. It has the advantage of isolating the first term, capturing selection, and then following this with a series of terms that collapse to moments of the individual fitness distributions.

For other applications, though, it is sometimes useful to write the result so that the first term does not involve the harmonic mean. To do this, we simply substitute the series expansion in Equation 16 directly into the far righthand part of Equation 11. Denoting the i^{th} central moment of b by $\langle\!\langle {}^i b\rangle\!\rangle$ (so that $\langle\!\langle {}^1 b\rangle\!\rangle=0$), this yields:

$$\mathrm{E}\left(\frac{a}{b}\middle|b\neq 0\right)=\frac{\widehat{a}}{\widehat{b}}+\sum_{i=1}^{\infty}(-1)^i\frac{\widehat{a}\langle\!\langle {}^i b\rangle\!\rangle+\langle\!\langle a,{}^i b\rangle\!\rangle}{\widehat{b}^{i+1}} \tag{20}$$

Comparing the Orthogonal Polynomial and Taylor Series Methods

Figure 1 shows the results of applying Equation 1 (on the left) and Equation 16 (on the right) to a set of values (black dots) lying within the region in which the Taylor series converges. For each graph, the weighted sum of squares (wss) gives the overall fit of the curve to the points (a smaller value of wss means a better fit).

For each order, the orthogonal polynomial series fits the values better than does the Taylor series. In fact, the first order orthogonal polynomial approximation (a straight line) fits the data better than does the second order Taylor series approximation (a quadratic function). Similarly, the second order orthogonal polynomial function fits the values better than does the third order Taylor series.

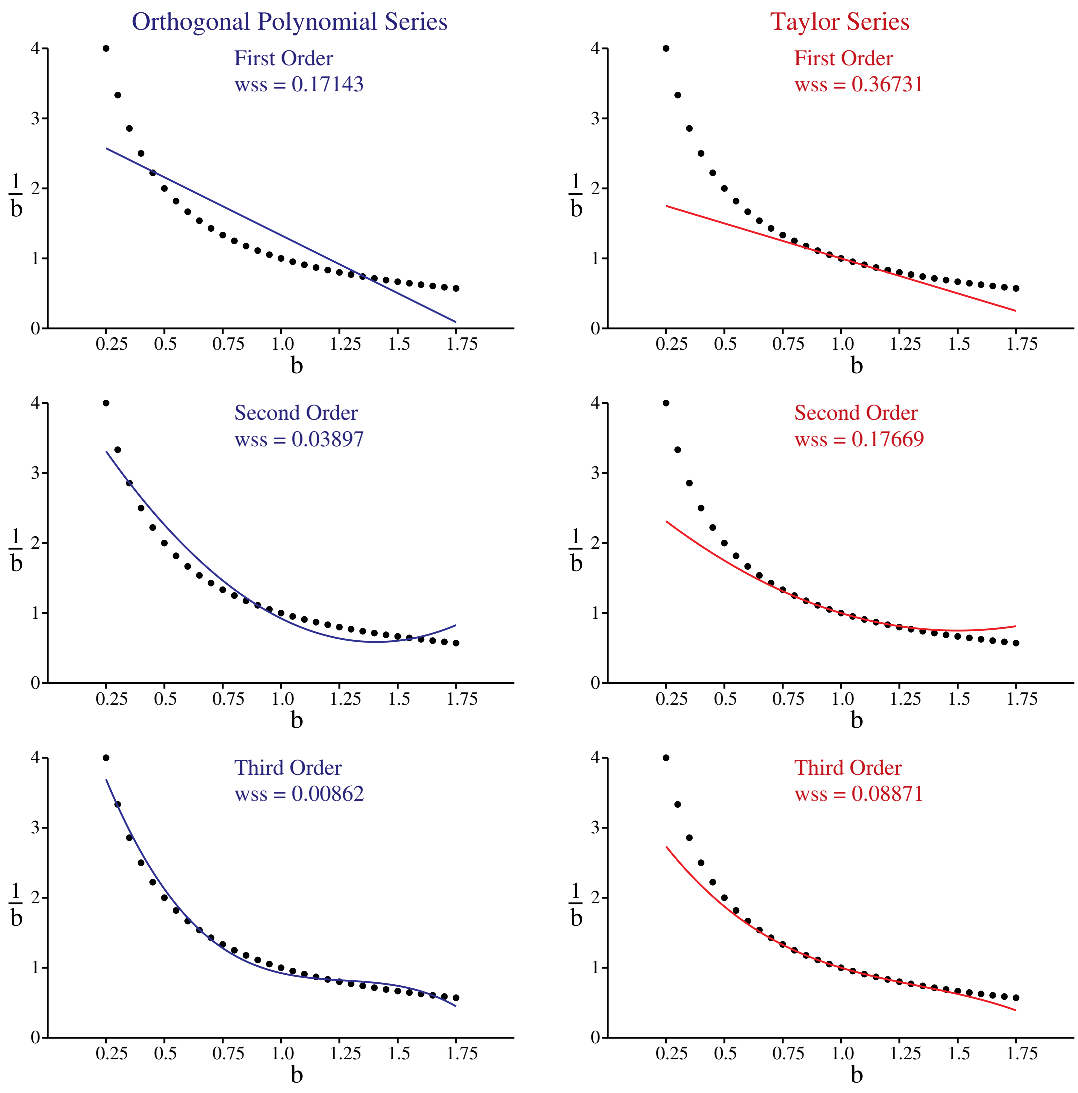

Figure 1: Comparison of orthogonal polynomial (blue, left) and Taylor series (red, right) approximations (up to 3^{rd} order) to the function $\frac{1}{b}$, given a distribution of values of b (black dots) with a mean of 1. The weighted sum of squares (wss) is the sum of squared deviations of the actual values from those predicted by the series, weighted by the density of points at a given value of b. A smaller value of wss indicates a better fit.

The figure also illustrates why the Taylor series approximation is biased. The first order Taylor series approximation is a tangent line to the actual function. Since the function $f(b) = \frac{1}{b}$ is convex, the tangent line lies below the curve everywhere other than the point of contact. Higher order Taylor series approximations fit the points better, but they remain biased in the sense that the mean of the Taylor series approximations will not be the same as the mean of the actual values.

By contrast, the mean of the orthogonal polynomial approximation will always be equal to the mean of the actual values, and the n^{th} order orthogonal polynomial series will be the best fit n^{th} order polynomial to the actual values.

The reason for the difference is that the orthogonal polynomial series is, at every step, finding the best fit polynomial to all of the given values. The Taylor series, however, is building approximations to the function in the *neighborhood* of a chosen point.

If our goal is to build axiomatic theories, then the orthogonal polynomial series is preferable, because – for any finite population – it will converge exactly in a finite number of terms.

If our goal is to find an acceptable approximation that balances accuracy with tractability, then which approach we should use depends on the nature of the problem.

For questions in population biology, we want to consider the entire population, not just a few individuals that are close to the mean. We also want to avoid systematic bias. We therefore should use the orthogonal polynomial approach even when building approximate models.

For some problems, however, our goal really is to locally approximate a function in immediate vicinity of a given point. For problems such as these, which include local stability analysis of a dynamical system, the Taylor series approach may be better.

Appendix

Proof of Equation 2

Because the polynomials $\mathbf{B}^{[i]}$ for $i > 0$ all have means of 0, we need only find the products of $\frac{1}{b}$ and each polynomial.

Multiplying the polynomial $\mathbf{B}^{[n]}$ by $\frac{1}{b}$ reduces the order of each term by 1. This new expression is not a polynomial, since $\mathbf{B}^{[n]}$ has a constant term, k_n, which becomes $\frac{k_n}{b}$. We can write the result as:

$$\frac{1}{b}\mathbf{B}^{[n]} = \mathbf{B}^{[n-1]} + \mathbf{Q}^{[n-2]} + \frac{k_n}{b} \tag{21}$$

Where $\mathbf{Q}^{[n-2]}$ is some polynomial of order $n-2$ (not necessarily orthogonal to the $\mathbf{B}$ polynomials).

Note that, since $\mathbf{B}^{[n]}$ is orthogonal to all polynomials of order $< n$:

$$
\begin{aligned}
\left\langle\!\!\left\langle \frac{1}{b}\mathbf{B}^{[n]}, \mathbf{B}^{[n-1]} \right\rangle\!\!\right\rangle &= \left\langle\!\!\left\langle \mathbf{B}^{[n-1]} + \mathbf{Q}^{[n-2]} + \frac{k_n}{b}, \mathbf{B}^{[n-1]} \right\rangle\!\!\right\rangle \\
&= \langle\!\langle {}^2\mathbf{B}^{[n-1]} \rangle\!\rangle + k_n \left\langle\!\!\left\langle \frac{1}{b}, \mathbf{B}^{[n-1]} \right\rangle\!\!\right\rangle
\end{aligned} \tag{22}
$$

Because $\widehat{\mathbf{B}^{[n]}} = 0 \ \forall n$, we can switch the position of $\frac{1}{b}$:

$$
\left\langle\!\!\left\langle \frac{1}{b}\mathbf{B}^{[n]}, \mathbf{B}^{[n-1]} \right\rangle\!\!\right\rangle = \left\langle\!\!\left\langle \mathbf{B}^{[n]}, \frac{1}{b}\mathbf{B}^{[n-1]} \right\rangle\!\!\right\rangle \tag{23}
$$

Applying Equation 21 to $\frac{1}{b}\mathbf{B}^{[n-1]}$ yields:

$$
\begin{aligned}
\left\langle\!\!\left\langle \frac{1}{b}\mathbf{B}^{[n]}, \mathbf{B}^{[n-1]} \right\rangle\!\!\right\rangle &= \left\langle\!\!\left\langle \mathbf{B}^{[n]}, \mathbf{B}^{[n-2]} + \mathbf{Q}^{[n-3]} + \frac{k_{n-1}}{b} \right\rangle\!\!\right\rangle \\
&= k_{n-1} \left\langle\!\!\left\langle \frac{1}{b}, \mathbf{B}^{[n]} \right\rangle\!\!\right\rangle
\end{aligned} \tag{24}
$$

Combining Equations 22 and 24, we can write the covariance of $\frac{1}{b}$ with $\mathbf{B}^{[n]}$ in terms of its covariance with $\mathbf{B}^{[n-1]}$:

$$
\left\langle\!\!\left\langle \frac{1}{b}, \mathbf{B}^{[n]} \right\rangle\!\!\right\rangle = \frac{\langle\!\langle {}^2\mathbf{B}^{[n-1]} \rangle\!\rangle}{k_{n-1}} + \frac{k_n}{k_{n-1}} \left\langle\!\!\left\langle \frac{1}{b}, \mathbf{B}^{[n-1]} \right\rangle\!\!\right\rangle \tag{25}
$$

We can now use Equation 25 recursively to write $\langle\!\langle \frac{1}{b}, \mathbf{B}^{[n-1]} \rangle\!\rangle$ in terms of $\langle\!\langle \frac{1}{b}, \mathbf{B}^{[n-2]} \rangle\!\rangle$. Substituting this back into Equation 25 we get:

$$
\left\langle\!\!\left\langle \frac{1}{b}, \mathbf{B}^{[n]} \right\rangle\!\!\right\rangle = \frac{\langle\!\langle {}^2\mathbf{B}^{[n-1]} \rangle\!\rangle}{k_{n-1}} + \frac{k_n \langle\!\langle {}^2\mathbf{B}^{[n-2]} \rangle\!\rangle}{k_{n-1}k_{n-2}} + \frac{k_n}{k_{n-2}} \left\langle\!\!\left\langle \frac{1}{b}, \mathbf{B}^{[n-2]} \right\rangle\!\!\right\rangle \tag{26}
$$

Repeating this process, we work down to terms with $\mathbf{B}^{[1]}$:

$$
\left\langle\!\!\left\langle \frac{1}{b}, \mathbf{B}^{[n]} \right\rangle\!\!\right\rangle = \frac{\langle\!\langle {}^2\mathbf{B}^{[n-1]} \rangle\!\rangle}{k_{n-1}} + \frac{k_n \langle\!\langle {}^2\mathbf{B}^{[n-2]} \rangle\!\rangle}{k_{n-1}k_{n-2}} + \frac{k_n \langle\!\langle {}^2\mathbf{B}^{[n-3]} \rangle\!\rangle}{k_{n-2}k_{n-3}} + \cdots + \frac{k_n \langle\!\langle {}^2\mathbf{B}^{[1]} \rangle\!\rangle}{k_2 k_1} + \frac{k_n}{k_1} \left\langle\!\!\left\langle \frac{1}{b}, \mathbf{B}^{[1]} \right\rangle\!\!\right\rangle \tag{27}
$$

Now note that, since $k_1 = -\widehat{b}$ and $\mathbf{B}^{[1]} = b - \widehat{b}$:

$$\frac{k_n}{k_1}\left\langle\!\!\left\langle \frac{1}{b}, \mathbf{B}^{[1]} \right\rangle\!\!\right\rangle = -\frac{k_n}{\widehat{b}}\left(1 - \frac{\widehat{b}}{{}^{h}\widehat{b}}\right) = k_n\left(\frac{1}{{}^{h}\widehat{b}} - \frac{1}{\widehat{b}}\right) = k_n\left(\frac{\widehat{b} - {}^{h}\widehat{b}}{{}^{h}\widehat{b}\,\widehat{b}}\right) \tag{28}$$

Combining Equations 27 and 28 yields Equation 2

Solving for the k terms (Equation 4)

The constant term, k_n, contains all of the terms in $\mathbf{B}^{[n]}$ that do not contain b. For $\mathbf{B}^{[0]}$ and $\mathbf{B}^{[1]}$, these are just:

$$\begin{array}{rclcl} \mathbf{B}^{[0]} & = & 1 & \Rightarrow & k_0 = 1 \\ \mathbf{B}^{[1]} & = & b - \widehat{b} & \Rightarrow & k_1 = -\widehat{b} \end{array} \tag{29}$$

For higher order polynomials, we need to sum the contributions of the polynomials within them. So, for k_2, we have:

$$\begin{aligned} \mathbf{B}^{[2]} &= \mathbf{B}^{[1]^2} - \frac{\langle\!\langle {}^3 b \rangle\!\rangle}{\langle\!\langle {}^2 b \rangle\!\rangle}\mathbf{B}^{[1]} - \langle\!\langle {}^2 b \rangle\!\rangle \\ &= b^2 - 2\widehat{b}b + \widehat{b}^2 - \frac{\langle\!\langle {}^3 b \rangle\!\rangle}{\langle\!\langle {}^2 b \rangle\!\rangle} b + \frac{\langle\!\langle {}^3 b \rangle\!\rangle}{\langle\!\langle {}^2 b \rangle\!\rangle}\widehat{b} - \langle\!\langle {}^2 b \rangle\!\rangle \end{aligned} \tag{30}$$

The constant terms, in blue in Equation 30, constitute k_2:

$$k_2 = \widehat{b}^2 + \frac{\langle\!\langle {}^3 b \rangle\!\rangle}{\langle\!\langle {}^2 b \rangle\!\rangle}\widehat{b} - \langle\!\langle {}^2 b \rangle\!\rangle \tag{31}$$

In general, the **B** polynomials are given by:

$$\mathbf{B}^{[n]} = \mathbf{B}^{[1]^n} - \sum_{i=1}^{n-1}\left\langle\!\!\left\langle \begin{matrix} \mathbf{B}^{[1]^n} \\ \mathbf{B}^{[i]} \end{matrix} \right\rangle\!\!\right\rangle \mathbf{B}^{[i]} - \langle\!\langle {}^n b \rangle\!\rangle \tag{32}$$

for which the constant terms are:

$$k_n = \left(-\widehat{b}\right)^n + \sum_{i=1}^{n-1}\left\langle\!\!\left\langle \begin{matrix} \mathbf{B}^{[1]^n} \\ \mathbf{B}^{[i]} \end{matrix} \right\rangle\!\!\right\rangle k_i - \langle\!\langle {}^n b \rangle\!\rangle \tag{33}$$

Finite Differences solution for $b \geqslant 2\widehat{b}$

As noted above, the Taylor series expansion of Equation 15, given in Equation 16, converges to the function only in the range $-\widehat{b} < b^* < \widehat{b}$ or, in the original variables, $0 < b < 2\widehat{b}$ (Figure 2). The lower bound of this range, $b = 0$, corresponds to the case in which the expected value of the ratio is undefined (and which we excluded by conditioning on $b \neq 0$). There is no biological reason, though, to treat the upper bound as different from any other value of b.

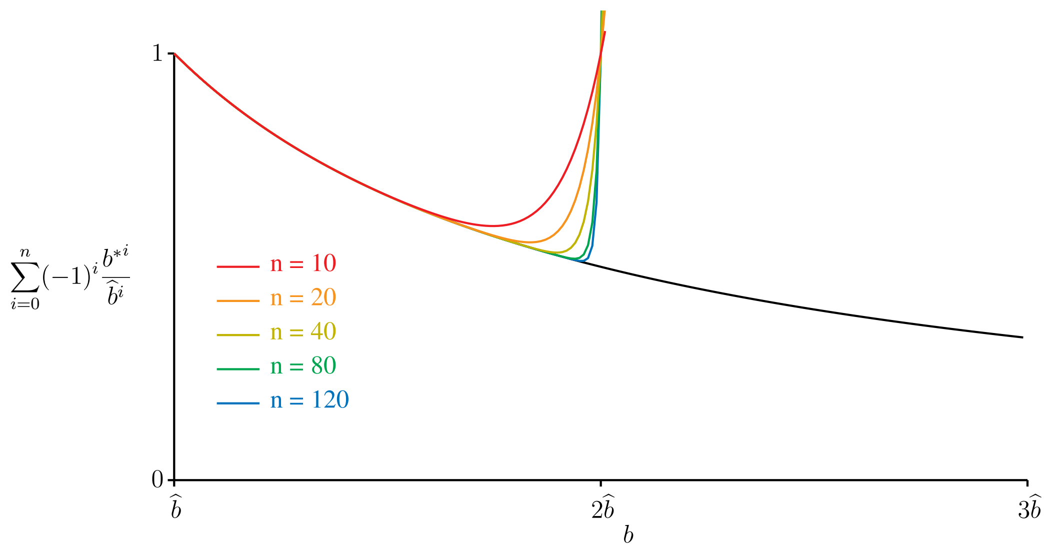

Figure 2: Result of adding increasong numbers of terms to the Taylor expansion.

Figure 2 shows how the Taylor series behaves at $b = 2\widehat{b}$; the series in Equation 16 is clearly not appropriate for values of b^* beyond this threshold.

One way around this problem is to use finite differences. For a function, f, that is bounded on $[0, \infty]$, Hille and Phillips Hille & Phillips (1957) (pg. 533) showed that the following series of finite differences converges to f.

$$f(a + x) = \lim_{h \to 0} \sum_{i=0}^{\infty} \frac{x^i}{i!} \frac{\Delta_h^i f(a)}{h^i} \tag{34}$$

where Δ_h^i is the finite difference opperator of degree i and step size h, defined as:

$$\Delta_h^i f(a) = \sum_{j=0}^{i} (-1)^j \binom{i}{j} f(a + (i - j) \cdot h) \tag{35}$$

(Note that some authors, including (Hille & Phillips, 1957), include the $\frac{1}{h}$ term in the definition of Δ_h^i).

As $h \to 0$, each term in the series in Equation 34 converges to the corresponding term in in the Taylor series. However, the convergence behavior of the entire series is different, with the series in Equation 34 converging to f even where the Taylor series does not.

Applying Equation 35 to Equation 15, evaluated at $b^* = 0$, yields:

$$\Delta_h^i f(0) = (-1)^i \frac{i! \cdot \widehat{b} \cdot h^i}{\prod_{j=0}^{i} (\widehat{b} + j \cdot h)} \tag{36}$$

We can now use Equation 34, with the finite difference operator from Equation 36, to get a series approximation for Equation 15:

$$f_{b^*} = \lim_{h \to 0} \sum_{i=0}^{\infty} (-1)^i \frac{(b^*)^i \widehat{b}}{\prod_{j=0}^{i} (\widehat{b} + j \cdot h)} \tag{37}$$

Note that the series in Equation 37 would be identical to the Taylor series, Equation 16, if the $\lim_{h \to 0}$ was moved inside the summation. This example is thus a good illustration of the fact that series that converge pointwise do not necessarily converge uniformly; and that the order of limits matters (since the summation is basically a limit as $i \to \infty$).

This is illustrated in Figure 3, which shows the behavior of the series in Equation 37 for different values of h. As $h \to 0$, an increasing number of terms must be added to the series to get a good approximation. However, for any nonzero h, the series will eventually converge to the function.

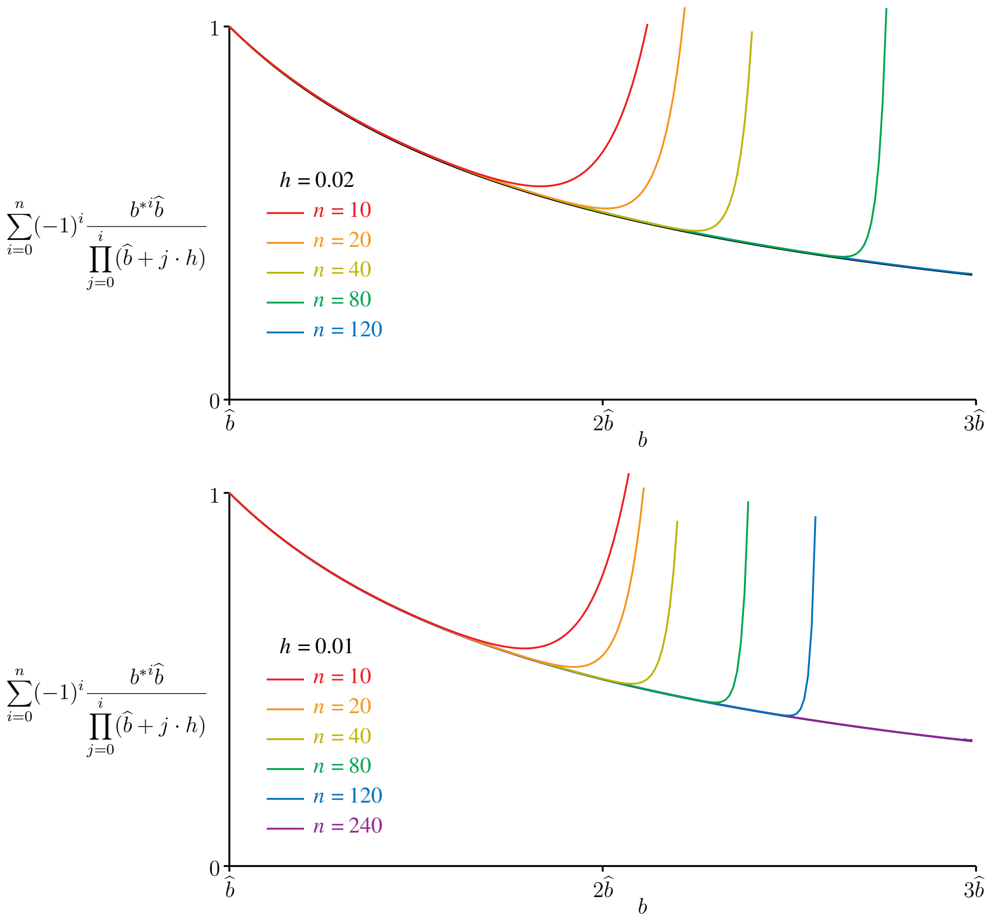

Figure 3: Result of adding increasing numbers of terms to the series in Equation (), for different values of h.

References

Hille, E. & Phillips, R. S. 1957: Functional Analysis and Semi Groups. American Mathematical Society, Providence, Rhode Island.

Rice, S. H. & Papadopoulos, A. 2009: Evolution with Stochastic Fitness and Stochastic Migration. PLoS ONE 4(10):e7130.