0.1 Moment matching

The concept of shape parameters is unfamiliar to ecologists trained in statistics classes emphasizing methods based on the normal distribution, which is to say, most ecologists. The two shape parameters of the normal distribution are its first and second central moments, the mean and the variance, motivating students and colleagues to ask us, "Why are shape parameters necessary? Why not simply use the moments as parameters for distributions?"

The answer is important, if not obvious. For all distributions except the normal, multivariate normal and the Poisson, the variance of the distribution is a function of the mean (Figure showing mean / variance relationships). Shape parameters for the binomial, multinomial, negative binomial, beta, gamma, lognormal, exponential and Dirichlet are functions of the mean *and* the variance, which allows the relationship between the mean and variance to change as the shape parameters change. The only conditions that allow the moments to be used as shape parameters occur mean and the variance are constant, as is the case for the normal, multivariate normal, and the Poisson.

This creates a problem for the ecologist who seeks to use the toolbox of distributions that we described above, a problem that can be easily seen in the following example. Assume you want to model the influence of growing season rainfall (x_i) on the mean aboveground net primary production in grasslands (μ_i). You might be inclined to reach for a simple linear model $\mu_i = \gamma_0+\gamma_1 x_i$ to represent this relationship (e.g. Knapp paper). However, there are structural problems with a linear model because it predicts values that can be negative, which makes no sense for production. Moreover, it predicts that growth increases infinitely with increasing rainfall, which clearly isn't right on biological grounds. So, using your knowledge of deterministic models (section __), you choose

$$\mu_i = \frac{\kappa \exp\left(\gamma_0 + \gamma_1 x_i\right)}{1+\exp\left(\gamma_0 + \gamma_1 x_i\right)}, \tag{1}$$

thereby deftly assuring that the model's estimate is non-negative and asymptotically approaches a maximum, κ.

Equation 1 is purely deterministic. You would like to represent the uncertainty that arises because the model isn't perfect and because you fail to observe net primary production perfectly.[1] Your first thought about modeling the uncertainty might be to use a normal distribution,

[1] In this case, these sources of uncertainty will be inseparable. Later, we will develop models that separate them.

normal $[y_i|\mu_i, \sigma]$. So, your model predicts the mean (μ_i) of the distribution of of observations of growth (y_i) and the uncertainty surrounding that prediction depends σ. This is the traditional framework for regression. It is convenient because the prediction of the model is the first argument to the distribution.[2].

However, informed by the section on continuous distributions, you decide that the normal is a poor choice for your model for two reasons. First, the support is wrong – production cannot be negative, so you need a distribution for data that are continuous and strictly positive. Moreover, a plot of the data shows that the spread of the residuals increases with increasing production, casting doubt on the assumption variance is constant. As an alternative, you choose the gamma distribution because it is [3]strictly non-negative and is parameterized such that the variance increases in proportion to μ^2.

This is entirely sensible, but now you have a problem. How do you get the prediction of your model, the mean prediction of production at a given level of rainfall (μ_i), into the gamma probability density function if the function doesn't contain an argument for the mean? How do you represent the uncertainty σ? The solution to this problem is *moment matching.* You need an equation for the shape parameters in terms of the moments to allow you to use the gamma distribution to represent the uncertainty in your model. Equations for moments as functions of the shape parameters can be found in any mathematical statistics text. The converse is not true; it is uncommon to see the shape parameters expressed as functions of the moments. However, obtaining these functions is easy and useful. We simply solve two equations in two unknowns. Illustrating this solution using the gamma distribution with shape parameters shape $= \alpha$ and rate $= \beta$:

$$\mu = \frac{\alpha}{\beta} \tag{2}$$

$$\sigma^2 = \frac{\alpha}{\beta^2}, \tag{3}$$

[2]You are probably more familiar with the equivalent formulation, $y_i = \gamma_0 + \gamma_1 x_i + \epsilon_i, \; \epsilon_i \sim \text{normal}\left(0, \sigma^2\right)$. We avoid this additive arrangement for representing stochasticity because it cannot be applied to distributions that cannon be centered on zero.

[3]The lognormal would be another logical choice and it is unlikely that it would make any difference which of these two distributions you chose.

so,

$$\alpha = \frac{\mu^2}{\sigma^2} \tag{4}$$

$$\beta = \frac{\mu}{\sigma^2}. \tag{5}$$

You are now equipped to use the gamma distribution to represent the uncertainty in your model of net primary production,

$$\mu_i = \frac{\kappa \exp(\gamma_0 + \gamma_1 x_i)}{1 + \exp(\gamma_0 + \gamma_1 x_i)}$$

$$\alpha_i = \frac{\mu_i^2}{\sigma^2} \tag{6}$$

$$\beta_i = \frac{\mu_i}{\sigma^2} \tag{7}$$

$$[y_i|\mu_i, \sigma] = \text{gamma}\left(y_i \,\middle|\, \alpha_i, \beta_i\right). \tag{8}$$

As a second example, imagine that you wanted to model the probability of survival of juvenile birds μ_i as a function of population density x_i. Removing the κ term from equation 1 does the trick for your deterministic model, which now makes predictions strictly between zero and 1, that is $\mu_i = \frac{\exp(\gamma_0+\gamma_1 x_i)}{1+\exp(\gamma_0+\gamma_1 x_i)}$. But the gamma distribution is no longer appropriate for representing the uncertainty because it applies to random variables that can exceed one. The normal is even worse because it includes negative values *and* values that exceed 1. A far better choice is the beta, which models continuous random variables with support over $(0, 1)$. Solving for the shape parameters in terms of the moments (using equations **??** and **??**), we can now make a prediction of survival with our deterministic model and properly represent uncertainty using the beta distribution,

$$\alpha_i = \frac{\mu_i^2 - \mu_i^3 + \mu_i \sigma^2}{\sigma_i} \tag{9}$$

$$\beta_i = \frac{\mu_i - 2\mu_i^2 + \mu_i^3 - \sigma^2 + \mu_i \sigma^2}{\sigma^2} \tag{10}$$

$$[y_i|\mu_i, \sigma] = \text{beta}(y_i|\alpha_i, \beta_i) \tag{11}$$

Equations 4,5 and 9, 10 are examples of moment matching. We use the functional relationship between the shape parameters and the moments to allow us to match the predictions of a model to the arguments of the distribution that is best suited to the model and the data. It is important to see how moment matching allows us to specify characteristics of distributions for which the variance is a function of the mean. These matching relationships are broadly useful for the ecological modeler because they allow us to use all of the distributions we have described above to represent the stochasticity regardless of the form of the arguments to those distributions. It is easy enough to derive the moment matching relationships yourself, but we saved you the trouble in the distribution cheat sheet.