Sums, integrals and probability distributions.

Matthew Stapleton

October 19, 2005

Consider a discrete probability distribution $P(s,p)$. For instance, if we were talking about percolation it could be the probability that a site belongs to a cluster of size s, i.e.

$$P(s,p) = sn(s,p). \tag{1}$$

Now imagine we wish to calculate the first moment of this distribution (the average cluster size),

$$\chi(p) \propto \sum_{s=1}^{\infty} sP(s,p). \tag{2}$$

Unfortunately, except for some special cases, we do not know the exact form of $P(s,p)$. However, often our aim is to calculate how $\chi(p)$ behaves (scales) as a function of a quantity we call s_ξ, which is a function of p and diverges when $p \to p_c$. In this case we will have the scaling ansatz,

$$n(s,p) \sim s^{-\tau}\mathcal{G}\left(s/s_\xi\right) \qquad \text{for } s \ll 1,\ p \to p_c \tag{3}$$

and we might study $\chi(p)$ by approximating the sum in Eq. (2) by an integral and looking at how it behaves as a function of s_ξ. The function $\mathcal{G}(s/s_\xi)$ is an example of a scaling function and its main role is to cut off the distribution when $s \gtrsim s_\xi$. This means that the sum in Eq. (2) can be thought of as from $s = 1$ to s_ξ because all higher terms are killed by $\mathcal{G}(s/s_\xi)$.

Approximating a sum as an integral is then quite simple, yet the level of approximation it entails is not always obvious. Consider a sum of the form

$$\text{Sum} = \sum_{s=1}^{s_\xi} f(s) \tag{4}$$

where s_ξ represents the upper limit of this sum, which is perhaps imposed by the kind of scaling function discussed above. Now, we may interpret this as an approximation to an integral using the trapezium rule,

$$I = \int_1^{s_\xi} ds\, f(s) \approx \sum_{s=1}^{\frac{s_\xi}{\delta s}} \delta s f(s\delta s) \tag{5}$$

and by setting $\delta s = 1$, we have

$$\left.\sum_{s=1}^{\frac{s_\xi}{\delta s}} \delta s f(s\delta s)\right|_{\delta s=1} = \sum_{s=1}^{s_\xi} f(s). \tag{6}$$

Hence, it follows that $I \approx$ Sum, where the quality of the approximation depends on the behaviour of $f(s)$. Note that if we were trying to approximate the integral by the sum then we would expect using $\delta s = 1$ gives a rather poor approximation. However, if the main contribution to the sum comes from the large s terms (where $\delta s \ll s$), the approximation is only slight and our hope is that these approximations only alter multiplicative factors and corrections to scaling, leaving the leading order scaling behaviour unchanged.

To illustrate the level of approximation this step has entailed, we calculate

$$\text{Sum} = \sum_{s=1}^{s_\xi} s^\alpha \tag{7}$$

for some values of α and compare them to their integral equivalents,

$$I = \int_1^{s_\xi} ds\ s^\alpha. \tag{8}$$

Exercise: Plot $y = \sum_{s=1}^{x} s^\alpha$ and $y = \int_1^x s^\alpha$ as a function of x for $\alpha = -1/2$, $\alpha = -1$ and $\alpha = -2$. For what values of α does y scale with x? For each of the three cases, how does the result for the sum compare with the integral? Can you offer any kind of explanation? (Hint: You should be able to do the integrals yourself, but the sums are best done on a computer using a program such as Matlab. To determine if y scales, try plotting, e.g. $\log(y)$ vs. $\log(x)$. Compare the behaviour of the sums with the integrals.)

When $\alpha = 1$ we have Sum $= 1/2\ s_\xi(s_\xi + 1)$ and $I = 1/2\ s_\xi^2 - 1/2$. Evidently, this approximation gets better with increasing s_ξ since the relative error, $\epsilon \equiv |I - S|/S \sim 1/s_\xi$. Similarly, when $\alpha = -1$ we have $\lim_{s_\xi \to \infty} S = \left(\lim_{s_\xi \to \infty} \ln s_\xi\right) + \gamma$, where $\gamma \approx 0.577$ is Euler's constant, and $I = \ln s_\xi$. The approximation again improves as $s_\xi \to \infty$. However, as α becomes more negative the approximation gets worse. For instance,

$$\left.\begin{array}{rcl} \sum_{s=1}^{\infty} \frac{1}{s^2} & = & \frac{\pi^2}{6} \approx 1.64 \\ \int_1^\infty ds\ \frac{1}{s^2} & = & 1 \end{array}\right\} \Rightarrow \epsilon \approx 0.39 \tag{9}$$

is the best it will ever get for $\alpha = -2$, and

$$\left.\begin{array}{rcl} \sum_{s=1}^{\infty} \frac{1}{s^4} & = & \frac{\pi^4}{90} \approx 1.0823 \\ \int_1^\infty ds\ \frac{1}{s^4} & = & \frac{1}{3} \end{array}\right\} \Rightarrow \epsilon \approx 0.69 \tag{10}$$

is the best it will ever get for $\alpha = -4$.

This trend may be easily understood by considering Fig. 1. It is clear that as the exponent α becomes more negative the integral misses out more and more of the sum.

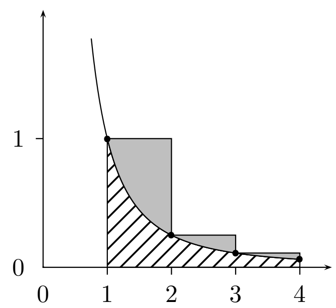

Figure 1: Plot of $1/s^2$ showing the difference between the area calculated from the integral and that calculated in the sum. The hashed area is counted by both, but the grey areas are missed by the integral.