How to Fit Multiple Geometric Tails

November 11, 2005

The exponential PDF comes in the form

$$\Pr(x) = \frac{1}{\lambda}e^{-\frac{x}{\lambda}} \tag{1}$$

where λ is the expected value of the PDF and s is used in our case to scale the area under the curve.

To fit a geometric tail from A to B with density N_i (i representing the current piece of the distribution) and start probability $\Pr_{i-1}(A)$, we can view this as a system of equations.

Our constraints are that

$$\begin{aligned} \Pr_{i-1}(A) &= \Pr_i(A) & (2) \\ \int_A^B \Pr_i(x)dx &= N_i & (3) \end{aligned}$$

Each, respectively, imply

$$\begin{aligned} \Pr_{i-1}(A) &= \frac{s_i}{\lambda_i}e^{-\frac{A}{\lambda_i}} & (4) \\ \int_A^B \frac{s_i}{\lambda_i}e^{-x}\lambda_i dx &= N_i & (5) \end{aligned}$$

The variable s_i is a constant factor different from N_i which scales the exponential such that the area under the curve between A and B is N_i, and not the area of the whole curve from 0 to infinity.

We know $\Pr_{i-1}(A)$ and N_i, and we hope to find solutions for λ_i and s_i in terms of these, so we can each piece one at a time without any fancy global inference.

Integrating (5), we get

$$s_i(e^{-\frac{A}{\lambda_i}} - e^{-\frac{B}{\lambda_i}}) = N_i \tag{6}$$

From (4), we can get s_i alone.

$$s_i = \lambda_i \Pr_{i-1}(A)e^{\frac{A}{\lambda_i}} \tag{7}$$

Now we can substitute this back into (6).

$$\lambda_i \Pr_{i-1}(A)(1 - e^{\frac{A-B}{\lambda_i}}) = N_i \tag{8}$$

This must be heuristically solved, as it is a transcendental equation. We put each instance of λ_i on either side of the equation:

$$\Pr_{i-1}(A)(1 - e^{\frac{A-B}{\lambda_i}}) = \frac{N_i}{\lambda_i} \tag{9}$$

Then set λ_i for a number of reasonable (say, 1 to 1,000,000) values on each side of the equation, and pick the one pair that is closest. We can s_i easily find s_i once λ_i is found with equation (7).

In the simple case, where there is only one tail, the parameters can be found in closed-form: Your constraints are as follows:

$$N_n = \int_A^\infty f_n(x)dx \tag{10}$$

$$f_{n-1}(A) = f_n(A) \tag{11}$$

Integrating (10) gives us

$$N_n = s_n e^{-\frac{A}{\lambda_n}} \tag{12}$$

Expanding (11) gives us

$$f_{n-1}(A) = \frac{s_n}{\lambda_n} e^{-\frac{A}{\lambda_n}} \tag{13}$$

Dividing equation (13) by equation(12) gives us

$$\frac{N_n}{f_{n-1}(A)} = \lambda_n \tag{14}$$

which we can use to solve for s_n in the equation (13).