

ALEKSANDER CIANCIARA*

**RESULTS OF THE VERIFICATION OF THE STATISTICAL DISTRIBUTION MODEL OF
MICROSEISMICITY EMISSION CHARACTERISTICS****WYNIKI WERYFIKACJI MODELU OPISUJĄCEGO ROZKŁAD STATYSTYCZNY CECH
EMISJI SEJSMOAKUSTYCZNEJ**

The paper presents the results of research aimed at verifying the hypothesis that the Weibull distribution is an appropriate statistical distribution model of microseismicity emission characteristics, namely: energy of phenomena and inter-event time. It is understood that the emission under consideration is induced by the natural rock mass fracturing. Because the recorded emission contain noise, therefore, it is subjected to an appropriate filtering. The study has been conducted using the method of statistical verification of null hypothesis that the Weibull distribution fits the empirical cumulative distribution function. As the model describing the cumulative distribution function is given in an analytical form, its verification may be performed using the Kolmogorov-Smirnov goodness-of-fit test. Interpretations by means of probabilistic methods require specifying the correct model describing the statistical distribution of data. Because in these methods measurement data are not used directly, but their statistical distributions, e.g., in the method based on the hazard analysis, or in that that uses maximum value statistics.

Keywords: microseismicity emission; rock burst; empirical distribution; Weibull distribution; Kolmogorov-Smirnov goodness-of-fit test; hazard analysis; maximum value statistics

Problematyka oceny stopnia zagrożenia tapaniami jest niezwykle ważnym zagadnieniem i do tej pory nie w pełni rozwiązany. Wstrząsy występują głównie w rejonach zrobów (Cianciara & Cianciara, 2006). Mogą również występować na wybiegu ściany, na skutek uginania się stropu (Marcak, 2012), obserwuje się wówczas wzmożoną aktywność pęknięcia górotworu, co jest przyczyną powstawania emisji sejsmicznej. W pracy przedstawiono wyniki badań mających na celu weryfikację hipotezy, że rozkład Weibull'a stanowi właściwy model opisujący rozkłady statystyczne cech emisji sejsmoakustycznej, a mianowicie: energii zjawisk, odstępów czasu między zjawiskami. Przyjmuje się, że emisja, będąca przedmiotem rozważań, wywołana jest naturalnym pękaniem górotworu. Jednak w praktyce rejestrowana emisja, oprócz zjawisk związanych z pękaniem, może zawierać zakłócenia. Dlatego, na potrzebę badania modelu, jest ona poddawana odpowiednim zabiegom celem usunięcia tych zakłóceń. Badanie prowadzone jest metodą statystycznej weryfikacji hipotezy zerowej o zgodności dystrybuant (rozkładów)

* AGH UNIVERSITY OF SCIENCE AND TECHNOLOGY, FACULTY OF GEOLOGY, GEOPHYSICS AND ENVIRONMENTAL PROTECTION, AL. A. MICKIEWICZA 30, 30-059 KRAKOW, POLAND

empirycznych z dystrybuantą zadaną a priori w formie rozkładu Weibull'a. Ponieważ model opisujący dystrybuantę hipotetyczną jest zadany w formie analitycznej, dlatego jego weryfikację można prowadzić stosując test λ *Kolmogorowa*. Interpretacje prowadzone metodami probabilistycznymi wymagają określenia właściwego modelu opisującego rozkład statystyczny danych pomiarowych. Ponieważ w metodach tych nie wykorzystuje się bezpośrednio danych pomiarowych, lecz ich rozkłady statystyczne, np. w metodzie opartej na analizie hazardu, czy też wykorzystującej statystyki wartości maksymalnych. W trakcie badań stwierdzono, że w około 95% badanych przypadków nie było podstaw do odrzucenia hipotezy zerowej o zgodności rozkładów empirycznych z modelem w formie rozkładu Weibull'a.

Słowa kluczowe: emisja sejsmoakustyczna, pękanie górotworu, cechy emisji sejsmoakustycznej, dystrybuanta empiryczna, rozkład Weibull'a, test λ *Kolmogorowa*, analiza hazardu, statystyki wartości maksymalnych

1. Introduction

The problem of assessing the threat degree of rock burst is a very important issue and has not been fully solved yet. Tremors occurs mainly in the areas of abandoned workings (Cianciara & Cianciara, 2006). They can also occur on the catwalk wall as a result of ceiling sagging (Marcak, 2012), then there are observed the increased activity of cracking of the rock mass, which is the cause of seismic emission. In the literature devoted to the problem of rock-burst hazard and tremor prediction, papers can be found where it is proposed to use microseismicity emission, i.e., low-energy seismic emission (Hardy, 2003), recorded at a frequency band of a more than 10 to 1000 Hz, as an information channel. Interpretation is conducted on the basis of the analysis of emission characteristics. Under consideration are the following emission features: energy of phenomena, time intervals between successive phenomena. Microseismicity phenomena are defined by a set of signals generated in the same source and registered by a number of measurement sensors. These characteristics under consideration have random structure and are treated as random variables. Therefore, to solve the discussed problems, probabilistic methods are applied, such as: hazard method adapted to microseismicity emission analysis (Cianciara et al., 2005, Cianciara & Cianciara, 2006), analysis based on the study of the process of rock mass fracturing (Cianciara, 2010), stochastic analysis of maximum values (Cianciara & Isakow, 2006). Interpretation performed with probabilistic methods is not based directly on the measured data, but on the parameters describing data statistical distributions (Cesca et al., 2012). Therefore, it is necessary to determine (to develop) a proper model describing these distributions. Specifically the interpretation is based on the analysis of the above-mentioned two main features, the energy and time intervals between successive phenomena. In the case of microseismicity emission the logarithms of energy of phenomena are connected with the intervals between the phenomena by means of linear statistical relationships (Pilecki, 1992; Cianciara et al., 2004). From the basics of probability it is known that the distributions of random variables that are interrelated by linear statistical relationships are described by the same model (Papoulis, 1972). In the case of microseismicity emission, it is one universal model in the form of the Weibull distribution. However, distribution (model) parameters can have different values for individual characteristics. We encounter here a difficulty due to the fact that the energies of microseismicity phenomena are not physical quantities, because it is not possible to locate the sources of these phenomena. This energy, called mathematical energy, is determined at the place of recording and defined as the square of the norm of recorded vibration signals. This is why it has additional statistical dispersion associated with random distribution of sources in the rock

mass. The energy, after having been brought to the sources of phenomena, is used in mining seismology to assess the physical energy of tremors (Dubinski, 1995). From here it follows the truthfulness of the statement that the mathematical energy is related by linear statistical dependence with physical energy of the phenomena. It is therefore fully suitable for use in the interpretation of microseismicity emission by means of statistical methods. In contrast, the inter-event time intervals are physical quantities and practically do not depend on the distribution of emission sources in the rock mass. The intervals, as already mentioned above, are linked with the logarithms of the physical energy by a linear statistical relationship (Cianciara, 2000; Cianciara et al., 2004). As a result, the analysis of emission can be performed independently on the basis of these two characteristics. A significant reduction in the uncertainty of interpretation is then achieved.

It is possibly with a large measure of certainty to describe the course of processes preceding the moments of tremors, e.g., increase of size of phenomena during the process of rock mass fracturing. The effectiveness of interpretation carried out with stochastic methods is conditional on whether the adopted a priori model of probability distribution of the characteristics under consideration is appropriate. The correctness of the model, in a statistical sense, can be assessed by testing the hypothesis that the model fits the empirical distribution.

2. A model of the statistical distribution of the main features of microseismicity emission

The main problem in the interpretation by probabilistic methods is to determine the adequate model that describes the data statistical distribution. In the case considered here, as already mentioned in the introduction, it is a model that describes the distribution of the two main characteristics of microseismicity emission. This emission is caused by rock mass cracking and for that reason it is assumed that the characteristic of its probabilistic structure is the same as the seismic emission. Within the range of global seismology and mining seismology, there are known models describing statistical distributions of earthquake magnitude and energy of mining tremors. Therefore, as a first approximation, the Gutenberg-Richter law describing the relationship between the logarithm of the frequency of occurrence of earthquakes and their magnitudes is considered (Lomnitz, 1974). This law is also used in mining seismology, under the name of “energy distribution”, with magnitude replaced by the logarithm of tremor energy. After an appropriate transformation of the law, we obtain an expression that describes the statistical distribution of logarithm of mining tremors (cumulative distribution function) in the form of an exponential model (Lasocki, 1993). It is assumed that microseismicity is a natural extension of seismology towards lower energies (Pilecki, 1995).

Therefore, following the example of seismology and taking into account that in microseismicity there is a linear statistical relationship between the energy of phenomena and the time intervals between events, a generalized model describing the probability distribution of these characteristics can be formulated:

$$-\ln[1 - F(\zeta)] = \begin{cases} \alpha^{-1} \cdot \zeta & \text{for } \zeta \geq 0 \\ 0 & \text{for } \zeta < 0 \end{cases} \quad (1)$$

It is a single-parameter exponential model, where α is the model parameter, and the random variable ζ represents values of particular characteristics, i.e., $\zeta = \ln \frac{E}{E_0}$ for energy E and $\zeta = u - u_0$ for inter-event time u , where E_0 and u_0 are the respective reference values.

Natural modification of the exponential model (1), stemming from practice, is a two-parameter Weibull distribution, which can be expressed in the following form (Cianciara, 2006):

$$-\ln[1 - F(\zeta)] = \begin{cases} \lambda^{-1} \cdot \zeta^\gamma & \text{for } \zeta \geq 0 \\ 0 & \text{for } \zeta < 0 \end{cases} \quad (2)$$

where: λ and γ are model parameters, the other symbols as above.

This model, described by the two-parameter exponential function, is much more effective, than the previous one, in its ability to approximate empirical distributions. Thus, it describes more adequately statistical distributions of microseismic measurement data. The Weibull model is also used in mining seismology to describe the statistical distribution of the logarithm of tremor energy (Lasocki, 1993). When parameter $\gamma = 1$, the Weibull distribution (2) becomes the exponential model (1). In Figures 1 and 2 exemplary Weibull distribution functions against empirical distribution functions are presented: energy of phenomena in Fig. 1 and the time intervals between events in Fig. 2. The calculations were carried out on the basis of the measurement data recorded in December 2005, in the area of longwall 306, seam 507 of Bielszowice Coal Mine. The size of the window from which the measurement data was collected was 8 hours. In the figures good agreement between the two courses is seen, which may evidence that this model is an adequate expression describing the a priori cumulative distribution function. To make this statement more reliable, in the following section a Weibull model verification method is presented as well as the results of fitting the Weibull model to empirical distributions of microseismicity emission characteristics.

Testing the statistical distribution model of the microseismicity emission characteristics

Testing the statistical distribution model was conducted on the basis of microseismicity emission registered in mine workings (galleries). It is assumed that this emission is due to the rock mass fracturing. However, in practice it may contain various types of disturbance. Thus, to get the proper material for testing, measurement series should be cleaned of disturbances. One way to eliminate interference resulting from practice, is operating of maximum events, i.e., those whose discussed characteristics take maximum values. This may be done already during the measurement by recording in a multiple-sensor system. Then those phenomena are recorded, whose signals coming from the same source were registered on a number of sensors, at least two. Later, it is additionally checked whether durations of vibration signals are in the proper relation to the value of the energy of phenomena. On the basis of the data series obtained this way, verification is carried out of the model describing the statistical distributions of the emission characteristics under consideration.

Since the exponential model (1) is a special case of model (2), then the model in the form of a Weibull distribution is tested. First, it is necessary to identify the model, which consists in estimating the parameters of model (2). The calculations were based on the measurement series of

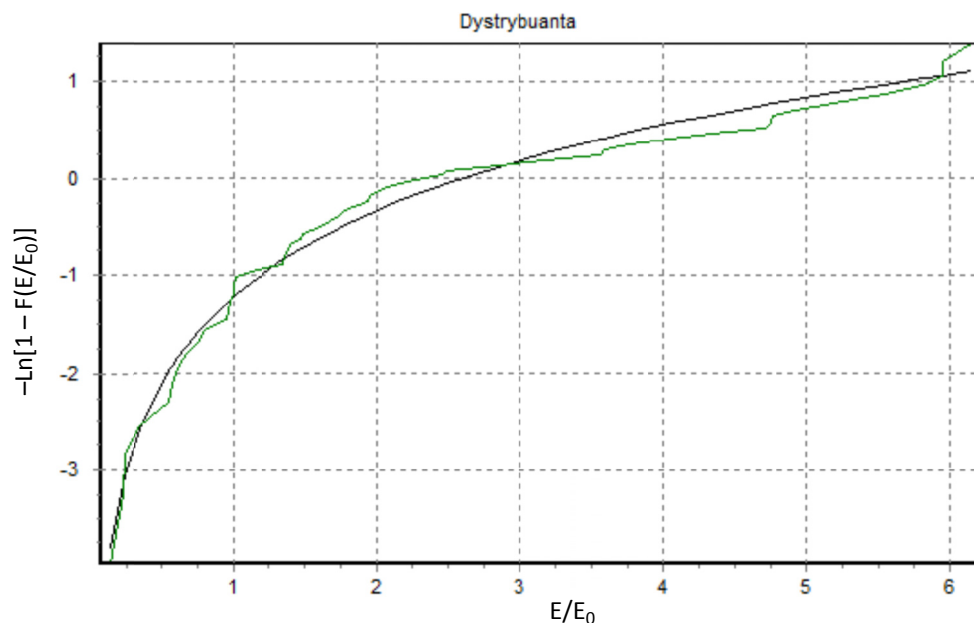


Fig. 1. The exemplary theoretical energy of phenomena Weibull distribution functions against empirical distribution functions. Measurement data recorded in the area of longwall 306, seam 507 of Bielszowice Coal Mine

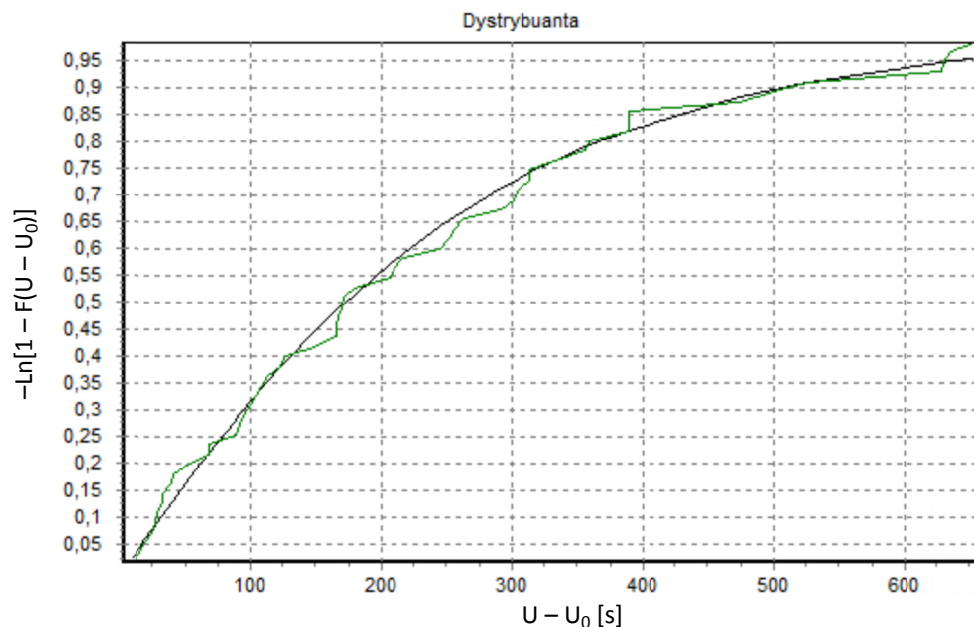


Fig. 2. The exemplary theoretical Weibull distribution functions of the time intervals between events against empirical distribution functions. Measurement data recorded in the area of longwall 306, seam 507 of Bielszowice Coal Mine

individual emission characteristics $\{\zeta_k\}$, $(k = 1, \dots, K)$, taken from the windows (time intervals) of size T . The estimation may be carried out in two ways, by the ML (*maximum likelihood*) method or on the basis of the LS (*least squares*) method. The method of maximum likelihood leads to the solution of the following non-linear system of two equations (Lehman, 1991; Lasocki, 1993):

$$\gamma^{-1} = \frac{\sum_{k=1}^K \zeta_k^\gamma \ln \zeta_k}{\sum_{k=1}^K \zeta_k^\gamma} - \frac{1}{K} \sum_{k=1}^K \ln \zeta_k, \quad \lambda = \frac{1}{K} \sum_{k=1}^K \zeta_k^\gamma \tag{3}$$

This system is solved by successive approximations, starting with the starting value $\gamma_0 = 1$ (Cianciara, 2000):

$$\gamma_l^{-1} = \frac{\sum_{k=1}^K \zeta_k^{\gamma_{l-1}} \ln \zeta_k}{\sum_{k=1}^K \zeta_k^{\gamma_{l-1}}} - \frac{1}{K} \sum_{k=1}^K \ln \zeta_k, \quad \lambda_l = \frac{1}{K} \sum_{k=1}^K \zeta_k^{\lambda_{l-1}} \tag{4}$$

The calculations are carried out until the appropriate (desired) accuracy of estimates $\hat{\gamma}$ and $\hat{\lambda}$ is obtained.

In the second method of estimating the parameters of the discussed model the empirical cumulative distribution function is used. Appropriately converting expression (2) a linear form related to parameters γ and λ can be obtained:

$$\ln \{-\ln [1 - F(\zeta)]\} = \gamma \ln \zeta - \ln \lambda \quad \text{for } \zeta > 0 \tag{5}$$

Replacing in the last expression $F(\zeta)$ with empirical distribution function P_k , we obtain the representation of the expression in the form of observation equation system:

$$\ln [-\ln (1 - P_k)] = \gamma \ln \zeta_k - \ln \lambda + \varepsilon_k, \dots, \zeta_k > 0, \quad \text{for } k = 1, \dots, K \tag{6}$$

where P_k are empirical distribution function values, ε_k random deviations.

The empirical distribution function P_k is defined as follows: the values of the random variable ζ_k are sorted in a non-decreasing order, i.e., $\zeta_n \leq \zeta_m$ for $m > n$ and then the k -th value is assigned the value P_k of the cumulative empirical distribution function (Smirnov & Dunin-Barkowski, 1973):

$$P_k = \frac{k}{K + 1}, \quad k = 1, \dots, K \tag{7}$$

Assuming that the random deviations ε_k have the normal distribution $N(0, \sigma_\varepsilon)$, then the solution of the system of equations (6) can be carried out by the LS method. So the question comes down to the system of two linear equations whose solution is the sought estimates $\hat{\lambda}$ and $\hat{\gamma}$. Based on these estimates a hypothetical cumulative distribution function is determined in the form the Weibull model (2). For its practical verification the Kolmogorov-Smirnov goodness-of-fit test was applied. The subject of the verification is to determine whether the discussed model is the adequate description of statistical distributions of seismoacoustic emission characteristics.

Thus, the null hypothesis $H_0 : P_K = F(\zeta_k)$ is verified on the goodness of fit of the hypothetical cumulative distribution to the empirical cumulative distribution using the Kolmogorov statistics:

$$D = \sup_k |P_k - F(\zeta_k)| \quad (8)$$

where: P_k is empirical distribution function, $F(\zeta_k)$ is the hypothetical distribution.

The verification is carried out calculating characteristic $\lambda = D\sqrt{K}$, where K is the length of the data series. The calculated characteristic λ is referred to the critical value λ_0 read from the tables of critical values $D_\alpha(K)$, where α is the significance level. If $\lambda < \lambda_0$ then there are no reason to reject the null hypothesis, while in the opposite case, when $\lambda \geq \lambda_0$, the null hypothesis should be rejected.

In order to obtain reliable results of the verification of this model, studies have been conducted on the basis of monthly sets of measurement data obtained in various mining and geological conditions. The study was conducted on the basis of the recorded sets of emission phenomena in the following mines, namely: Marcel Coal Mine, Wesola Coal Mine and Bielszowice Coal Mine. The input for research consists of series of the energy $\{E_k\}$ of phenomena and time intervals between events, $\{u_k\}$, ($k = 1, \dots, K$). Sizes T of the windows from which the measurement information was collected were chosen so that the size K of the data in each window was greater than 30 elements ($K > 30$). Also the effect of window T on the verification results was studied by changing window size in the range from 4 to 24 hours, subject to the condition that the size K of the data set was greater than 30.

3. Conclusions

Verification of the null hypothesis H_0 of good fit of the Weibull model to the empirical distributions was carried out on the basis of the measurement data recorded in the aforementioned mines. The model was verified for different emission characteristics sampled in the 4, 8, and 24-hour windows T , while respecting the principle that the size of the data set in each window was greater than thirty ($K > 30$). The verification of the H_0 hypothesis was performed in 300 windows, 100 cases per each 4, 8, 24-hour window. The two characteristics were tested independently. As a result of testing the null hypothesis was verified positively in 95% of cases. The term "positively verified" means that there is no reason to reject the hypothesis H_0 of fit of the Weibull model (2) to empirical distributions. Probably the reason for negative result of the verification, i.e., rejection of the hypothesis H_0 in the remaining 5% of cases, was insufficient elimination of disturbances. It is known that the disturbances have statistical distribution different from model (2). In practice sufficient elimination of disturbances is not always possible, so it is appropriate to verify the hypothesis H_0 on the fit of distributions for each window T . If the hypothesis is rejected then the window should be removed from the measurement series. It was also shown that the size T of the window has no influence on the verification results if the amount of data in each window is greater than 30, $K > 30$.

In summary, it is concluded that the Weibull model (2) adequately describes statistical distributions of seismoacoustic emission characteristics under consideration.

The paper was prepared as part of the project no. PBS2 / B2 / 8/2013 acronym INGEO obtained under the second contest of Applied Research Programme, funded by the National Centre for Research and Development.

References

- Cianciara A., 2000. *System monitorowania zagrożenia wystąpieniem silnych wstrząsów w oparciu o analizę emisji sejsmoakustycznych*. Rozprawa doktorska Biblioteka Główna AGH. Kraków.
- Cianciara A., Cianciara B., Takuska-Węgrzyn E., 2004. *A Method of Evaluating the Threat of Tremors on the Basis of an Analysis of the Degree of Non-homogeneity of the Seismoacoustic Emission Process*. Arch. Min. Sci., Vol. 49, No 3, p. 405-416.
- Cianciara A., Cianciara B., 2006. *The Meaning of Seismoacoustic Emission for Estimation of Time of Mining Tremors Occurrence*. Arch. Min. Sci., Vol 51, No 4, p 463-575.
- Cianciara A., Cianciara B., Isakow Z., 2006. *Sposób monitorowania zagrożenia tąpnięciami oparty na analizie emisji sejsmoakustycznej metodą hazardu sejsmicznego*. Mechanizacja i Automatyzacja Górnictwa, 44, nr 11, s. 5-11.
- Cianciara A., 2010. *Matematyczny opis procesu pęknięcia górotworu na podstawie detekcji emisji sejsmicznej*. Przegląd Górniczy, nr 6/2010, s. 37-41.
- Cesca S., Monna S., Kaiser D., Dahm T., 2012. *Imaging velocity and attenuation anomalies in mining environments using Acoustic Emissions*. EGU.
- Dubiński J., 1995. *Metody obliczania energii sejsmicznej wstrząsów górniczych, „Wstrząsy górnicze – mechanizm, lokalizacja i energia”*. Szkoła Eksploatacji Podziemnej, Seria wykłady nr 8.
- Hardy H.R., Jr., 2003. *Acoustic Emission/Microseismic Activity: Volume 1: Principles, Techniques, and Geotechnical Applications*, Balkem Publishers.
- Lasocki S., 1993. *Weibull distribution as a model for sequense of seismic events induced by mining*. Acta Geophys., Pol. 41, no 2.
- Lehmann E.L., 1991. *Teoria estymacji punktowej*. PWN Warszawa.
- Lomnitz C., 1974. *Global Tectonics and Earthquake Risk*. Elsevier, Amsterdam.
- Marcak H., 2012. *Seismicity in mines due to roof layer bending*. Arch. Min. Sci., Vol. 57, No 1, p. 229-250.
- Papoulis A., 1972. *Prawdopodobieństwo, zmienne losowe i procesy stochastyczne*. WNT, Warszawa.
- Pilecki Z., 1992. *Statystyczna analiza emisji sejsmoakustycznej dla kontroli stanu zagrożenia tąpnięciami*. Studia i Rozprawy 21, CPPGSMiE PAN, Kraków, 91.
- Pilecki Z., 1995. *An Example of Rock Burst Hazard State Control Using a Zonal Seismoacoustic Observation*. Proc. 5th Conf. on Acoustic Emission/Microseismic Activity, Minneapolis, 313-332.
- Smirnow N., Dunin-Barkowski W., 1973. *Kurs rachunku prawdopodobieństwa i statystyki matematycznej dla zastosowań technicznych*. PWN Warszawa.