LOVEGROVE MATHEMATICALS

Ranked distributions on finite domains

RESEARCH REPORT 2013-02

Roger Lovegrove

Ranked distributions are some of the most important distributions encountered, yet very little is known about them as a whole because they have always been analysed by using closed, parametric forms which allow analysis of only a particular types. In this report, a set-theoretic approach is taken which allows the set of all *all* ranked distributions to be investigated.

This is the first reprint of 2010-20, which is now withdrawn.

The recommended citation is

Lovegrove R., (2013), Ranked distributions on finite domains, Lovegrove Mathematicals Research Report 2013-02. December 2013.

LONDON
UNITED KINGDOM
November 2013

www.lovegrovemaths.co.uk roger@lovegrovemaths.co.uk

Contents

1 Introduction

Ranked distributions have long been a source of interest because of their mixture of simplicity and complexity. Because they cover such a disparate range of data, the emphasis by various authors has historically been on particular types of ranked distribution which can be represented parametrically in various ways.

Zipf (1949) famously saw in the distribution of the commoner words in the English language the pattern called Zipf's Law, now often called the Zipf/Pareto Law in recognition of Pareto's earlier work, Pareto (1896), on economics. Here, the governing equation is $Z(i) = K.(1/i)$, for some K the rôle of which is to normalise the distribution by bringing the sum of the terms to 1; on the finite set $i \in \{1, \ldots, N\}$ this becomes

$$Z_N(i) = \frac{1}{\Sigma_{j=1}^{N}(1/j)}(\frac{1}{i}). \tag{1}$$

Mandelbrot (1982) extended this (with different notation) to

$$Z_N(i) = \frac{(i+V)^{-1/D}}{\Sigma_{j=1}^{N}(j+V)^{-1/D}}$$

for some D,V.

Such equations cannot apply to all ranked distributions, and so describe only a subset which will have special properties not held by all ranked distributions.

Rather than seeking to represent ranked distributions by closed, parametric equations, a set-oriented approach will be taken by considering the set of all ranked distributions (on a finite domain).

2 Notation and Terminology

Note: Because of problems with embedded fonts for the larger mathematical symbols

- The greek upper-case letter Σ is used instead of the summation sign
- The greek upper-case letter Π us used instead of the product sign
- Superscripts and subscripts are used instead of limits
- normal, line-height (and) have been used as brackets, wherever possible, instead of larger brackets.
- the Daniell integral is represented by the summation (Σ) sign.

There are, however, still problems with some symbols, such as large brackets around matrices.

Notation and terminology follow Lovegrove [2].

Let $\mathbb{R}^+$ be the non-negative reals, and $\mathbb{N}^+$ be the non-negative integers. For $N \in \mathbb{N}$ let $X_N = \{1, \ldots, N\}$. N is called the *degree*.

Let $G(N) = \{g | g : X_N \to \mathbb{N}^+\}$, $H(N) = \{h | h : X_N \to \mathbb{R}^+\}$, so $G(N) \subset H(N)$. The elements of H(N) are called *histograms* on X_N and those of G(N) *integer-valued histograms*, shortened to *integrams*, on X_N. The histogram h is identified with the point $(h(1), \ldots, h(N))$.

We define the integram $\underline{0}_N$ by $\underline{0}_N : X_N \to \mathbb{N}^+ : i \mapsto 0$. Provided no ambiguity results, we shorten this to $\underline{0}$.

If we roll a die and throw the number 2 then we have not only thrown a 2 once but have also thrown 1, 3, 4, 5 and 6 zero times each. So we can think of ourselves as having thrown the integram (0,1,0,0,0,0). Also, we have not actually thrown the number 2 but have, rather, thrown the face labelled "2". It will be very convenient to adopt notation which associates the symbol "2" with (0,1,0,0,0,0).

We define $''i''_N$ to be that integram $(x_1, \ldots, x_N)$ for which $x_i = 1$ but $x_n = 0$ otherwise; for example, $''2''_6 = (0, 1, 0, 0, 0, 0)$. It is usually possible to write $''i''$ rather than $''i''_N$ without introducing ambiguity.

Let $f : X_N \to]0, 1]$ be such that $\Sigma_{i=1}^N f(i) = 1$. Then f is called a *distribution on* X_N .

The set of all distributions on X_N is denoted by $S(N)$.

For $f \in S(N)$ and $h \in H(N)$, by f^h is meant $f(1)^{h(1)} f(2)^{h(2)} \ldots f(N)^{h(N)}$. We have $f^{''i''_N} = f(i)$.

The *Set of Ranked distributions of degree N* is
$R(N) = \{r \in S(N) |\ r(1) > r(2) > \cdots > r(N)\}$.

For $h \in H(N)$, the *sample size* of h is $\omega(h) = \Sigma_{i=1}^N h(i)$.

For $h \in H(N), i \in X_N$, the *Relative Frequency of i given h* is $RF(i|h)$ where

$$(h = \underline{0}) \quad RF(i|h) = \frac{1}{N};$$

$$(h \neq \underline{0}) \quad RF(i|h) = \frac{h(i)}{\omega(h)}.$$

and the *Ranked Relative Frequency of i given h* is

$$RRF(i|h) = \Sigma_{j=i}^N \frac{1}{j}.RF(j|h) \quad .$$

For $g \in G(N)$, the *Multinomial coefficient associated with g* is

$$M(g) = \frac{\omega(g)!}{\Pi_{i=1}^N g(i)!}.$$

We have $M(''i'') = 1$.

For g∈G(N), h∈ H(N) and $P \subset S(N)$ where $P \neq \emptyset$, we define

$$L_P(g|h) = M(g)\frac{\int_{f\in P} f^g f^h}{\int_{f\in P} f^h}$$

where $\int$ is the Daniell integral.

$L_P(g|h)$ is called the *likeliness, over P, of g given h.* Since P, g or h will usually be clear from the context, this terminology can normally be shortened by omitting appropriate terms.

h is called the *given histogram*, g the *required integram* and P the *underlying set.* More generally, any non-empty subset of S(N) is called an *underlying set* in S(N).

We have $L_P(\underline{0}|h) = 1$ for all (h,P). $L_P(g|\underline{0})$ is written as $L_P(g)$.

If P is a singleton set, $P = \{f\}$, then $L_P(g|h) = M(g)f^g$, which is denoted by $Pr(g|f,h)$: since this is independent of h the notation may be simplified to $Pr(g|f)$; however, the presence of the h, although technically unnecessary, can sometimes add clarity.

Let $V \subset S(N)$. Then the *likeliness of V, over P and given h*, is

$$L_P(V|h) = \frac{\int_{V\cap P} f^h}{\int_P f^h}.$$

For $x \in [0,1]$ let $V_x = \{f \in S(N)|Pr(g|f) < x\}$. Then $L_P(V_x|h)$ is the likeliness, over P and given h, of the set of those $f \in P$ for which $Pr(g|f) < x$. We denote this by $L_P(Pr(g|f) < x|h)$. That is, in the text string '$L_P(V|h)$' we replace the symbol 'V' by the definition of the set V.

The function $[0,1] \to [0,1] : x \mapsto L_P(Pr(g|f) \leq x|h)$ is the *expected CDF of* $Pr(g|f)$.

Likewise, if $0 \leq x_0 \leq x_1 \leq 1$ then we define $L_P(Pr(g|f) \in [x_0, x_1]|h)$ to be $L_P(V|h)$ where $V = \{f \in P|Pr(g|f) \in [x_0, x_1]\}$. By covering [0,1] by cells in this way, we obtain an *expected frequency distribution for* $Pr(g|f)$.

3 Linear bijection $S(N) \to R(N)$

3.1 The mapping ϕ_N

Let A_N be the square matrix

$$A_N = \begin{bmatrix} {}^1\!/_1 & {}^1\!/_2 & \dots & \dots & {}^1\!/_N \\ & {}^1\!/_2 & \dots & \dots & {}^1\!/_N \\ & & \dots & \dots & {}^1\!/_N \\ & & & \dots & {}^1\!/_N \\ & & & & {}^1\!/_N \end{bmatrix}.$$

Being triangular with non-zero entries on its leading diagonal, A_N is invertible; its inverse is

$$A_N^{-1} = \begin{bmatrix} 1 & -1 & & & \\ & 2 & -2 & & \\ & & \dots & \dots & \\ & & & (N-1) & -(N-1) \\ & & & & N \end{bmatrix}.$$

Theorem 1. *Let $\phi_N : S(N) \to R(N) : f \mapsto r$ where $r^T = A_N f^T$. Then ϕ_N is a linear bijection.*

Proof. The codomain of ϕ_N is well-chosen (that is, $r \in R(N)$) since

i. A_N is left stochastic, so $r \in S(N)$;

ii. $(\forall i \in X_{N-1})\ r(i) = \dfrac{1}{i} f(i) + r(i+1) > r(i+1)$

So $r \in R(N)$.

ϕ_N is clearly linear. It is invertible, and therefore bijective onto its image $Im[\phi_N]$, because A_N is invertible.

It remains to show that $Im[\phi_N] = R(N)$, that is

(a) $Im[\phi_N] \subset R(N)$ *and*

(b) $R(N) \subset Im[\phi_N]$.

Since the codomain is well-chosen, we already have (a) so all that is left to show is (b), that is that $(\forall r \in R(N))(\exists f \in S(N))\ r = \phi_N(f)$.

Let $r \in R(N)$ and then let $f : X_N \to \mathbb{R}$ be given by $f^T = A_N^{-1} r^T$. Then $A_N f^T = r^T$, so that $r = \phi_N(f)$ iff f is in the domain of ϕ_N, that is iff $f \in S(N)$. We need to show that:-

(A.) $(\forall i \in X_N) f(i) > 0$

(B.) $\Sigma_{i=1}^N f(i) = 1$.

We have

$$\begin{bmatrix} f(1) \\ f(2) \\ ... \\ ... \\ f(N) \end{bmatrix} = \begin{bmatrix} 1 & -1 & & & & \\ & 2 & -2 & & & \\ & & \cdots & \cdots & & \\ & & & (N-1) & -(N-1) & \\ & & & & N & \end{bmatrix} \begin{bmatrix} r(1) \\ r(2) \\ ... \\ ... \\ r(N) \end{bmatrix}.$$

From the final row, we have $f(N) = N.r(N) > 0$. Otherwise, for $i < N$, we have $f(i) = i.[r(i) - r(i+1)]$. But $r \in R(N)$ so $r(i) > r(i+1)$, that is $r(i) - r(i+1) > 0$; hence $f(i) > 0$. So we have (A.).

Also,$\Sigma_{i=1}^{N} f(i) = \Sigma_{i=1}^{N-1} f(i) + f(N) = \Sigma_{i=1}^{N-1} i.[r(i) - r(i+1)] + N.r(N) = r(1) + \cdots + r(N)$. But $r \in S(N)$ so $r(1) + \cdots + R(N) = 1$. So we have (B.) and the result is proved. □

3.2 Geometric interpretation of ϕ_N

Figure 1(a) shows a ranked distribution, $r \in R(N)$, as a bar chart consisting of N bars of width 1. That is, the outlined area is subdivided into columns by means of vertical lines. Column i has width 1 and area r(i) and therefore height r(i).

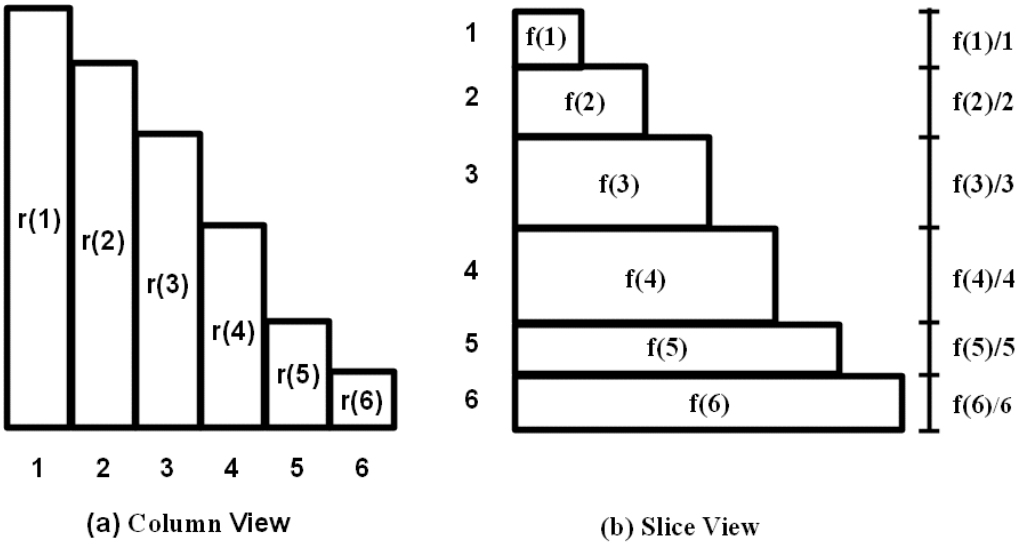


Figure 1: Column and Slice views

Figure 1(b) shows the same outlined area, but subdivided into slices by means of horizontal lines. The areas of those slices define an $f \in S(N)$, where f(i) is the area of slice i. Since slice i has area f(i) and length i it has height $f(i)/i$.

Comparing Figures 1(a) and 1(b) gives $r(i) = \frac{1}{i} f(i) + \cdots + \frac{1}{N} f(N)$, that is $r = \phi_N(f)$. As indicated in the Figure, we shall call Figure 1(a) the column view of r, and Figure 1(b) the slice view.

3.3 Numerical generation of elements of R(N)

The algorithm used to generate $r \in R(N)$ follows from the linear bijection.

- Use the algorithm given in Lovegrove [2] for S(N) to generate $f \in S(N)$;
- Set $r(N) = \frac{1}{N} f(N)$;
- For $i = (N-1), \ldots, 1$ put $r(i) = \frac{1}{i} f(i) + r(i+1)$.

4 Likelinesses over R(N)

4.1 h=0

Since $L_{S(N)}(i) = \frac{1}{N}$ for all i, it follows immediately from Theorem 1 that

$$(\forall i \in X_N) \quad L_{R(N)}(i) = \frac{1}{N}(\frac{1}{i} + \cdots + \frac{1}{N}), \tag{2}$$

a result that has previously been found [1] by using a geometric argument. Graphs of some $L_{R(N)}$ are shown in Figure 2.

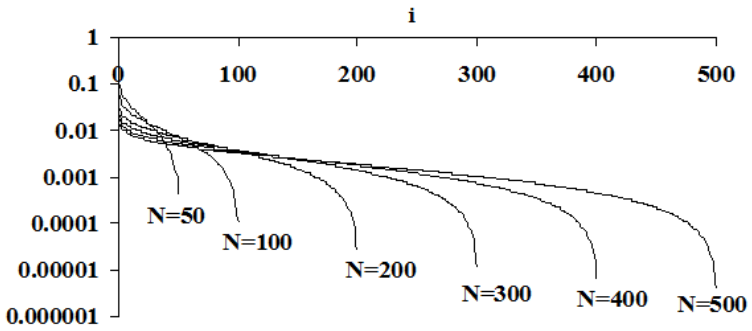


Figure 2: $L_{R(N)}(i)$ for various N

5 Reduction

We might sometimes want to analyse ranked data of an unknown degree or combine data of differing degrees. The purpose behind reduction is to enable these to be done, albeit with some loss of information, by reducing the degree to a smaller, but known, value.

Two types of reduction will be defined: S(M) reduction and R(M) reduction. S(M) reduction is commonplace; it is being given principally for the sake of the notation but also to give a context for R(M) reduction.

5.1 S(M) Reduction

Let $M \in X_N$. Let $f \in S(N)$ and then let $F \in S(M)$ be such that

$$(\forall i \in X_M)\ F(i) = \frac{f(i)}{\Sigma_{j=1}^{M} f(j)}.$$

Then we shall call F the *S(M) reduction* of f. Denote this by $F = \pi_M(f)$.

5.2 R(M) Reduction

Define $\eta_M = \phi_M \pi_M \phi_N^{-1}$ and let $r \in R(N)$. Then we shall call $\eta_M(r)$ the *R(M) reduction* of r.

Geometrically, $\eta_M(r)$ is obtained from r by switching from column view to slice view (ϕ_N^{-1}), carrying out S(M) reduction on the slices (π_M), and then switching back to column view with the reduced number of slices (ϕ_M).

$\eta_M(r)$ can also be interpreted as suppression of the Origin of the graph of r so that the horizontal axis of the graph passes through (M+1, r(M+1)), together with the renormalisation necessary to bring the sum back to 1 (this renormalisation is carried out by the π_M).

Since $R(M) \subset S(M)$ we could apply S(M) reduction to r to remove $r(M+1), \ldots, r(N)$. However, if we were to do so then the result would contain a 'memory' of the original distribution, as shown in Fig 3a. R(M) reduction removes that memory whilst still removing $r(M+1), \ldots, r(N)$, as in Fig 3b.

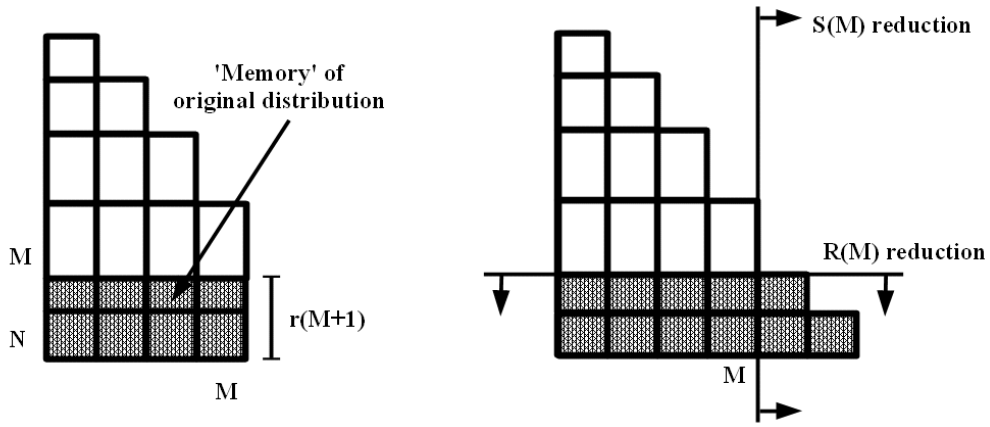
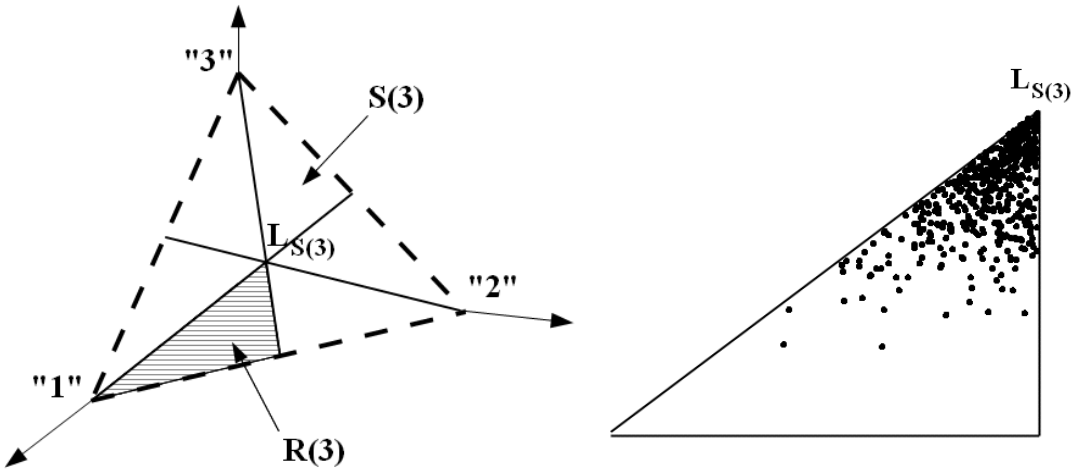


Figure 3: Reduction of ranked distributions

The effect of retaining this 'memory' would be to flatten out the graph of the reduced r, that is to move r towards uniformity, that is towards $L_{S(M)}$. The reason is purely arithmetic: for any given difference $r(1) - r(M)$, the slope after reduction would be less after S(M) reduction than under R(M) reduction because the presence of the 'memory' would mean that the normalisation process would involve division by larger numbers (Figure 3a). Figure 4a shows the position of R(3) within S(3). Figure 4b then shows the effect of S(3) reduction on a randomly-generated sample of elements of R(20): the migration towards $L_{S(3)}$ is clear.



(a) Position of R(3) in S(3)

(b) Effect of S(3) reduction

Figure 4: Effect of S(3) reduction on elements of R(20)

6 Distribution of distributions

If likelinesses are to form a practical means for quantifying the informal concept of "How Likely Something is to Happen" then actual distributions need to be, at least approximately, uniformly distributed over the underlying set.

We would, for example, certainly like to know how distributions generally are distributed over S(N). This question, however, cannot be answered directly because of the bias introduced by authors' predilection for presenting data in ranked order.

Instead, we shall investigate how ranked distributions are distributed over R(N) for some N. This then -because the fundamental linear bijection is invertible- implies the answer to the question about S(N).

In attempting to carry out this investigation we might consider analysing a sample of published elements of R(N). This would be made difficult, however, by the fact that the degree of original data is often not specified, and might not even be known to the original author.

To overcome this difficulty, we could take a sample of ranked data and carry out R(M) reduction for some M.

A sample of 500 ranked distributions of degree at least 10 was collected from published sources, and the R(9) reduction of each was found. The choice of 10 (and thence of the R(9)) was due to the popularity of 'Top 10' rankings in published data.

The selection criteria for inclusion in the sample were:-

- The basis of the ranking must have been quantified, and the quantified values given.
- Degree at least 10.
- The data to be a decreasing, rather than increasing, function of rank. As an important example, in athletics the faster the better so that smaller times rank above larger: although results could be converted into ranked form by giving speeds rather than times, no data was found where this had been done.

- To eliminate, so far as was possible, bias on the part of this Author when carrying out the selection, the data must have been ranked by the original author and described as being "Top 10", "Top 20", "Largest", "Most" etc or as "ranked" and given a rank number starting at 1.

One source of data was rejected because it was described as being 'The Top Ten worst cases', which does not seem to be well-defined and, even if it were, would apparently imply that the data actually formed the bottom ten. Otherwise, there was no selection or rejection by subject matter or because of the actual values.

The 'no selection or rejection by subject matter' resulted in data being included which would not normally be used in scientific analyses but which could not, and indeed should not, have been eliminated. A small selection (not random) of the data sources is given at the end of the References.

The means over all 500 R(9) reductions are compared with $L_{R(9)}$ in Table 1.

i=	1	2	3	4	5	6	7	8	9
Sample Mean	0.35	0.21	0.14	0.10	0.08	0.05	0.04	0.02	0.01
$L_{R(9)}(i)$	0.31	0.20	0.15	0.11	0.08	0.06	0.04	0.03	0.01

Table 1: Means of ranked data compared with ranked likelinesses

Because they are readily plotted on a 2-dimensional page, the R(3) reductions were also found: these are shown in Figure 5.

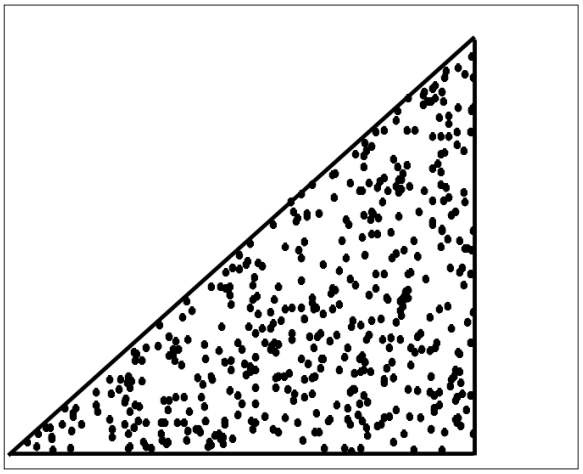
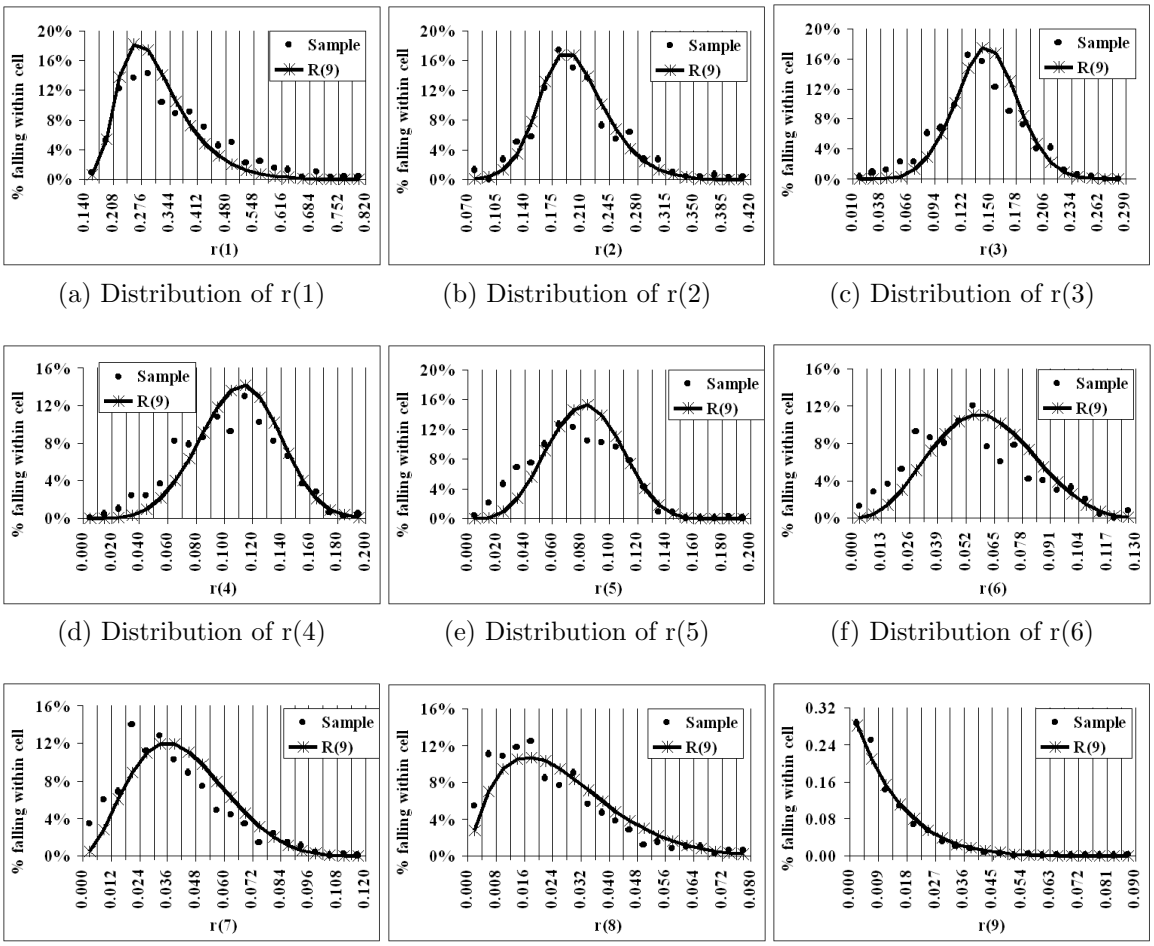


Figure 5: Distribution of R(3)-reductions of data

For $i = 1, \ldots, 9$, the distributions of the R(9)-reductions of the sample are presented in Figure 6, where they are compared with algorithmically-generated results for R(9).



(a) Distribution of r(1) (b) Distribution of r(2) (c) Distribution of r(3)

(d) Distribution of r(4) (e) Distribution of r(5) (f) Distribution of r(6)

(g) Distribution of r(7) (h) Distribution of r(8) (i) Distribution of r(9)

Figure 6: Actual and theoretical distributions

7 Horse racing results

7.1 Wins by a favourite

Since $L_{R(4)}(1) = \frac{1}{4}\{\frac{1}{1} + \ldots \frac{1}{4}\} = 0.52 > 0.5 > 0.46 = \frac{1}{5}\{\frac{1}{1} + \ldots \frac{1}{5}\} = L_{R(5)}(1)$ it follows that, in a competition, the favourite is on average likely to win if there are precisely 4 competitors, but not if there are precisely 5.

Since $L_{R(N)}(1)$ is a strictly decreasing function of N, it follows that the favourite is likely to win if there are 4 *or fewer* competitors, but not if there are 5 *or more*.

The Author collected the results of 30,000 UK horse races, taking the results from various newspapers, teletext and web sites. The proportion of races won by the Favourite is compared with $L_{R(N)}(1)$ in Figure 7a

7.2 Wins by a second favourite

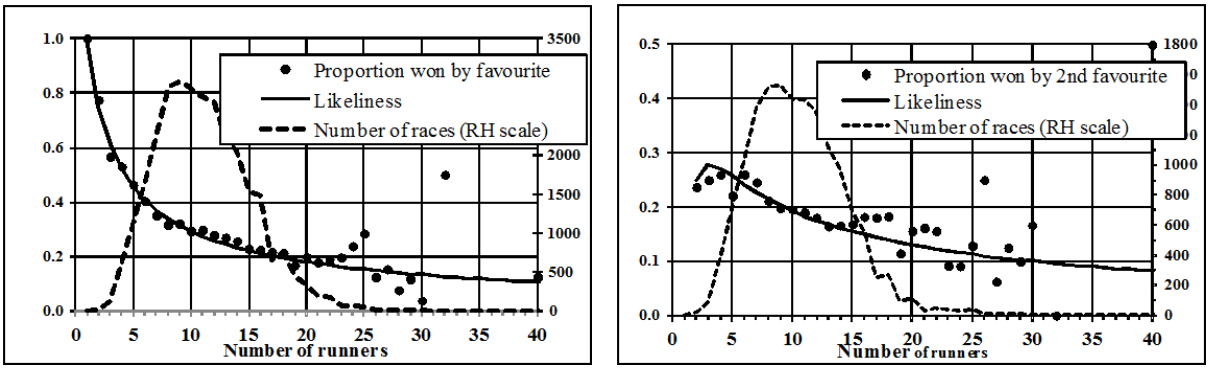
The second favourite can never be likely to win.

$L_{R(N)}(i)$, viewed as a function of N, is strictly decreasing only when $i = 1$. In the case of second favourites ($i = 2$), we have $L_{R(3)}(2) = 0.28 > 0.25 = L_{R(2)}(2)$, so the second favourite is on average more likely to win if there are 3 competitors than if there are 2.

Not all of the sources used to provide the horse-race data in 7.1 indicated when a race had been won by the second favourite. Those that did accounted for approximately half (15,160) of the 30,000 results collected. For those 15,160 races, the proportion won by the second favourite is compared with $L_{R(N)}(2)$ in Figure 7b.

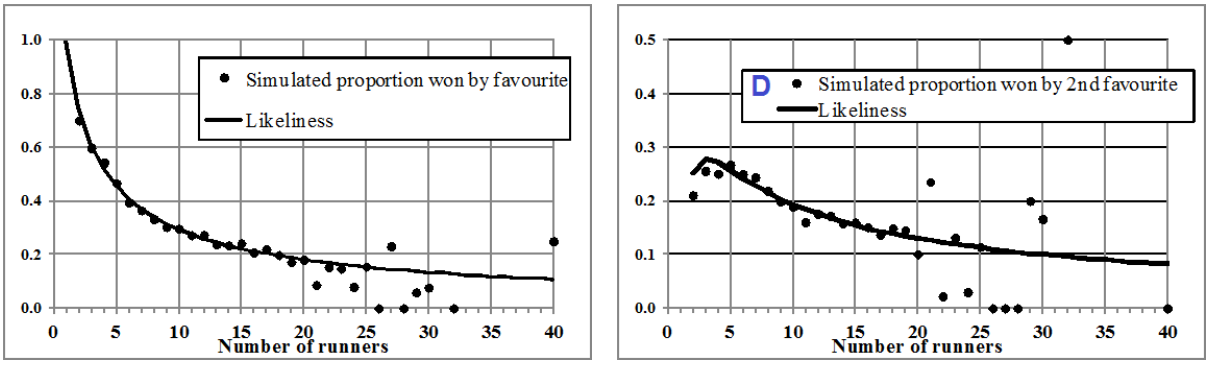
7.3 Simulated results

Since, in the data used to produce Figures 7a & 7b, the number of races is known for each degree (number of runners), it is possible to simulate the data by using the sampling files produced by the Author's program 'Great Likelinesses'. The results are shown in Figures 8a & 8b.



(a) Wins by favourites (b) Wins by second favourites

Figure 7: Actual wins by favourites and second favourites



(a) Simulated wins by favourites (b) Simulated wins by second favourites

Figure 8: Simulated wins by favourites and second favourites

8 Discussion

Likelinesses are a perfect predictor of the mean of a set of relative frequencies if that set of relative frequencies is concentric with the underlying set, P.

In practice, normal observational scatter would result in the two sets not being concentric even if in the limit, as the sample size increases, they would be. In addition, they might not be concentric even in the limit.

How good a predictor the Likelinesses are then depends upon the distance between the centres of the two sets. Strictly, it depends upon the vector joining them.

Even if the Likelinesses are judged to be a good predictor, it does not follow that the relative frequencies are randomly distributed over P: they might, for example, be randomly distributed over some circle which is concentric with P. The extent to which they are, or are not, randomly distributed over P depends upon the distribution of the individual relative frequencies, and how well that distribution compares with the distribution which would be expected from the uniform distribution on P.

With some data sets the concept of individual 'relative frequencies' is meaningless since this requires repetitions, which are not always possible. Under these circumstances, data can be collected, and relative frequencies found, only at the global, rather than individual, level.

This is the case with horse races. If we could repeat a race, with the same runners and riders under identical conditions then we might be able to find the relative frequencies of wins by the favourite (second favourite, etc) for that race. Of course, we can do no such thing; we can merely make a single observation of which horse won, and then collect observations together at the global level -which is where the relative frequencies have to be found. The theoretical background to ranked distributions gives us a way to do this by defining an appropriate classification system.

Figure 7 seems to suggest that ranked likelinesses are, for horse races, a reasonably good predictor of the relative frequencies of wins by the favourites and 2nd favourites. Why this should be so is another matter. It certainly has interesting implications for the handicapping and odds-setting systems, and -on the philosophical side- for the 'wisdom of crowds'.

The comparison of actual and simulated proportions of wins by favourites and second favourites seems to be very good. It is certainly pleasing that the simulated results replicate several details from the actual data. The most noteworthy of these is perhaps the fairly sudden increase in scatter when the sample size (number of races) falls below about 600. This suggests that 600 might be considered as the defining lower limit of 'large sample', but this would need verification from other scenarios.

The situation is different with the sample of 500 ranked distributions. Here, we do have the individual distributions and so can compare the distributions of their individual values with what would be expected on the assumption of a uniform distribution. Figure 6, which does this, does suggest consistency with such an assumption, and is supported by the R(3) reductions in Figure 5.

The distributions are, however, of arbitrary data so it should perhaps not be surprising that, being arbitrary, they are random. It is hard to see how they could be otherwise.

On the theoretical side of things, the linear bijection is of importance because it lies at the heart of the numerical generation of randomly-selected ranked distributions. This

in turn is the basis for the numerical distribution of other types of distribution, such as unimodal, bell-shaped, piecewise unimodal, step-down. So the linear bijection lies at the heart of what is in effect a new subject area.

Another consequence of the linear bijection is that the form of the matrix A_N implies the importance of using R(M)-reduction rather than S(M)-reduction. Figure 4b and the associated discussion make it clear that to use S(M)-reduction on ranked distributions could be misleadingly incorrect.

References

[1] Lovegrove, R, 2008, 'Olympic Games' Medals tables', Mathematics Today, Volume 44, Number 3, pp 121-122 Institute of Mathematics and Its Applications, Southend-On-Sea, Essex

[2] Lovegrove R., 2013, 'The Fundamentals of Likelinesses', Lovegrove Mathematical Services Research Report 2013-01. September 2013.

[3] Mandelbrot, B.B., 1982 'The fractal geometry of nature', W.H.Freeman and Company

[4] Pareto,V., 1986, 'Cours deconomie politique', Droz, Geneva Switzerland

[5] Zipf, G.K., 1949, 'Human Behaviour and the Principle of Least Effort', Addison-Wesley

~

Selection of sources of ranked data

FTSE 100 companies ranked by market capitalisation, downloaded 19 September 2009, http://investing.thisismoney.co.uk/company_performance?country_selected=1&security_classification_id=179097

Top 100 [geographically] largest countries in the World, downloaded 19 Sep 2009, http://www.aneki.com/top_100_largest.html

Top 10 most popular Hardback fiction titles for the week ending 19 Sep 2009, downloaded 20 Sept 2009
http://www.guardian.co.uk/books/table/2009/sep/19/bestsellers-hardback-fiction

UK's Top Ten most popular company names used by email fraudsters. Dowloaded 20 September 2009,
http://www.clearmymail.com/press/natwest_is_phishers_favourite.aspx

Top Ten Most Popular Registered Breeds of Dog in the UK, 2006,
downloaded 20 February 2010,
http://www.top-ten-10.com/recreation/pets/uk-dogs.htm [ex The Kennel Club]

Highest earning TV Characters, Guiness World Records 2008; (2007 Guiness World Records Ltd,) ISBN: 978-1-904994-18-3, P187.

UK investment banking fee ranking. The Times, p 41, 1 January 2010.

Top Ten Most Popular Surnames in Germany, 1995, downloaded 23 January 2010,
http://www.top-ten-10.com/society/family/surnames-german.htm [ex Pratt]

Journals Ranked by Impact: Obstetrics & Gynecology 2008, Thomson Reuters

ScienceWatch.com, 11 October 2009, downloaded 26 January 2010,
http://sciencewatch.com/dr/sci/09/oct11-09_1/

Top Ten Sexiest Tennis Players - Female, December 2008, downloaded 21 February 2010,
http://www.top-ten-10.com/sports/tennis/sexiest-tennis-players.htm
[exTennisReporters.net

A Useful likelinesses

Likeliness $L_P(g)$	Underlying set P	Integram g	Description	Symbolic
0.75	R(2)	$''1''$	top of two ("likely")	$L_{R(2)}(''1'')$
0.61	R(3)	$''1''$	top of three	$L_{R(3)}(''1'')$
0.52	R(4)	$''1''$	top of four	$L_{R(4)}(''1'')$
0.50	S(2)	$''1''$, $''2''$	either of two	$L_{S(2)}(''1'')$, etc
0.46	R(5)	$''1''$	top of five	$L_{R(5)}(''1'')$
0.33	S(3)	$''1''$, $''2''$, $''3''$	any of three	$L_{R(3)}(''1'')$, etc
0.28	R(3)	$''2''$	middle of three	$L_{R(3)}(''2'')$
0.25	R(2)	$''2''$	bottom of two ("unlikely")	$L_{R(2)}(''2'')$
0.11	R(3)	$''3''$	bottom of three	$L_{R(3)}(''3'')$
0.07	R(4)	$''4''$	bottom of four	$L_{R(4)}(''4'')$
0.04	R(5)	$''5''$	bottom of five	$L_{R(5)}(''5'')$

Table 2: Milestone likelinesses

Degree	Rank order, i						
N	1	2	3	4	5	...	N
1	1						
2	0.75	0.25					
3	0.61	0.28	0.11				
4	0.52	0.27	0.146	0.063			
5	0.46	0.26	0.156	0.090	0.040		
6	0.41	0.24	0.158	0.103	0.061		0.028
10	0.29	0.19	0.143	0.110	0.085	...	0.010
12	0.26	0.18	0.134	0.106	0.085	...	0.010
30	0.133	0.100	0.083	0.072	0.064	...	0.0011
50	0.090	0.070	0.060	0.053	0.048	...	0.00040
100	0.052	0.042	0.037	0.034	0.031	...	0.00010
360	0.018	0.015	0.014	0.013	0.012	...	0.000008
500	0.016	0.0116	0.0106	0.0099	0.0094	...	0.000004

Table 3: Ranked likelinesses