Scaling in cell number variability.

Paul Grayson

December 14, 2001

1 Introduction

The species-area and mass-metabolism relationships were discussed in lecture. Both relationships are simple log-linear relationships that extend over many orders of magnitude, suggesting that they are caused by very simple mechanisms. Researchers have suggested many possible mechanisms, but none are very convincing.

A new scaling "law" can now be added to these two: a power-law relationship identified in a recent study [1]. In this study, R. Azevedo *et al* show that variations in the number of cells in an organ scale with the size of that organ over twelve orders of magnitude. The researchers propose a simple theoretical model of cell number variability that closely predicts the observed law.

In Section 2, I will explain how the researchers analyzed data on cell number variability. Section 3 presents the theory and compares it to models seen in lecture, and Section 4 discusses implications and remaining questions.

2 Analysis of experimental data

The researchers discovered a new scaling law by analyzing and combining the results of many studies. They examined 138 experiments that resulted in 2,177 characterizations of cells in particular organs. These estimates were sorted by species and organ type, then compared in various ways. The number n of cells in each organ is given by a particular probability distribution, for which the researchers obtained two numerical estimates: $M \equiv \langle n \rangle$, the

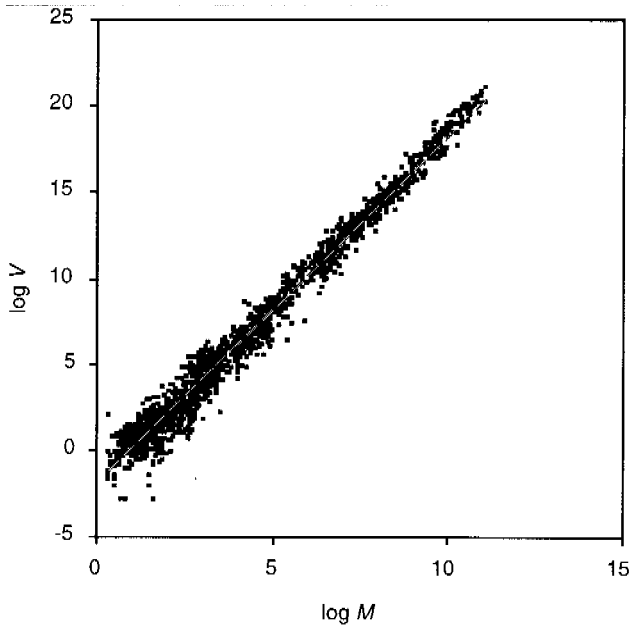

Figure 1: The power law relationship between mean cell number (M) and variance (V). From [1].

mean number of cells in the organ; and $V \equiv \langle (n - M)^2 \rangle$, the variance of the cell number.

A plot of the relationship between M and V (Figure 1) shows a clear power-law that covers almost 12 orders of magnitude, extending from slime mold and nematode organs containing around 10 cells to the major organs of large mammals:

$$V \propto M^{2.03} \,. \tag{1}$$

To physicists, an exponent of two is immediately surprising. We are used to Poisson processes, like radioactive decay, that produce a linear relationship:

$$V \propto M \,. \tag{2}$$

This study shows that cells behave differently — their relative variability

$$\frac{\sigma}{M} = \frac{\sqrt{V}}{M} \tag{3}$$

is approximately constant, and it was found to be especially constant when restricted to a single organ type across several species. In the next section

we will see how a simple theoretical model of cell division can explain this constant relative variability.

3 Theoretical model

The researchers presented a simple and reasonable theoretical model that is capable of predicting the relationship between M and V. I will present a simplified version of it here, with a nearly complete explanation of how it predicts the observed power law.

We will model organ development from a single initial cell as a random process consisting of successive generations. During a generation, each cell has the opportunity to divide, and we assume it does so with probability q. Let the probability of having n cells at generation i be given by $P_i(n)$; then we can find the probability distribution in generation $i+1$ by extending the tree on generation backward. If the initial cell does not divide in its first generation (probability $1-q$), then the probability at generation $i+1$ will just be $P_i(n)$. If it does divide (probability q), we are left with two cells that proceed through the next n generations independently — the probability distribution in this case will be the convolution of $P_i(n)$ with itself. Adding these two cases gives us the probability distribution at generation $i+1$:

$$P_{i+1}(n) = (1-q) \cdot P_i(n) + q \cdot \sum_{m=0}^{n-1} P_i(m) P_i(n-m)\,. \tag{4}$$

For large values of i, the probability distribution can be replaced by a continuous function:

$$p_{i+1}(n) = (1-q) \cdot p_i(n) + q \cdot \int_0^\infty p_i(m) p_i(n-m)\, dm\,. \tag{5}$$

We may attempt to use the Renormalization Group methods discussed in class to find fixed points p of this relation satisfying

$$\lambda\, p(\lambda\, n) = (1-q) \cdot p(n) + q \cdot \int_0^\infty p(m) p(n-m)\, dm\,. \tag{6}$$

The Fourier transform $\tilde{p}(k)$ of the probability distribution satisfies the simpler equation

$$\tilde{p}(k/\lambda) = (1-q)\, \tilde{p}(k) + q\, \tilde{p}(k)^2\,; \tag{7}$$

this is similar to but more complicated than the relation we derived for the Central Limit Theorem. I do not know how to solve it, but I have simulated the distribution using small values of i with $q \approx 1$, and it seems to approach a interesting limiting shape like that shown in Figure 2. This distribution is self similar: it has a small peak that is a near-duplicate of the entire graph, caused by the $(1 - q)$ term in the recursion relation.

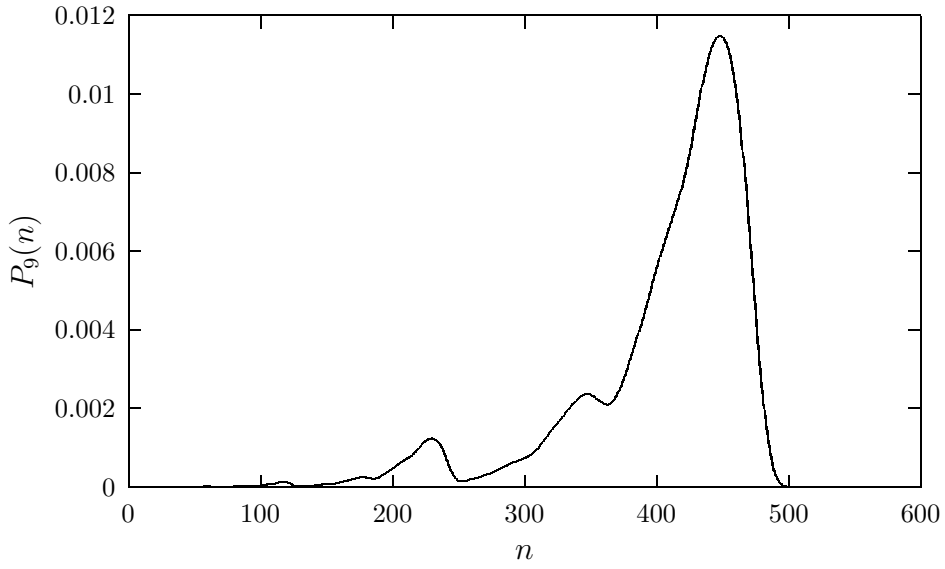

Figure 2: The cell number distribution for an organ with $q = 95\%$ after nine generations.

If we assume that the probability distribution does approach a fixed point as i increases, then it follows that the standard deviation will scale linearly with the mean. Then we find from Equation 3 that the relative variability of n stays constant as the number of generations increases. A simple model of independent cells duplicating randomly therefore predicts the observed scaling. Using this model, the authors found that the duplication probabilities in observed cells correspond to an error rate $(1 - q)$ of approximately two percent[1].

[1]Since they also include probabilities that the cell will die or duplicate more than once in a generation, their result does not directly give a value of q.

4 Discussion

The theory presented in Section 3 predicts the data very well, but does not match reality in several ways. As with any biological theory, there are details that a simple model does not capture. The authors admit that the assumptions of independent cells and fixed generation number are both violated in reality — but it is not clear that incorporating these effects would significantly change the prediction. In particular, cells in an organ are unlike animals in an environmental region: they are not confined to a bounded volume, and they do not move around. Instead, they duplicate in place, causing the volume of the developing organ to expand proportionally to itself. This indicates that space is not a particularly important issue, and that the cells can be treated independently.

The theory, in addition, predicts some details that are not observed in real organisms. As can be seen in Figure 2, there is a probability of $1 - q$ that a cell will not divide in its first generation, causing the organ to have half the number of expected cells. The authors quote a cell division accuracy rate of 98% as exceptionally high, but it predicts that one out of fifty mature organs will be half as small as average. But we know, for example, that very few people are born with scaled-down hearts or brains. This suggests that the first few generations must be controlled very precisely. The source of this precise control has probably already been identified by biologists — it would be interesting to see how it affects the scaling law.

References

[1] Ricardo B. R. Azevedo and Armand M. Leroi. A power law for cells. *PNAS*, 98(10):5699–5704, May 2001.