Log transformed logseries

Luís Borda-de-Água

8 February 2017

The formula for the logseries is

$$S = \alpha \frac{x^n}{n}.$$

Since we are interested in a mass function we will work, instead, with

$$f(n) = \frac{\alpha}{S} \frac{x^n}{n}.$$

α and x are related, so there is only one parameter.

Often, we plot the log (base 2) transformed values, in the following format

$$y = \lfloor \log_2(n) \rfloor.$$

The question is: Which is the formula for the distribution of y?

Because we are dealing with discrete variables, $f(y)$ is

$$f(y) = \frac{\alpha}{S} \sum_{j=2^y}^{2^{y+1}-1} \frac{x^j}{j}.$$

The sum can be simplified. Consider $k = 2^y$ and $p = 2^{y+1} - 1$, hence

$$f(y) = \frac{\alpha}{S} \sum_{j=k}^{p} \frac{x^j}{j}$$

The result is

$$f(y) = \frac{\alpha}{S}(x^k \Phi(x, 1, k) - x^{p+1} \Phi(x, 1, p+1)),$$

where Φ is the *Lerch transcendent function*, defined as

$$\Phi(z, s, a) = \sum_{k=0}^{\infty} \frac{z^k}{(a+k)^s}.$$

In R the Lerch transcendent function is calculated in the package VGAM:

```
lerch(z,s,a,tolerance=1.0e-10,iter=100)
```

However, I don't think we need to use this function. In general, it should be easy to just calculate

$$f(y) = \frac{\alpha}{S} \sum_{j=k}^{p} \frac{x^j}{j}.$$

Equally, we can easily calculate, and then plot, the *raw moments*

$$M_n(y) = \sum_{y=0}^{N} y^n \frac{\alpha}{S} \sum_{j=k}^{p} \frac{x^j}{j}$$

and recall that $k = 2^y$ and $p = 2^{y+1} - 1$.

We know the mode of the logarithmic distribution is for 1. What is the mode for $f(y)$?

To calculate the mode, calculate $f(y+1)/f(y)$ and see whether it is < 1:

or

$$\frac{f(y+1)}{f(y)} = \frac{\sum_{j=2^{y+1}}^{2^{y+2}-1} \frac{x^j}{j}}{\sum_{j=2^y}^{2^{y+1}-1} \frac{x^j}{j}} < 1.$$

For instance if $y = 2$, is it true that

$$\frac{x^8}{8} + \frac{x^9}{9} + \frac{x^{10}}{10} + \frac{x^{11}}{11} + \frac{x^{12}}{12} + \frac{x^{13}}{13} + \frac{x^{14}}{14} + \frac{x^{15}}{15} < \frac{x^4}{4} + \frac{x^5}{5} + \frac{x^6}{6} + \frac{x^7}{7}.$$

To show it is, separate the right hand-side into two terms

$$(\frac{x^8}{8} + \frac{x^9}{9} + \frac{x^{10}}{10} + \frac{x^{11}}{11}) + (\frac{x^{12}}{12} + \frac{x^{13}}{13} + \frac{x^{14}}{14} + \frac{x^{15}}{15}) < (\frac{x^4}{4} + \frac{x^5}{5} + \frac{x^6}{6} + \frac{x^7}{7}).$$

We could now re-order the sum, and ask for each term

$$\frac{x^{2n+k}}{2n+k} + \frac{x^{3n+k}}{3n+k} < \frac{x^{n+k}}{n+k}$$

with $0 \leq k \leq n-1$ and $x < 1$. For instance,

$$\frac{x^8}{8} + \frac{x^{12}}{12} < \frac{x^4}{4}$$

Consider the two extreme cases $k = 0$ and $k = n-1$. When $k = 0$

$$\frac{x^{2n}}{2n} + \frac{x^{3n}}{3n} < \frac{x^n}{n}$$

or

$$\frac{x^n}{2} + \frac{x^{2n}}{3} < 1$$

and this is always true. Notice that the largest value occurs for $x = 1$ and we have $1/2 + 1/3 < 1$.