Appendix: algebra and calculus basics

September 28, 2005

1 Logarithms

Logarithms are the solutions to equations like $y = e^x$ or $y = 10^x$. *Natural* logs, ln or $\log_e$, are logarithms base e ($e = 2.718\ldots$); *common* logs, $\log_{10}$, are typically logarithms base 10. When you see just log it's usually in a context where the difference doesn't matter (although in R $\log_{10}$ is `log10` and $\log_e$ is `log`).

1. $\log(1) = 0$. If $x > 1$ then $\log(x) > 0$, and vice versa. $\log(0) = -\infty$ (more or less); logarithms are undefined for $x < 0$.
2. Logarithms convert products to sums: $\log(ab) = \log(a) + \log(b)$.

3. Logarithms convert powers to multiplication: $\log(a^n) = n \log(a)$.

4. You can't do anything with $\log(a+b)$.

5. Converting bases: $\log_x(a) = \log_y(a)/\log_y(x)$. In particular, $\log_{10}(a) = \log_e(a)/\log_e(10) \approx \log_e(a)/2.3$ and $\log_e(a) = \log_{10}(a)/\log_{10}(e) \approx \log_{10}(a)/0.434$. This means that converting between log bases just means multiplying or dividing by a constant. You can prove this relationship as follows:

$$\begin{aligned} y &= \log_{10}(x) \\ 10^y &= x \\ \log_e(10^y) &= \log_e(x) \\ y \log_e(10) &= \log_e(x) \\ y &= \log_e(x)/\log_e(10) \end{aligned}$$

   (compare the first and last lines).

6. The derivative of the logarithm, $d(\log x)/dx$, equals $1/x$. This is always positive for $x > 0$ (which are the only values for which the logarithm means anything anyway).

7. The fact that $d(\log x)/dx > 0$ means the function is *monotonic* (always either increasing or decreasing), which means that if $x > y$ then $\log(x) > \log(y)$ and if $x < y$ then $\log(x) < \log(y)$. This in turn means that if you find the maximum likelihood parameter, you've also found the maximum log-likelihood parameter, and *vice versa*.

2 Differential calculus

1. Notation: differentation of a function $f(x)$ with respect to x can be written, depending on the context, as $\frac{df}{dx}$; f'; $\dot{f}$; or f_x. I will stick to the first two notations, but you may encounter the others elsewhere.

2. Definition of the derivative:

$$\frac{df}{dx} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{(x+\Delta x) - x} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}. \quad (1)$$

   In words, the derivative is the slope of the line tangent to a curve at a point, or the "instantaneous" slope of a curve. The second derivative, d^2f/dx^2, is the rate of change of the slope, or the curvature.

3. The derivative of a constant (which is a flat line if you think about it as being a curve) is zero (zero slope).

4. The derivative of a line, $y = ax$, is the slope of the line, a.

5. Derivatives of polynomials: $\frac{d(x^n)}{dx} = nx^{n-1}$.

6. Derivatives of sums: $\frac{d(f+g)}{dx} = \frac{df}{dx} + \frac{dg}{dx}$ (and $d(\sum_i y_i)/dx = \sum_i (dy_i/dx)$).

7. Derivatives times constants: $\frac{d(cf)}{dx} = c\frac{df}{dx}$, if c is a constant ($\frac{dc}{dx} = 0$).

8. Derivative of the exponential: $\frac{d(\exp(ax))}{dx} = a\exp(ax)$, if a is a constant. (If not, use the chain rule.)

9. Derivative of logarithms: $\frac{d(\log(x))}{dx} = \frac{1}{x}$.

10. Chain rule: $\frac{d(f(g(x)))}{dx} = \frac{df}{dg} \cdot \frac{dg}{dx}$ (thinking about this as "multiplying fractions" is a good mnemonic but don't use that in general!) *Example:*

$$\frac{d(\exp(x^2))}{dx} = \frac{d(\exp(x^2))}{d(x^2)} \cdot \frac{dx^2}{dx} = \exp(x^2) \cdot 2x. \tag{2}$$

   Another example: people sometimes express the proportional change in x, $(dx/dt)/x$, as $d(\log(x))/dt$. Can you see why?

11. *Critical points* (maxima, minima, and saddle points) of a curve f have $df/dx = 0$. The sign of the second derivative determines the type of a critical point (positive = minimum, negative = maximum, zero = saddle).

3 Partial differentiation

1. Partial differentiation acts just like regular differentiation except that you hold all but one variable constant, and you use a curly d ∂ instead of a regular d. So, for example, $\partial(xy)/\partial(x) = y$. Geometrically, this is taking the slope of a surface in one particular direction. (Second partial derivatives are curvatures in a particular direction.)

2. You can do partial differentiation multiple times with respect to different variables: order doesn't matter, so $\frac{\partial(f)}{\partial(x)\partial(y)} = \frac{\partial(f)}{\partial(y)\partial(x)}$.

4 Integral calculus

For the material in this book, I'm not asking you to remember very much calculus, but it would be useful to remember that

1. the (definite) integral of $f(x)$ from a to b, $\int_a^b f(x)\,dx$, represents the area under the curve between a and b; the integral is a limit of the sum $\sum_{x_i=a}^{b} f(x_i)\Delta x$ as $\Delta x \to 0$.

2. You can take a constant out of an integral (or put one in): $\int af(x)\,dx = a\int f(x)\,dx$.

3. Integrals are additive: $\int (f(x) + g(x))\,dx = \int f(x)\,dx + \int g(x)\,dx$.

5 Factorials and the gamma function

A *factorial*, written as (say) $k!$, means $k \times k - 1 \times \ldots 1$. For example, $2! = 2$, $3! = 6$, and $6! = 720$ (in R a factorial is `factorial()` — you can't use the shorthand ! notation, especially since `!=` means "not equal to". Factorials come up in probability calculations all the time, e.g. as the number of permutations with k elements. The *gamma function*, usually written as Γ (`gamma()` in R) is a generalization of factorials. For integers, $\Gamma(x) = (x-1)!$. Factorials are only defined for integers, but for positive, non-integer x (e.g. 2.7), $\Gamma(x)$ is still defined and it is still true that $\Gamma(x+1) = x \cdot \Gamma(x)$.

Factorials and gamma functions get very large, and you often have to compute ratios of factorials or gamma functions (as in the binomial coefficient, $k!/(N!(N-k)!)$. Numerically, it is more efficient and accurate to compute the logarithms of the factorials first, add and subtract them, and then exponentiate the result: $\exp(\log k! - \log N! - \log(N-k)!)$. R provides the log-factorial (`lfactorial()`) and log-gamma (`lgamma()`) functions for this purpose.

About the only reason that the gamma function ever comes up in ecology is that it is the *normalizing constant* (see ch. 4) for the gamma *distribution*, which is usually denoted as Gamma (not Γ): $\mathrm{Gamma}(x,a,s) = \frac{1/(s^a\Gamma(a))}{x}^{(} a - 1)e^{-}(x/s)$.

6 Probability

1. Probability distributions always add or integrate to 1 over all possible values.

2. Probabilities of independent events are multiplied: $p(A \text{ and } B) = p(A)p(B)$.

3. The *binomial coefficient*,

$$\binom{N}{k} = \frac{N!}{k!(N-k)!}, \tag{3}$$

is the number of different ways of choosing k objects out of a set of N, without regard to order. ! denotes a factorial: $n! = n \times n-1 \times ... \times 2 \times 1$. (Proof: think about picking k objects out of N, without replacement but keeping track of order. The number of different ways to pick the first object is N. The number of different ways to pick the second object is $N-1$, the third $N-2$, and so forth, so the total number of choices is $N \times N-1 \times ...N-k+1 = N!/(N-k)!$. The number of possible orders for this set (permutations) is $k!$ by the same argument (k choices for the first element, $k-1$ for the next ...). Since we don't care about the order, we divide the number of ordered ways ($N!/(N-k)!$) by the number of possible orders ($k!$) to get the binomial coefficient.)

7 The delta method: formula and derivation

The formula for the delta method of approximating variances is:

$$\text{Var}(f(x,y)) \approx \left(\frac{\partial f}{\partial x}\right)^2 \text{Var}(x) + \left(\frac{\partial f}{\partial y}\right)^2 \text{Var}(y) + 2\left(\frac{\partial f}{\partial x}\frac{\partial f}{\partial y}\right) \text{Cov}(x,y) \tag{4}$$

Lyons [**?**] gives a very readable alternative description of the delta method; Oehlert [**?**] gives a short technical description of the formal assumptions necessary for the delta method to apply.

This formula is exact in a bunch of simple cases:

- Multiplying by a constant: $\text{Var}(ax) = a^2\text{Var}(x)$
- Sum or difference of independent variables: $\text{Var}(x \pm y) = \text{Var}(x) + \text{Var}(y)$
- Product or ratio of independent variables: $\text{Var}(x \cdot y) = y^2\text{Var}(x) + x^2\text{Var}(y) = x^2y^2\left(\frac{\text{Var}(x)}{x^2} + \frac{\text{Var}(y)}{y^2}\right)$: this also implies that $(\text{CV}(x \cdot y))^2 = (\text{CV}(x))^2 + (\text{CV}(y))^2$
- I believe (check!!) that the formula is exact if $P(x,y)$ is bivariate normal (and the function is not too weird??)

You can also extend the formula to more than two variables if you like.

Derivation: use the (multivariable) Taylor expansion of $f(x, y)$ including *linear terms only*:

$$f(x, y) \approx f(\bar{x}, \bar{y}) + \frac{\partial f}{\partial x}(x - \bar{x}) + \frac{\partial f}{\partial y}(y - \bar{y})$$

where the derivatives are evaluated at $(\bar{x}, \bar{y})$.

Substitute this in to the formula for the variance of $f(x, y)$:

$$\begin{aligned}
\mathrm{Var}(f(x,y)) &= \int P(x,y)(f(x,y) - f(\bar{x},\bar{y}))^2 \, dx \, dy && (5)\\
&= \int P(x,y) \left(f(\bar{x},\bar{y}) + \frac{\partial f}{\partial x}(x - \bar{x}) + \frac{\partial f}{\partial y}(y - \bar{y}) - f(\bar{x},\bar{y}) \right)^2 dx \, dy && (6)\\
&= \int P(x,y) \left(\frac{\partial f}{\partial x}(x - \bar{x}) + \frac{\partial f}{\partial y}(y - \bar{y}) \right)^2 dx \, dy && (7)\\
&= \int P(x,y) \left(\left(\frac{\partial f}{\partial x}\right)^2 (x - \bar{x})^2 + \left(\frac{\partial f}{\partial y}\right)^2 (y - \bar{y})^2 + 2\frac{\partial f}{\partial x}\frac{\partial f}{\partial y}(x - \bar{x})(y - \bar{y}) \right) dx \, dy && (8)\\
&= \int P(x,y) \left(\frac{\partial f}{\partial x}\right)^2 (x - \bar{x})^2 \, dx \, dy \\
&\quad + \int P(x,y) \left(\frac{\partial f}{\partial y}\right)^2 (y - \bar{y})^2 \, dx \, dy \\
&\quad + \int P(x,y) \, 2\frac{\partial f}{\partial x}\frac{\partial f}{\partial y}(x - \bar{x})(y - \bar{y}) \, dx \, dy && (9)\\
&= \left(\frac{\partial f}{\partial x}\right)^2 \int P(x,y)(x - \bar{x})^2 \, dx \, dy \\
&\quad + \left(\frac{\partial f}{\partial y}\right)^2 \int P(x,y)(y - \bar{y})^2 \, dx \, dy \\
&\quad + 2\frac{\partial f}{\partial x}\frac{\partial f}{\partial y} \int P(x,y)(x - \bar{x})(y - \bar{y}) \, dx \, dy && (10)\\
&= \left(\frac{\partial f}{\partial x}\right)^2 \mathrm{Var}(x) + \left(\frac{\partial f}{\partial y}\right)^2 \mathrm{Var}(y) + 2\frac{\partial f}{\partial x}\frac{\partial f}{\partial y}\mathrm{Cov}(x,y) && (11)
\end{aligned}$$