

Extreme Value Analysis of Diamond Sizes and Values

Rombouts, Luc

Terraconsult, Oosterveldlaan 273, B-2640 Mortsel-Antwerp, Belgium

At the exploration and evaluation stage, early diamond grades or values are often strongly influenced by the occurrence of a few large stones. At the exploration stage, the initial samples from a kimberlite sent for microdiamond analysis are often strongly influenced by one or two larger stones. At the evaluation stage, diamond parcels, obtained from a small bulk sample of a few hundred tonnes, may contain one or two larger good quality stones that contribute heavily to the average carat price of the parcel. For instance, it is not uncommon in a parcel of a few hundred carats to have one stone contributing for 30% to the total value of the parcel. With initial grade and value estimates so dependent on the presence of one or two stones, it is essential to assess the statistical significance of these extreme values. Diamond size and value distribution have previously been shown to belong to two main classes of distributions, the lognormal and loghyperbolic distributions (Rombouts, 1997). The loghyperbolic distribution is obtained as a mixture of lognormal distributions. In alluvial or marine deposits, diamond size and value distributions are often best fitted to lognormal distributions, because of sorting over long transport distances. In primary deposits, diamond size and value distributions are often loghyperbolic, as the erupting kimberlite magma sampled and mixed many different diamond populations, each with its own lognormal size distribution. Diagnosing lognormal from loghyperbolic distributions is essential in interpreting the tail of the size and value distribution and the significance of a few extreme values. Extreme value theory analyses the statistical significance of extreme values or the tail of the distribution. Classical statistical techniques mainly focus on the behaviour of the average and tend to treat extreme values as undesirable outliers. Extreme value theory on the other hand readily accepts that the extremes form an integral part of the distribution and tries to characterize the tail of the distribution with a single parameter, called the extreme value index and denoted as γ . The extreme value index can be less than, equal to, or greater than zero, with only the two latter cases of importance in diamond geology. If $\gamma=0$, then the distribution has a lognormal tail, while, if $\gamma>0$, the distribution has a loghyperbolic or Pareto (power law) tail. If $\gamma>1$, the distribution has no mean and the average grade or value will depend on the sample size.

Two types of plots are useful in analysing the extreme value index: the QQ- and UH-plots. The QQ-plot is a Pareto-quantile plot and readily detects Pareto-type tails. The UH-plot is more sensitive and general, as it detects both Pareto-type (including loghyperbolic) and lognormal tail behaviour. The UH-plot is preferably used in fitting γ to the tail of the distribution (Caers and Rombouts, 1996).

Very early at the exploration stage, the grade of commercial-sized diamonds in a kimberlite deposit can be roughly estimated by extrapolating the size distribution of the microdiamonds, recovered from a few hundred tonnes of drill samples. If the size distribution has a lognormal tail ($\gamma=0$), the grade of the commercial-sized diamonds can be estimated by integrating the lognormal size distribution of the microdiamonds within the range of 0.01 to 100 carats. If $\gamma>1$, the tail is loghyperbolic or Pareto-like, and an estimate of the grade of the commercial-sized diamonds can be obtained by extrapolating the extreme value distribution defined by γ (giving cumulative grade versus size) beyond the stone sizes larger than 0.01 carat. These estimates are rough, but reliable enough to prioritize kimberlite targets for bulk sampling.

At the evaluation stage, the value distribution of a parcel of diamonds will show a tail, characterized by $\gamma=0$, $0<\gamma<1$ or $\gamma\geq 1$. If the extreme value index of the value distribution equals zero, the average carat price can be estimated by dividing the mean or the t-estimator of the average stone values by the mean of the average stone sizes. If the extreme value index is greater than 1, the average value does not exist, but depends on the sample size. The greater the sample size the more likely a better average carat price is obtained. In such a case, estimates of the average carat price can be obtained for different sample sizes or degrees of mining by extrapolating the extreme value graph and calculating the contribution of extreme values, more likely to occur if the sample is larger or mining is more advanced, to the average value. In case the extreme value index remains between zero and one, the average exists and can be calculated by several methods. The average can be obtained by simply taking the arithmetic averages of the stone sizes and values, or, if the value distribution can be fitted to a three-parameter lognormal distribution, by the t-estimator, or thirdly, by extrapolating the tail of the value distribution using the extreme value graph and adding the contribution of the extreme value tail to the rest of the distribution to obtain the average carat price.

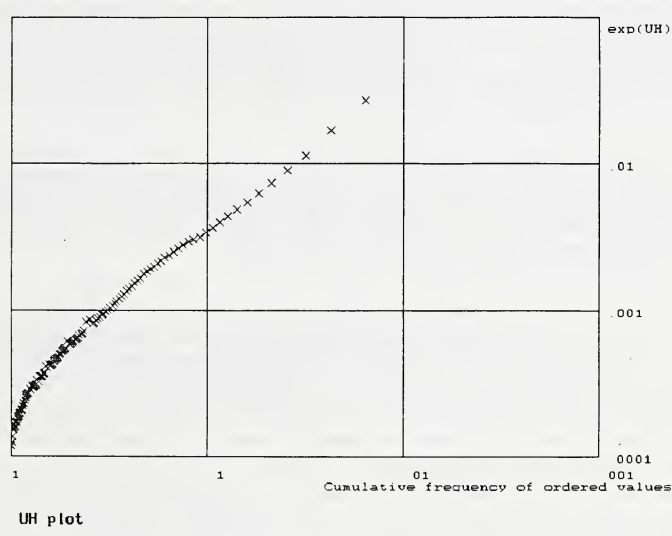


Fig. 1: The microdiamond size distribution of a kimberlite: the UH plot shows a linear trend towards the largest microdiamonds, pointing to a loghyperbolic size distribution.

References

Caers, J., and Rombouts, L., 1996, Valuation of primary diamond deposits by extreme value statistics: *Economic Geology*, Vol. 91, pp. 841-854.

Rombouts, L., 1997, Size and value distributions of diamonds in kimberlites and lamproites: *Russian Geology and Geophysics*, Vol. 38, pp. 599-611.