Statistical Distributions in Empirical Study of Meteorite Fragments

Vladimir V. Vinnikov Maria I. Gritsevich
vvinnikov@list.ru gritsevich@list.ru

Abstract

Hypervelocity atmospheric entry is a complex multiscale physical process lacking straightforward mathematical description due to the large number of factors. Ground-based meteor observations do not directly provide us with all necessary information such as, for example, meteoroid bulk and grain densities, and number of individual meteoroid fragments. Thus, numerical simulations are computationally extensive and involve simplifying assumptions as assumption that meteoroid has fragmented into number of absolutely equal pieces, or consideration of constant main body shape along the whole luminous segment of its trajectory. This paper is concerned with the mathematical tests for widely exploited distribution laws on an available data for meteorite fragments [1]. We start with normal and lognormal distributions of fragment masses in logarithmic and linear scales respectively. Our investigation shows good agreement of the obtained sample with the anticipated statistical functions, especially when bimodality is assumed. Next, we consider more advanced distributions commonly used in fragmentation theory. These functions with carefully calibrated parameters are also fit quite well with the data points. Thus we obtain some integral characteristics of the underlying fragmentation event. At the end we formulate recommendations how to account for meteoroid mass distribution law during fragmentation process.

Introduction

The present paper is concerned with statistical study of meteorite fragments found in expeditions [1, 2, 3, 4]. With account for the meteorite fragments found during illegal searches, the sample counts 218 obtained fragments, what brings it far beyond other instrumentally recorded meteorite falls in terms of number of found pieces. The recovered fragments provide unprecedented opportunity to develop a novel approach involving proved statistical laws in fragmentation modeling [5]. So a reliable fragmentation model is proposed to interpret number of similar fireball events and will significantly decrease the number of free parameters.

Methodology and Statistical Distributions

Available meteorite fragments are described by the values of their absolute mass. Since fragment masses span over four orders of magnitude, it is convenient to treat them in logarithmic scale. Figure 1 shows empirical cumulative distribution function (CDF) for this set of data. The experimental points on the plot have clearly distinguishable shape, which can be fitted by a number of distributions, including normal, logistical and other continuous sigmoid cumulative functions like Weibull.

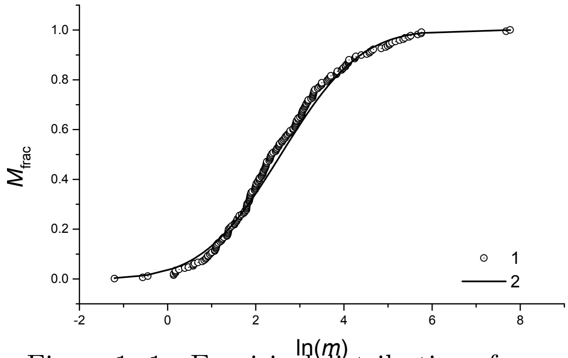

Figure 1: 1 – Empirical distribution of mass fraction M_{frac} less or equal than m; 2 – normal distribution with the mean $\mu = 2.52$ and the standard deviation $\sigma = 1.41$.

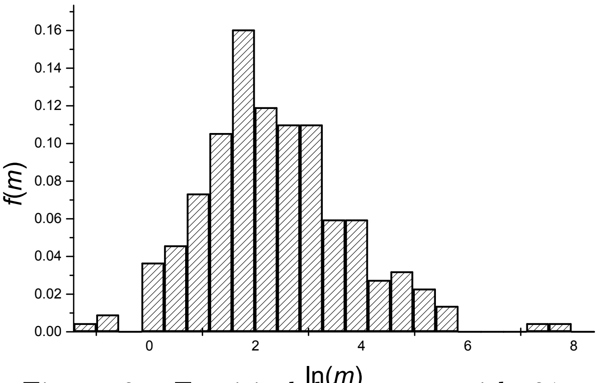

Figure 2: Empirical histogram with 21 uniform sampling mass - $ln\,(m)$ subintervals vs. fraction of total mass $f\,(m)$.

For example the data points can be approximated via normal CDF, which has the form: $F\left(x, \mu, \sigma^2\right) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$. In order to confirm or reject goodness of fit for the selected theoretical distribution we use Pearson's chi-squared test. The main idea behind this test is to compute a kind of normalized relative error between the obtained sample and assumed distribution to compare it against chi-square quantile with desirable significance level. It is important to keep in mind that the null-hypothesis about assumed distribution is not a simple, but a composite one, since CDF parameters (e.g. mean and standard deviation) are a-priori unknown. The abovementioned estimations of these parameters are not applicable directly to the chi-squared test, still there is a way to circumvent a clumsy technique of maximum likelihood. One can first take sample mean and unbiased sample standard deviation as initial values for the required parameters, then vary them to вЂњminimizeвЂќ the chi-squared distance ${\chi^2}_{emp}$ (as stated in Fisher's theorem), and then compare that obtained вЂњminimalвЂќ distance to the desired chi-squared quantile $\left({\chi^2}_{emp}\right)^{-1}(\alpha, K-p-1)$ with specified significance level and degrees of freedom. We consider following parameters for the Pearson's test on the obtained sample. The number of subintervals is $K = 19$ and they are constructed such a way that their estimated sample frequency is no less than 10. Normal distribution has two independent parameters, so $p = 2$. For the chi-squared quantile we take $\alpha = 0.05$. The result for empirical chi-squared statistics is: ${\chi^2}_{emp} = 14.68$ and chi-squared quantile equals: $\left({\chi^2}_{emp}\right)^{-1}(\alpha, K-p-1) = 26.3$. Thus, the null-hypothesis for normal distribution can be accepted since it is valid with 95% probability. Taking into account that a continuous probability distribution of a random variable, whose logarithm is normally distributed, is a log-normal distribution, our finding looks quite reasonable and is in agreement with [6]. However, the first obtained sample permits other similar-shaped continuous distributions, for example logistic function: $F\left(x, \mu, s\right) = 0.5 + 0.5 tanh\left(0.5\left(x-\mu\right)s^{-1}\right)$, where $s = \sigma\sqrt{3}\pi^{-1}$. In this case, Pearson's chi-squared statistics is even less and equals: ${\chi^2}_{emp} = 14.26$. The logistic CDF often considered as a more simple alternative for normal distribution. It has no other advantages per se except of more simple form not involving integration. There are several other goodness-of-fit tests. For the normal distribution one can also use modified Kolmogorov-Smirnov test [7]. It is fulfilled for the presented sample and the above mentioned sample mean μ

and unbiased standard deviation σ. The use of other tests is highly dependent on their complexity. Among simple tests one can also apply G-test, which is slightly more accurate than Pearson chi-squared statistics.

Regardless of acceptance of null-hypothesis for log-normal distribution, we consider other CDFs for better goodness-of-fit values. For a start one can look deeper into the dataset and construct a histogram. The shape of it can hint that the sample points satisfy some kind of superposition of two or more distributions. The plot on Fig.2 shows no significant secondary peaks, though the sample exhibits small local maxima on the left and right tails. In our case this form of statistical representation offers little insight into the possible multimodal nature of the underlying theoretical distribution. Still, there is another approach to statistical investigation of meteorite fragments. It deals with the cumulative number of fragments N instead of the mass fraction of fragments $F(m)$ as in classical CDF. First it‘s crucial to determine if collected data are statistically valid. To check this we construct the $N(\geq m)$ (in ln or log scale) vs. $log(m)$ plot, where $N(\geq m)$ is the number of fragments with mass greater than m (Fig. 3). One can observe a significant gap of data for massive fragments. Actually, the two last points stand aside from remaining empirical distribution and correspond to the weak local maximum on the Fig.2. It is discussible whether to consider this maximum as a second mode, since two data points do not provide any feasible statistics. On contrary, these two data points don‘t distort goodness of fit test significantly. There are still few gaps in the mass samplings which can serve as possible delimiters for various modes.

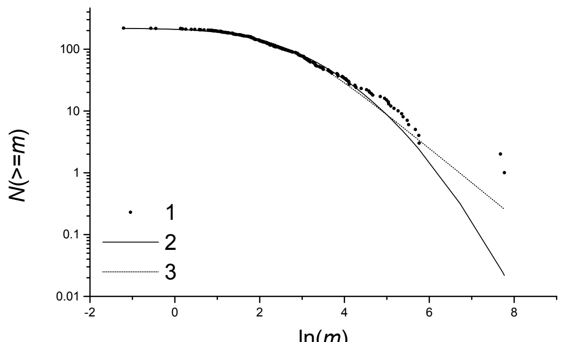

Figure 3: Cumulative number of fragments $N(\geq m)$ vs $ln(m)$ for the sample. 1 – Experimental data, 2 – Normal distribution for $ln(m)$ with the mean $\mu = 2.52$ and the standard deviation $\sigma = 1.41$, 3 – Logistic distribution for $ln(m)$ with the same mean and deviation.

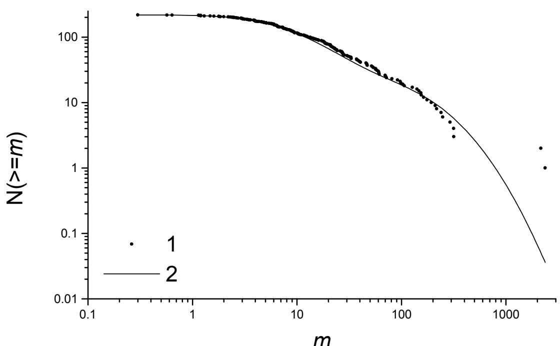

Figure 4: Cumulative number of fragments $N(\geq m)$ vs m (decimal logarithm scale) for the sample. 1 – Experimental data, 2 – Bimodal lognormal distribution with the means $\mu_1 = 2.22$, $\mu_2 = 5.19$, the standard deviations $\sigma_1 = 1.05$, $\sigma_2 = 0.88$ and $\omega = 0.9$.

The common practice for constructing multimodal distribution is to limit the number of modes. Usually bimodal and trimodal CDFs are chosen: $F_{BM}(x) = \omega_1 F_1(x) + (1-\omega_1) F_2(x)$, $F_{TM}(x) = \omega_1 F_1(x) + \omega_2 F_2(x) + (1-\omega_1-\omega_2) F_3(x)$, where $\omega_\bullet$ – are appropriate weight coefficients: $\omega \in [0;1]$, $\omega_1 \geq 0$, $\omega_2 \geq 0$, $\omega_1 + \omega_2 \leq 1$. The reason for such limitation is following. The number of independent parameters for the distribution increases with each additional mode. While such flexibility can be handy to approximate given samples, it also decreases the total degrees of freedom for the chi-squared test and lowers the threshold of the quantile. We apply bimodal log-norm function to the sample, so the resulting shape can be tuned to conform sample data with greater accuracy. Minimizing the functional of empirical chi-squared statistics we obtain suboptimal values for independent parameter $\omega = 0.9$, the means $\mu_1 = 2.22$, $\mu_2 = 5.19$ and the standard deviations $\sigma_1 = 1.05$, $\sigma_2 = 0.88$, that provide empirical chi-square estimation equal to $\chi^2_{emp} = 9.44$ which is well below the threshold of 23.36 for the number of independent parameters $p = 5$, and the number of subintervals $K = 19$. This distribution better follow original sample points than unimodal one (see Fig. 3,4).

Various aspects of fragmentation processes are quite commonly discussed in scientific literature [8, 9, 10, 11, 12, 13, 14]. To simplify our approach we omit theoretical issues involving Rosen-Rammler equation and its implications and focus on mainstream distributions. Apart from above-mentioned extensively used log-norm function, there are other well-known statistical laws dealing with fragmentation and size distribution of particles. The Weibull distribution provides successful empirical description for lifetimes of objects, fatigue data and the size of particles generated by grinding, milling and crushing operations. The CDF for the mass of fragments has the form: $F_W(m, \gamma, \mu) = 1 - M(\geq m)/M_0 = 1 - exp\left(-\left(m/\mu\right)^{\gamma}\right)$, where M_0 is the total expected mass of the sample and μ – the mean mass. The linear exponential distribution, known as Grady distribution [9, 10] represents cumulative number of fragments as $F_{GK}(x, \mu) = 1 - exp\left(-m/\mu\right)$, where μ is the expected mean mass. The Gilvarry distribution [11, 12] is proposed for the same purpose as beforementioned statistical functions. However, it's CDF is defined as the integral of probability density: $F_G(m) = \int_0^m \left(1/\mu\right)\left(M_0/x\right) exp\left(-x/\mu\right) dx$. It's common to evaluate this function numerically by trapeze method on sufficiently fine grids.

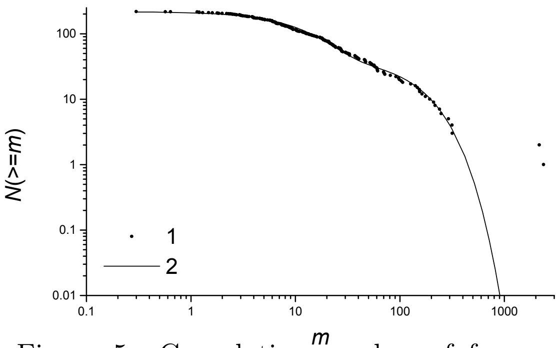

Figure 5: Cumulative number of fragments $N(\geq m)$ vs m for the sample. 1 – Experimental data, 2 – Bimodal Weibull distribution with the weighting factor $omega = 0.8$, $\gamma_1 = \gamma_2 = 1.14$ and $\mu_1 = 13.1$, $\mu_2 = 140$.

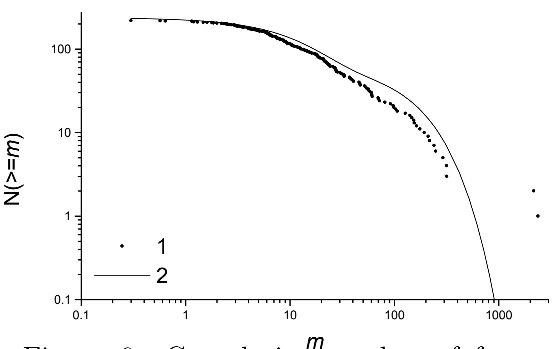

Figure 6: Cumulative number of fragments $N(\geq m)$ vs m for the sample. 1 – Experimental data, 2 – Bimodal Grady distribution with $M_1 = 2047.59$, $\mu_1 = 12$, $M_2 = 9237.09$, $\mu_2 = 140$.

These above mentioned distributions in their bimodal forms are applied to the sample. First we investigate the bimodal Weibull distribution:

$$F_W(m, \omega, \gamma_1, \mu_1, \gamma_2, \mu_2) = \omega\left[1 - e^{-\left(m/\mu_1\right)^{\gamma_1}}\right] + (1-\omega)\left[1 - e^{-\left(m/\mu_2\right)^{\gamma_2}}\right].$$

The parameters for this CDF are the weighting factor ω, the shape and scale γ_1, μ_1 for the first mode and γ_2 and μ_2 for the second one. Therefore, one gets $p = 5$ for computing theoretical chi-squared quantile with significance level α and $K - p - 1$ degrees of freedom. The number of subintervals K is 19. The quantile value is 22.36. The parameters are tuned manually to find suboptimal local minimum for empirical ${\chi^2}_{emp}$. The chi-squared goodness of fit test gives ${\chi^2}_{emp} = 9.89$ for the values $\omega = 0.8$, $\gamma_1 = \gamma_2 = 1.14$, $\mu_1 = 13.1$, $\mu_2 = 140$. This is clearly below the threshold of 22.36, so the Weibull distribution is

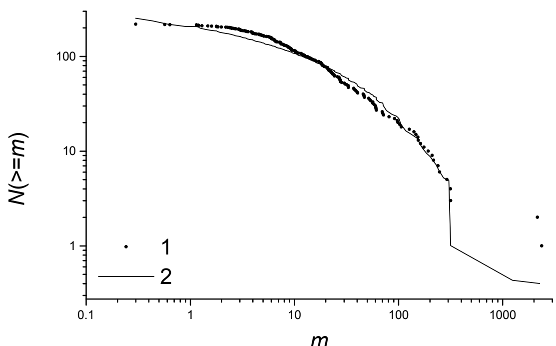

Figure 7: Cumulative number of fragments $N(\geq m)$ vs m for the sample. 1 – Experimental data, 2 – Bimodal Gilvarry distribution with $M_1 = 6743.28$, $\mu_1 = 150$, $M_2 = 4541.4$, $\mu_2 = 2270.7$.

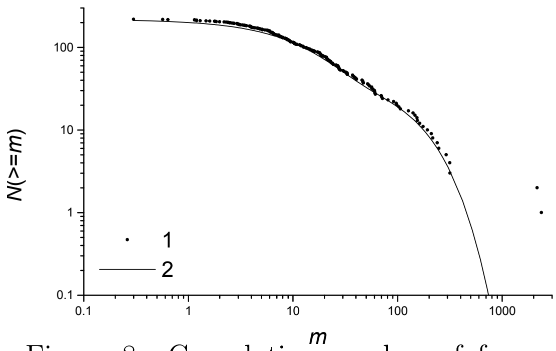

Figure 8: Cumulative number of fragments $N(\geq m)$ vs m for the sample. 1 – Experimental data, 2 – Bimodal sequential fragmentation distribution with $\omega = 0.8$, $\gamma_1 = -0.09$, $\mu_1 = 13.1$, $\gamma_2 = -0.01$, $\mu_2 = 121$.

also suitable for approximation (Fig.5). The cumulative number of fragments distribution is defined as $N_W(\geq x, \bullet) = N \cdot \left(1 - F_W(m, \bullet)\right)$. Next we use the bimodal Grady distribution $F_{GK}(x, M_1, \mu_1, M_2, \mu_2) = \left[\frac{M_1}{\mu_1}\left[1 - e^{-m/\mu_1}\right] + \frac{M_2}{\mu_2}\left[1 - e^{-m/\mu_2}\right]\right]\left(\frac{M_1}{\mu_1} + \frac{M_2}{\mu_2}\right)^{-1}$, where $M_\bullet$, $\mu_\bullet$ – the subtotal mass and the average mass for the first and second modes respectively. The cumulative number of fragments is described as $N_{GK}(x, M_1, \mu_1, M_2, \mu_2) = \frac{M_1}{\mu_1}e^{-m/\mu_1} + \frac{M_2}{\mu_2}e^{-m/\mu_2}$. Chi-squared test yields the value ${\chi^2}_{emp} = 15.11$ for the following arguments: $M_1 = 2047.59$, $\mu_1 = 12$, $M_2 = 9237.09$, $\mu_2 = 140$. The threshold for $K = 20$ and $p = 4$ is $\left(\chi^2\right)^{-1}(\alpha, K - p - 1) = 25$, so the null-hypothesis about Grady distribution is also acceptable with significance level α (Fig.6). The bimodal version of Gilvarry CDF is and the corresponding function for cumulative number of fragments is $F_G(m, M_1, \mu_1, M_2, \mu_2) = \left[\frac{M_1}{\mu_1}\int_0^m \frac{1}{x}e^{-x/\mu_1}dx + \frac{M_2}{\mu_2}\int_0^m \frac{1}{x}e^{-x/\mu_2}dx\right]\left(\frac{M_1}{\mu_1} + \frac{M_2}{\mu_2}\right)^{-1}$. The Gilvarry distribution is has one special aspect in comparison of other considered distributions. As stated in [14] Gilvarry theory overestimates the number of small lightweight fragments. One can see this on the Fig.7. Goodness-of-fit test yields the value $\chi^2_{emp} = 89.85$ for the $M_1 = 6743.28$, $\mu_1 = 150$, $M_2 = 4541.4$, $\mu_2 = 2270.7$. It‘s quite beyond the threshold of $\left(\chi^2\right)^{-1}(\alpha, K - p - 1) = 26.3$, with $K = 21$, $p = 4$. Thus, the obtained sample can‘t be approximated by this distribution. However, the question remains about the completeness of fragments recovery. Some small meteorite parts can completely burn down during descent resulting in the underestimation of true number of fragments. Both Grady and Gilvarry distributions are correct under assumption of nearly-instant singular breaking [13, 14]. If material is exposed to multiple successive fragmentation events, then the above-mentioned statistical laws are no longer applicable. However there is a recent development of the fragmentation theory [8], which can be used. The CDF for sequential fragmentation is $F(m, \mu, \gamma) = 1 - N(\geq m)/N_0 = 1 - exp\left(-\left(m/\gamma\right)^{\gamma+1}\right)$. The corresponding mass distribution has the form: $F(m, \mu, \gamma) = 1 - \frac{M(\geq m)}{M_0} = 1 - \Gamma\left(\frac{\gamma+2}{\gamma+1}, \frac{1}{\gamma+1}\left(\frac{m}{\gamma}\right)^{\gamma+1}\right)\left(\Gamma\left(\frac{\gamma+2}{\gamma+1}\right)\right)^{-1}$, where $\Gamma(\alpha, x) = \int_x^\infty t^{\alpha-1}e^{-t}dt$ is a complementary incomplete gamma function. We should specially note one significant fact about this mass CDF. The ratio of these two gamma functions is essentially the continuous Poisson distribution with parameter $\frac{1}{\gamma+1}\left(\frac{m}{\mu}\right)^{\gamma+1}$ [15], obtained by spreading initial discrete probabilistic measure continuously onto $[0; \infty)$. We use binormal variant of this CDF: $F_{SF}(m, \mu_1, \gamma_1, \mu_2, \gamma_2, \omega) = 1 - \left[\omega \exp\left(-\frac{1}{\gamma_1+1}\left(\frac{m}{\mu_1}\right)^{\gamma_1+1}\right) + (1-\omega)\exp\left(-\frac{1}{\gamma_2+1}\left(\frac{m}{\mu_2}\right)^{\gamma_2+1}\right)\right]$. Corresponding cumulative number of fragments is obtained via relation: $N_{SF}(\geq m, \mu_1, \gamma_1, \mu_2, \gamma_2, \omega) = N_0\left[1 - F_{SF}(m, \mu_1, \gamma_1, \mu_2, \gamma_2, \omega)\right]$. The goodness of fit gives for the values $\omega = 0.8$, $\gamma_1 = -0.09$, $\mu_1 = 13.1$, $\gamma_2 = -0.01$, $\mu_2 = 121$. The number of independent parameters is $p = 5$, and the number of subintervals K is 18. The value of theoretical chi-squared quantile with significance level α and $K - p - 1$ degrees of freedom is 21.03. This threshold is slightly larger then empirical distribution, so the null-hypothesis is still rejected. However the plot shows good agreement (see Fig.8). The considered distributions can provide us some additional information about the sample. According to bimodal lognormal, bimodal Grady and bimodal sequential fragmentation distributions we can assume that two processes took place. One process with the mean fragments mass about 12 g, and another – with the mean mass around 140 g. It can hint on primary singular prefragmentation and consequent atmospheric entry as two independent meteorites with residual masses of approximately 2 and 9 kg respectively (the latter includes two last anomalously heavy pieces of 2 kg each).

Conclusions

The apparatus of mathematical statistics allowed us to build a robust theory for meteoroid mass distribution during fragmentation. In our particulary studied case we recommend the following models for the fragment distributions starting from the best: bimodal Weibull, bimodal Grady and bimodal lognormal. The Weibull and Grady functions have more extensive physical basis than lognormal, despite of the acceptable accuracy of the latter. The model of sequential fragmentation is a quite recent development and still needs more validation. In our current research we assume that meteorite fragments originate from at least two statistically effective pre-entry вЂњbodiesвЂќ or from two major fragmentation events. However, we question the formation of two largest fragments. They neither form their own statistics nor comply with abundant lightweight pieces. The most obvious use of all presented models is the estimation of completeness for obtained collection. One can assume that few large fragments can be missing due to variety of reasons (say two-three pieces of 500-1000 g each).

Acknowledgements

The authors are grateful to Valery N. Tutubalin, Professor of the Department of Probability Theory, Faculty of Mechanics and Mathematics at the Lomonosov Moscow State University, for his preliminary review of the manuscript, fruitful discussion and useful comments he has made.

References

[1] Tóth J., Svoreň J., Borovička J., Spurný P., Igaz A., Kornoš L., Vereš P., Husárik M., Koza, J., Kučera A., Zigo P., Gajdoš Š., Világi J., Čapek D., Krišandová Z., Tomko D., Šilha J., Schunová E., Bodnárová M., Búzová D., Krejčová T. 2013. The Košice meteorite fall: Recovery and strewn field. Meteorit. Planet. Sci, to be submitted

[2] Borovička J., Tóth J., Igaz A., Spurný P., Kalenda P., Haloda J., Svoreň J., Kornoš L., Silber E., Brown P., Husárik M. 2013. The Košice meteorite fall: Atmospheric trajectory, fragmentation, and orbit. Meteorit. Planet. Sci, accepted

[3] Kohout T., Havrila K., Tóth J., Husárik M., Gritsevich M., Britt D., Borovička J., Spurný P., Igaz A., Kornoš L., Vereš P., Koza J., Zigo P., Gajdoš Š., Világi J., Čapek D., Krišandová Z., Tomko D., Šilha J., Schunová E., Bodnárová M., Búzová D., Krejčová T. 2013. Density, porosity and magnetic susceptibility of the Košice meteorites and homogeneity of H chondrite showers. Meteorit. Planet. Sci, submitted

[4] Ozdín D., Uher P., Porubčan V., Tóth J., Svoreň J., Konečný P., Povinec P., Veis P., Sitek J. 2012. Mineralogy, petrography, geochemistry and classification of the Košice meteorite (2010), Meteorit. Planet. Sci, submitted

[5] Gritsevich M., Vinnikov V., Kohout T., Tóth J., Peltoniemi J., Turchak L., A comprehensive study of distribution laws for the fragments of Košice meteorite, 2013, Meteorit. Planet. Sci, to be submitted

[6] Kolmogorov, A.N., On the log-normal distribution of particles sizes during break-up process. // Dokl. Akad. Nauk. SSSR. 1941. XXXI 2, P. 99-101.

[7] Turin Yu.N., Makarov A.A., Statistical data analysis via computer, Moscow, Finance and Statistics, Infra-M, 1998, 528 p.

[8] Brown, W.K., A theory of sequential fragmentation and its astronomical applications// J. Astrophys. Astr. 1989. V. 10, P. 89-112.

[9] Grady D.E., Kipp M.E., Geometric statistics and dynamic fragmentation // J. Appl. Phys. 1985. V. 58, N 3. P. 1210-1222.

[10] Grady D.E., Particle size statistics in dynamic fragmentation // J. Appl. Phys. 1990. V. 68, N. 12, P. 6099-6105.

[11] Gilvarry J.J., Fracture of brittle solids. I. Distribution function for fragment size in single fracture (Theoretical) // J. Appl. Phys. 1961. V. 32, N 3. P. 391-399.

[12] Gilvarry J.J., Bergstrom B.H., Fracture of brittle solids. II. Distribution function for fragment size in single fracture (Experimental) // J. Appl. Phys. 1961. V. 32, N 3. P. 400-410.

[13] V.V. Silvestrov, Application of the Gilvarry Distribution to the Statistical Description of Fragmentation of Solids under Dynamic Loading, Combustion, Explosion, and Shock Waves, V. 40, N. 2, 2004, P. 225-237

[14] V.V. Silvestrov, Fragmentation of a Steel Sphere by a High-Velocity Impact on a Highly Porous Thin Bumper, Combustion, Explosion, and Shock Waves, V. 40, N. 2, (2004), P. 238-252

[15] Ilienko A. On Continuous Versions of Poisson and Binomial Distributions, National Technical University of Ukraine, Kiev, January 12, 2011, electronic presentation `http://goo.gl/MUeVF`

Vladimir V. Vinnikov, Institute of Mechanics, LMSU, Michurinskii prt. 1, 119192, Moscow, Russia; Dorodnitsyn Computing Centre, Russian Academy of Sciences, ul. Vavilova 40, 119333 Moscow, Russia

Maria I. Gritsevich, Institute of Mechanics, LMSU, Michurinskii prt. 1, 119192, Moscow, Russia; Dorodnitsyn Computing Centre, Russian Academy of Sciences, ul. Vavilova 40, 119333 Moscow, Russia; Finnish Geodetic Institute, Geodeetinrinne 2, P.O. Box 15, FI-02431 Masala, Finland