The expected value of the ratio of correlated random variables

Sean H. Rice
Texas Tech University
July 15^{th}, 2015

The series equation for the expected value of a ratio of two random variables that are not independent of one another (such as w and $\overline{w}$) plays an important role in the analysis of the axiomatic theory. Since the derivation in Rice (2008) is rather brief and skips a number of steps, I give a detailed step-by-step derivation below.

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

$$\begin{aligned} a &= \mathrm{E}(a) + a^* \\ b &= \mathrm{E}(b) + b^* \end{aligned} \tag{1}$$

Essentially, we are replacing variables a and b with new variables, a^* and b^*. 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 $a^* = -1$).

Now, 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, 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)).

Using the definitions in Equation 1. we can now write:

$$\mathrm{E}\left(\frac{a}{b}\middle| b \neq 0\right) = \mathrm{E}\left(\frac{\mathrm{E}(a) + a^*}{\mathrm{E}(b) + b^*}\right) = \mathrm{E}\left(\frac{\mathrm{E}(a)}{\mathrm{E}(b)}\,\frac{1 + \frac{a^*}{\mathrm{E}(a)}}{1 + \frac{b^*}{\mathrm{E}(b)}}\right) \tag{2}$$

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 <u>*the expected values,* $\mathrm{E}(a)$ *and* $\mathrm{E}(b)$*, are not random variables*</u> - they thus can come outside the expectation on the righthand side of Equation 2, yielding:

$$\mathrm{E}\left(\frac{a}{b}\middle| b \neq 0\right) = \frac{\mathrm{E}(a)}{\mathrm{E}(b)}\,\mathrm{E}\left(\frac{1 + \frac{a^*}{\mathrm{E}(a)}}{1 + \frac{b^*}{\mathrm{E}(b)}}\right) = \frac{\mathrm{E}(a)}{\mathrm{E}(b)}\,\mathrm{E}\left[\left(1 + \frac{a^*}{\mathrm{E}(a)}\right)\left(1 + \frac{b^*}{\mathrm{E}(b)}\right)^{-1}\right] \tag{3}$$

Multiplying out the term in the square brackets yields:

$$\mathrm{E}\left(\frac{a}{b}\middle| b \neq 0\right) = \frac{\mathrm{E}(a)}{\mathrm{E}(b)}\ \mathrm{E}\left[\left(1+\frac{b^*}{\mathrm{E}(b)}\right)^{-1}\right] + \frac{1}{\mathrm{E}(b)}\ \mathrm{E}\left[a^*\left(1+\frac{b^*}{\mathrm{E}(b)}\right)^{-1}\right] \tag{4}$$

Now: Note that, by the definition of the harmonic mean, $\mathrm{E}(\frac{1}{b}) = \frac{1}{\mathrm{H}(b)}$, where $\mathrm{H}(b)$ is the harmonic mean of b. We can use Equation 3 to find $\mathrm{E}(\frac{1}{b})$ by setting $a = 1$ (so $\mathrm{E}(a) = 1$ and $a^* = 0$). Doing so, we find:

$$\mathrm{E}\left(\frac{1}{b}\middle| b \neq 0\right) \equiv \frac{1}{\mathrm{H}(b)} = \frac{1}{\mathrm{E}(b)}\ \mathrm{E}\left[\left(1+\frac{b^*}{\mathrm{E}(b)}\right)^{-1}\right] \tag{5}$$

We can now rewrite the first term on the righthand side of Equation 4 by using Equation 5:

$$\mathrm{E}\left(\frac{a}{b}\middle| b \neq 0\right) = \frac{\mathrm{E}(a)}{\mathrm{H}(b)} + \frac{1}{\mathrm{E}(b)}\ \mathrm{E}\left[a^*\left(1+\frac{b^*}{\mathrm{E}(b)}\right)^{-1}\right] \tag{6}$$

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

$$f_{b^*} = \left(1+\frac{b^*}{\mathrm{E}(b)}\right)^{-1} \tag{7}$$

Taylor's theorem yields:

$$f_{b^*} = 1 + \sum_{i=1}^{\infty}(-1)^i\frac{{b^*}^i}{\mathrm{E}(b)^i} \tag{8}$$

It is important to note that the use of Taylor's theorem here is not applicable in all cases. Specifically, Equation 8 does not converge to Equation 7 if $b^* \geqslant \mathrm{E}(b)$. (Thanks to Kevin Van Horn for pointing this out to me). In such cases, we use the calculus of Finite Differences, described in the last section below.

When we can use the Taylor expansion in Equation 8, then we get:

$$\mathrm{E}\left[a^*\left(1+\frac{b^*}{\mathrm{E}(b)}\right)^{-1}\right] = \mathrm{E}\left[a^* + \sum_{i=1}^{\infty}(-1)^i\frac{a^*{b^*}^i}{\mathrm{E}(b)^i}\right] \tag{9}$$

Given the definitions of a^* and b^* from Equation 1, we know that $\mathrm{E}(a^*) = 0$, $\mathrm{E}(a^*b^*) = \mathrm{cov}(a,b)$, and, in general, $\mathrm{E}(a^*{b^*}^i)$ is the the mixed central moment defined as $\mathrm{E}\{[a-\mathrm{E}(a)][b-\mathrm{E}(b)]^i\}$. 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 9 as:

$$\mathrm{E}\left[a^*\left(1+\frac{b^*}{\mathrm{E}(b)}\right)^{-1}\right]=\sum_{i=1}^{\infty}(-1)^i\frac{\langle\!\langle a,^i b\rangle\!\rangle}{\mathrm{E}(b)^i} \tag{10}$$

Substituting Equation 10 into Equation 6 gives the equation for the expected value of the ratio:

$$\mathrm{E}\left(\frac{a}{b}\middle|b\neq 0\right)=\frac{\mathrm{E}(a)}{\mathrm{H}(b)}+\sum_{i=1}^{\infty}(-1)^i\frac{\langle\!\langle a,^i b\rangle\!\rangle}{\mathrm{E}(b)^{i+1}} \tag{11}$$

An alternate representation

Equation 11 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 8 directly into the far righthand part of Equation 3. 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:

$$\mathrm{E}\left(\frac{a}{b}\middle|b\neq 0\right)=\frac{\mathrm{E}(a)}{\mathrm{E}(b)}+\sum_{i=1}^{\infty}(-1)^i\frac{\mathrm{E}(a)\langle\!\langle {}^i b\rangle\!\rangle+\langle\!\langle a,^i b\rangle\!\rangle}{\mathrm{E}(b)^{i+1}} \tag{12}$$

Series solution for $b \geqslant 2\mathrm{E}(b)$

As noted above, the Taylor series expansion of Equation 7, given in Equation 8, converges to the function only in the range $-\mathrm{E}(b) < b^* < \mathrm{E}(b)$ or, in the original variables, $0 < b < 2\mathrm{E}(b)$ (Figure 1). 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 1 shows how the Taylor series behaves at $b^* = \mathrm{E}(b)$; the series in Equation 8 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 & Phillips (1957) (pg. 533) showed that the following series of finite differences converges to f.

$$f(a+x)=\lim_{h\to 0}\sum_{i=0}^{\infty}\frac{x^i}{i!}\frac{\Delta_h^i f(a)}{h^i} \tag{13}$$

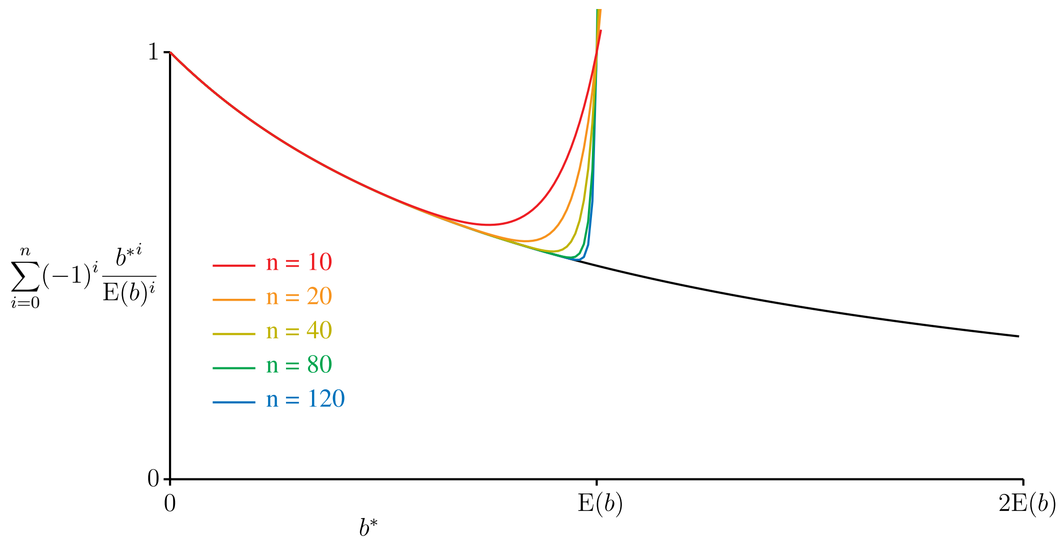

Figure 1: Result of adding increasing numbers of terms to the Taylor expansion.

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

$$\Delta_h^i f(a) = \sum_{j=0}^{i}(-1)^j \binom{i}{j} f(a + (i-j)\cdot h) \tag{14}$$

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

As $h \to 0$, each term in the series in Equation 13 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 13 converging to f even where the Taylor series does not.

Applying Equation 14 to Equation 7, evaluated at $b^* = 0$, yields:

$$\Delta_h^i f(0) = (-1)^i \frac{i! \cdot \mathrm{E}(b) \cdot h^i}{\prod_{j=0}^{i}(\mathrm{E}(b) + j \cdot h)} \tag{15}$$

We can now use Equation 13, with the finite difference operator from Equation 15, to get a series approximation for Equation 7:

$$\left(1 + \frac{b^*}{\mathrm{E}(b)}\right)^{-1} = \lim_{h\to 0}\sum_{i=0}^{\infty}(-1)^i \frac{(b^*)^i \mathrm{E}(b)}{\prod_{j=0}^{i}(\mathrm{E}(b) + j \cdot h)} \tag{16}$$

Note that the series in Equation 16 would be identical to the Taylor series, Equation 8, 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 2, which shows the behavior of the series in Equation 16 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.

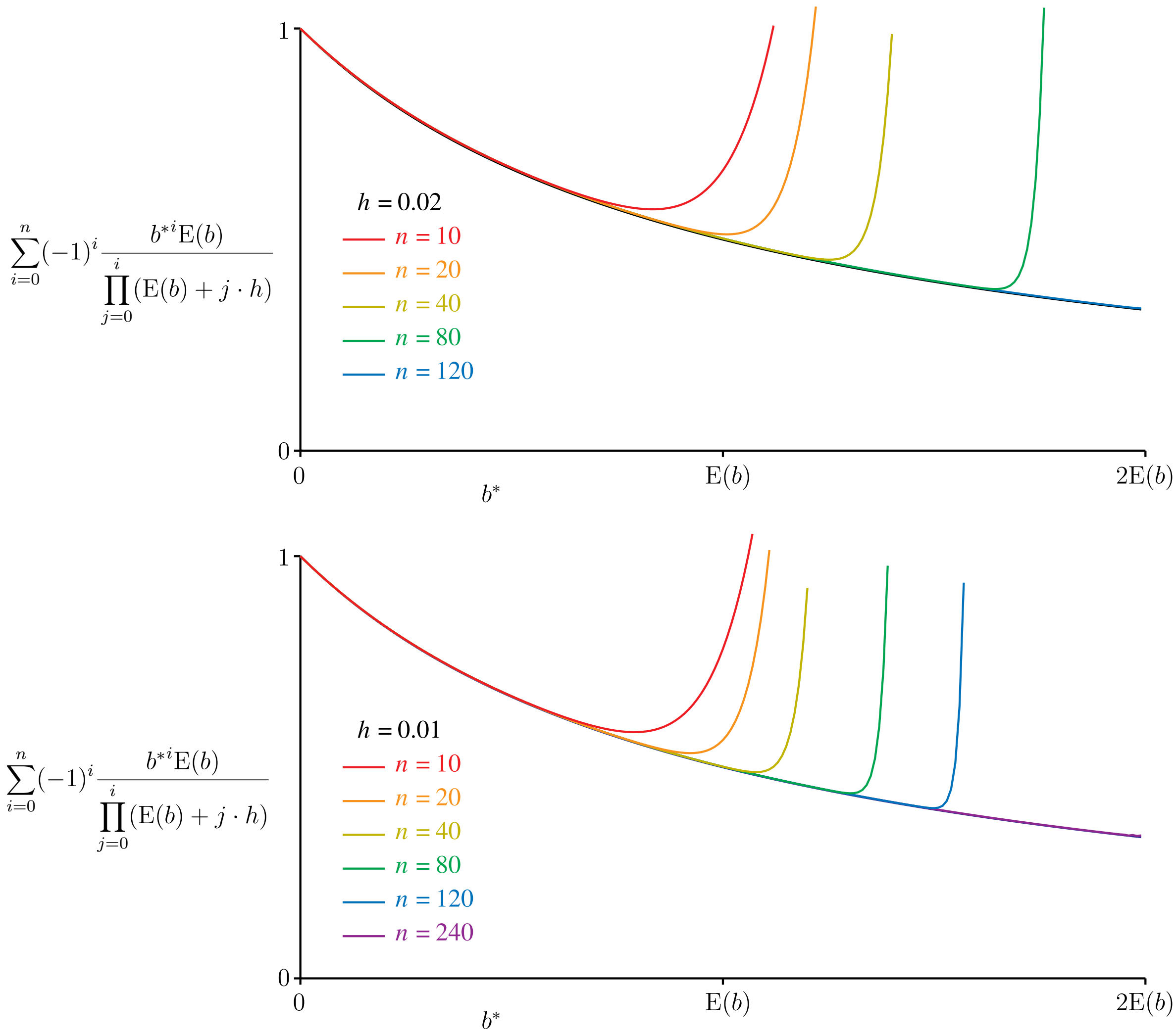

Figure 2: Result of adding increasing numbers of terms to the series in Equation 16, for two different values of h.

Substituting Equation 16 into Equation 4, we get the following two equations as generalizations of Equations 11 and 12:

$$\mathrm{E}\left(\frac{a}{b}\middle|b \neq 0\right) = \frac{\mathrm{E}(a)}{\mathrm{H}(b)} + \lim_{h\to 0}\sum_{i=1}^{\infty}(-1)^i \frac{\langle\!\langle a,^i b\rangle\!\rangle}{\prod_{j=0}^{i}(\mathrm{E}(b)+j\cdot h)} \tag{17}$$

and:

$$\mathrm{E}\left(\frac{a}{b}\middle|b \neq 0\right) = \frac{\mathrm{E}(a)}{\mathrm{E}(b)} + \lim_{h\to 0}\sum_{i=1}^{\infty}(-1)^i \frac{\mathrm{E}(a)\langle\!\langle {}^i b\rangle\!\rangle + \langle\!\langle a,^i b\rangle\!\rangle}{\prod_{j=0}^{i}(\mathrm{E}(b)+j\cdot h)} \tag{18}$$

For dealing with an arbitrary ratio of correlated random variables, the limit term, and the fact that we can not simply ignore the finite step size (h), makes these equations somewhat more awkward to work with than those based on the Taylor series. For computational work, Figure 2 shows that choosing a larger value of h reduces the number of terms that need to be considered (though too large a value will make the series a poor approximation).

For purposes of our evolutionary models, though, this is not so much of a problem. The reason is that, so long as different individuals have largely independent fitness distributions, each term in the series is divided by an increasing power of N (population size). For realistic population sizes and values of $\widehat{\overline{w}}$ (usually not close to zero for a population that is not about to go extinct), this insures that the final series will converge quickly. We can therefore still use the approximation:

$$\mathrm{E}\left(\frac{w}{\overline{w}}\middle|\overline{w} \neq 0\right) \approx \frac{\widehat{w}}{\mathrm{H}(\overline{w})} - \frac{\langle\!\langle {}^2 w\rangle\!\rangle}{N\widehat{\overline{w}}^2} + \frac{\langle\!\langle {}^3 w\rangle\!\rangle}{N^2\widehat{\overline{w}}^3} - O\left(\frac{1}{N^3}\right) \tag{19}$$

even when there is appreciable probability that $\overline{w}$ will be more than twice $\widehat{\overline{w}}$.

References

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

Rice, S. H. 2008: A stochastic version of the Price equation reveals the interplay of deterministic and stochastic processes in evolution. BMC Evolutionary Biology 8:262.

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