

Tomasz NIEDOBA
Akademia Górniczo-Hutnicza, Kraków

APROKSYMACJA KRZYWYCH SKŁADU ZIARNOWEGO ZA POMOCA BAYESOWSKICH ESTYMATORÓW PARAMETRÓW W ROZKŁADZIE WEIBULLA

Streszczenie. W artykule przedstawiono porównanie dwóch metod estymacji parametrów krzywych składu ziarnowego, które aproksymowane są za pomocą rozkładu Weibulla. Zastosowano technikę tradycyjnej linearyzacji za pomocą metody najmniejszych kwadratów oraz podejście do estymacji tych parametrów za pomocą twierdzenia Bayesa. Porównania metod dokonano na podstawie wyników procesu kruszenia porfiru w kruszarce szczękowej. Wyniki aproksymacji zostały poddane ocenie statystycznej, przy zastosowaniu jako miary dopasowania odchylenia resztowego.

APPROXIMATION OF PARTICLE SIZE COMPOSITION CURVES BY BAYESIAN ESTIMATORS FOR WEIBULL DISTRIBUTION FUNCTION PARAMETERS

Summary. The paper presents the comparison between two methods of the estimation of parameters of particle size composition curves approximated by the Weibull distribution function. The traditional linearization technique by means of the least squared method as well the modern approach to these parameters estimation by Bayes theorem were applied to that purpose. The comparison was done on the basis of porphyry crushing results given by the experiment conducted in jaw crusher. The results of approximation were evaluated statistically by means of rest deviation, as the measure of accordance.

Wstęp

W wielu badaniach mających na celu porównanie efektów rozdrabniania przeprowadza się aproksymację otrzymywanych krzywych składu ziarnowego. Do najbardziej znanych wzorów aproksymujących krzywe składu ziarnowego należą dystrybuanty rozkładu logarytmiczno-normalnego, Weibulla (RRB), Gaudina-Andrejewa-Schuhmanna (GAS) oraz uciętego rozkładu Weibulla. Podstawową metodą wyznaczania parametrów występujących we wzorach funkcji aproksymujących krzywą składu ziarnowego jest linearyzacja tych

wzorów i zastosowanie metody najmniejszych kwadratów. Jednakże, estymatory parametrów uzyskane tą drogą są często estymatorami obciążonymi. Należy również podkreślić, że krzywa uzyskana po linearyzacji metodą najmniejszych kwadratów minimalizuje sumę $\sum_i (w_i + (au_i + b))^2$, gdzie w_i, u_i oznaczają wartości nowych zmiennych po linearyzacji, a i b – parametry równania regresji, natomiast nie minimalizuje sumy $\sum_i (y_i - \Phi(x_i))^2$, gdzie krzywa $y = \Phi(x)$ jest szukaną funkcją aproksymującą. W efekcie, można szukać rozwiązań bardziej optymalnych.

Podstawy teoretyczne

W pracy zajmiemy się wyznaczeniem estymatorów parametrów występujących w rozkładzie Weibulla, nie za pomocą linearyzacji, lecz metodami bayerowskimi [Box i Tiao, 1992; Domański i Pruska, 2000; Grabski i Jaźwiński, 2001; Krzyśko, 1997; Niedoba, 2003].

Do opisu krzywej składu ziarnowego stosuje się dystrybuantę rozkładu Weibulla (RRB), o równaniu [Tumidajski, 1993]:

$$\Phi(d) = 1 - \exp\left[-\left(\frac{d}{d_0}\right)^c\right] \quad \text{dla } d \geq 0 \quad (1)$$

gdzie d oznacza średnicę ziarna, c oznacza parametr kształtu, a d_0 parametr skali.

Funkcja (1) w układzie $\left(\ln d; \ln \ln \frac{1}{1-\Phi}\right)$ przyjmuje postać liniową $y = ax + b$ i jej parametry można oszacować metodą najmniejszych kwadratów, gdzie $x = \ln d$, $y = \ln \ln \frac{1}{1-\Phi}$, $a = c$, $b = -n \ln d_0$.

Funkcją gęstości rozkładu Weibulla jest funkcja:

$$f(d) = \frac{c}{d_0^c} d^{c-1} \exp\left[-\left(\frac{d}{d_0}\right)^c\right] \quad \text{dla } d \geq 0 \quad (2)$$

Wartość średnia zmiennej o rozkładzie Weibulla zadana jest wzorem [Grabski i Jaźwiński, 2001; Stanis, 2007; Tumidajski, 1993]:

$$m = E(D) = d_0 \Gamma\left(\frac{1}{c} + 1\right) \quad (3)$$

gdzie $\Gamma(p) = \int_0^{\infty} x^{p-1} e^{-x} dx$ jest funkcją gamma.

Stąd otrzymujemy, że:

$$d_0 = \frac{m}{\Gamma\left(\frac{1}{c} + 1\right)} \quad (4)$$

Po podstawieniu do wzoru (2) otrzymujemy:

$$f(d) = \frac{c\alpha^c}{m^c} d^{c-1} \exp\left[-\left(\frac{d\alpha}{m}\right)^c\right] \quad (5)$$

gdzie $\alpha = \Gamma\left(\frac{1}{c} + 1\right)$.

Niech (d_1, \dots, d_n) oznacza realizację próbki losowej. Funkcja wiarygodności odpowiadająca danej próbce losowej ma postać:

$$f(d|c, m) = f(d_1; c, m) \cdot f(d_2; c, m) \cdot \dots \cdot f(d_n; c, m)$$

czyli

$$f(d|c, m) = \frac{c^n \alpha^{nc}}{m^{nc}} \cdot \beta^{c-1} \exp\left[-\left(\frac{\alpha}{m}\right)^c \cdot \gamma\right] \quad (6)$$

gdzie

$$\beta = d_1 \cdot d_2 \cdot \dots \cdot d_n, \quad \gamma = \sum_{i=1}^n d_i^c.$$

Dokonyjmy randomizacji parametrów. Załóżmy, że wartość średnia m ma przy danej wartości parametru c rozkład o gęstości $g_1(m|c)$, natomiast parametr c ma rozkład o gęstości $g_2(c)$. Na podstawie twierdzenia Bayesa łączny rozkład parametrów (m, c) , przy danej średnicy ziarna d ma postać

$$g(m, c|d) = \frac{f(d|m, c) g_1(m|c) \cdot g_2(c)}{\iint_{\Omega} f(d|m, c) g_1(m|c) g_2(c) dm dc} \quad (7)$$

gdzie Ω oznacza obszar zmienności parametrów m i c .

Dla kwadratowej funkcji strat bayesowski estymator parametru m wyraża się wzorem:

$$\hat{m} = E(m|d) = \iint_{\Omega} m g(m, c|d) dm dc \quad (8)$$

Jeżeli uzna się za zasadne przyjęcie różnych rozkładów *a priori* parametrów m i c , możemy wyznaczyć ich estymatory bayesowskie. W pracy przyjęto, że parametr c jest znany (wartość parametru c oszacowano za pomocą linearyzacji funkcji Weibulla). Można zastanowić się, czy nie byłoby korzystnie aproksymować go metodą bayesowską, ale wymaga to dalszego sprawdzenia.

Jako rozkład *a priori* parametru m zaproponowano rozkład trójkątny dany wzorem:

$$g_1(m) = \begin{cases} \frac{1}{\varepsilon^2}(m - m_0 + \varepsilon) & \text{jeżeli } m \in (m_0 - \varepsilon, m_0) \\ -\frac{1}{\varepsilon^2}(m - m_0 - \varepsilon) & \text{jeżeli } m \in (m_0, m_0 + \varepsilon) \end{cases} \quad (9)$$

$$\text{gdzie } m_0 = \frac{1}{n} \sum_{i=1}^n d_i, \varepsilon = \frac{u_p s}{\sqrt{n}}, s^2 = \frac{1}{n} \sum (d_i - m_0)^2,$$

u_p jest kwantylem rozkładu normalnego standaryzowanego.

Po podstawieniu do wzoru (7) otrzymujemy, że

$$g(m|d; c) = \frac{f(d; c|m)g_1(m)}{\int_{m_0-\varepsilon}^{m_0+\varepsilon} f(d; c|m)g_1(m)dm} \quad (10)$$

Natomiast bayesowski estymator parametru m ma postać:

$$\hat{m} = \int_{m_0-\varepsilon}^{m_0+\varepsilon} mg(m|d; c)dm$$

skąd

$$\hat{m} = \frac{\int_{m_0-\varepsilon}^{m_0+\varepsilon} mf(d; c|m)g_1(m)dm}{\int_{m_0-\varepsilon}^{m_0+\varepsilon} f(d; c|m)g_1(m)dm} \quad (11)$$

Wstawiając równania (6) i (7) do (11) i dokonując prostych przekształceń, otrzymujemy:

$$\hat{m} = \frac{\int_{m_0-\varepsilon}^{m_0+\varepsilon} c^n \frac{1}{m^{cn-1}} e^{-\frac{\alpha^c}{m^c} \gamma} u(m)(m - m_0 + \varepsilon u(m))dm}{\int_{m_0-\varepsilon}^{m_0+\varepsilon} \frac{1}{m^{cn}} e^{-\frac{\alpha^c}{m^c} \gamma} u(m)(m - m_0 + \varepsilon u(m))dm} \quad (12)$$

$$\text{gdzie } u(m) = \begin{cases} 1 & \text{dla } m \leq m_0 \\ -1 & \text{dla } m > m_0 \end{cases}$$

W celu implementacji różnych metod aproksymacji krzywych składu ziarnowego produktów kruszenia przeprowadzono proces kruszenia porfiru (w kruszarce szczękowej) i dokonano aproksymacji parametrów rozkładu Weibulla poprzez linearyzację oraz za pomocą metody bayesowskiej [Box i Tiao, 1992, Domański i Pruska, 2000; Grabski i Jaźwiński, 2001]. Wyniki aproksymacji porównano. Obliczeń dokonano za pomocą pakietu STATISTICA PL [Stanisz, 2007] oraz programu Microsoft Excel. Pomocne były też tablice statystyczne [Zieliński, 1972]

Jako miernik dopasowania przyjęto odchylenie resztowe (błąd średniokwadratowy) o wzorze ogólnym

$$s_r = \sqrt{\frac{\sum_{i=1}^n (\Phi_r(d_i) - \Phi_t(d_i))^2}{k-2}} \quad (13)$$

gdzie $\Phi_r(d_i)$ oznacza dystrybuantę teoretyczną Weibulla, a $\Phi_t(d_i)$ dystrybuantę empiryczną, k – liczba klas ziarnowych w próbie.

Wyniki i dyskusja

Badany materiał – porfir poddano 10-stadialnemu procesowi kruszenia. Dla każdego z etapów dokonywano analizy składu ziarnowego oraz aproksymowano go za pomocą rozkładu Weibulla. Do badań wybrano trzy etapy kruszenia, dla których aproksymacja rozkładem Weibulla, otrzymanym metodą linearyzacji, nie dawała zadowalających efektów (wartość odchylenia resztowego była zbyt duża).

W wybranych, kolejnych kruszeniach warunki rozdrabniania przedstawiały się następująco:

1. Porfir o klasie (20, 25) mm, 5 kg, kruszarka szczękowa wypustowa 11 mm (w kolejnych kruszeniach nadawą był produkt z poprzedniego kruszenia).
2. Kruszarka szczękowa, szczelina wypustowa 4,4 mm.
3. Kruszarka szczękowa, szczelina wypustowa 2 mm.

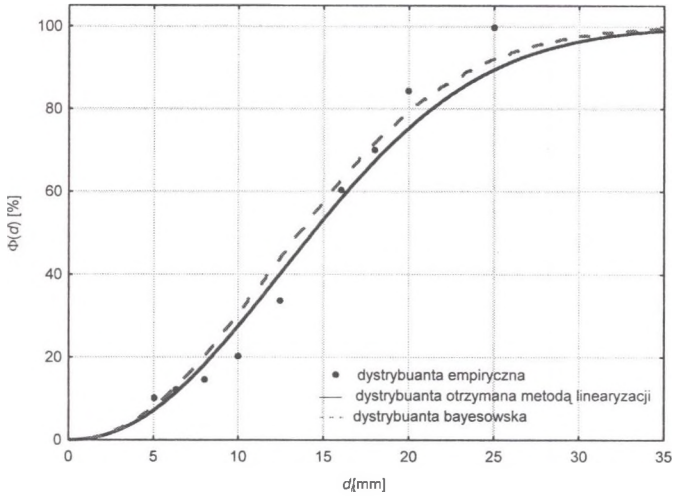
Oceny aproksymacji składu ziarnowego zestawiono w tabelach i na wykresach.

Tabela 1

Poziom 1 kruszenia

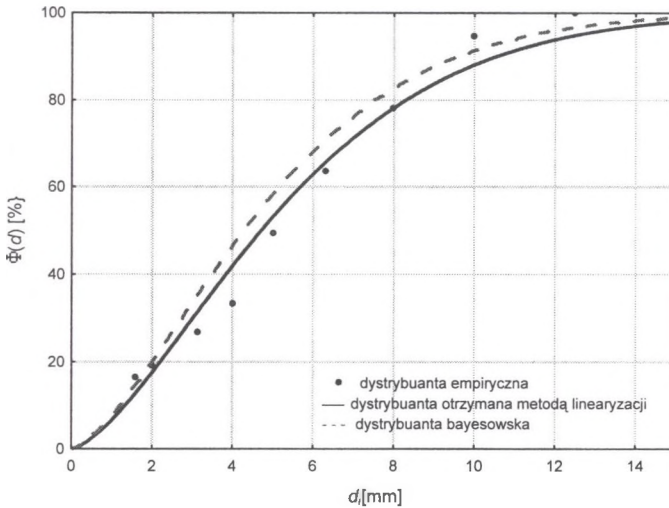
d_i	n_i [%]	\bar{d}_i	$\Phi_r(d)$ *	$\Phi_{t_1}(d)$	$\Phi_{t_2}(d)$
0-5	10,05	2,5	10,05	7,15	7,89
5-6,3	1,94	5,65	11,99	11,41	12,55
6,3-8	2,52	7,15	14,42	18,21	16,56
8-10	6,01	9	20,43	22,74	24,88
10-12,5	13,16	11,25	33,59	33,91	36,78
12,5-16	26,92	14,25	60,51	49,52	62
16-18	9,42	17	69,93	67,43	71,12
18-20	14,40	19	84,33	75,40	78,83
20-25	15,58	22,5	99,91	89,47	91,73

gdzie: $\Phi_r(d)$ - dystrybuanta empiryczna, $\Phi_{t_1}(d)$ - dystrybuanta rozkładu Weibulla, przy parametrach wyznaczonych przez linearyzację, $\Phi_{t_2}(d)$ - dystrybuanta rozkładu Weibulla, gdzie parametry rozkładu zastąpiono estymatorami bayesowskimi \bar{d}_i - średnia średnica ziarna w i -tej klasie ziarnowej.



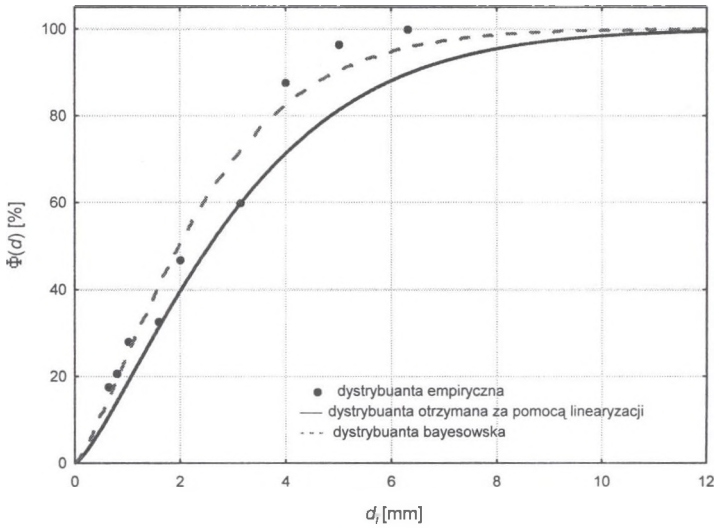
Rys. 1. Zestawienie dystrybuant składu ziarnowego porfiru, otrzymanych różnymi metodami, I poziom kruszenia

Fig. 1. Juxtaposition of porphyry particle size composition distribution functions obtained by various methods, I level of crushing



Rys. 2. Zestawienie dystrybuant składu ziarnowego porfiru, otrzymanych różnymi metodami, II poziom kruszenia

Fig. 2. Juxtaposition of porphyry particle size composition distribution functions obtained by various methods, II level of crushing



Rys. 3. Zestawienie dystrybuant składu ziarnowego porfiru, otrzymanych różnymi metodami, III poziom kruszenia

Fig. 3. Juxtaposition of porphyry particle size composition distribution functions obtained by various methods, III level of crushing

Tabela 2

Estymatory parametrów rozkładu Weibulla

Poziom kruszenia	Estymatory wyznaczone za pomocą linearyzacji		Estymatory bayesowskie	
	c	d_0	c	d_0
I	2,12	17,05	2,12	16,12
II	1,49	6,03	1,49	5,48
III	1,31	3,38	1,31	2,62

Tabela 3

Średnie błędy resztowe s_r

Poziom kruszenia	Metoda linearyzacji	Metoda bayesowska
I	5,49	4,47
II	4,98	3,75
III	8,04	4,13

Wnioski

Aby ocenić jakość aproksymacji (dopasowania), wybrano te poziomy kruszenia, dla których krzywa Weibulla, wyznaczona za pomocą linearyzacji nie była najlepiej dopasowana (wartość odchylenia resztowego była zbyt duża). Za każdym razem estymatory bayesowskie

dawały lepsze dopasowanie. Zasadne jest więc zastanowienie się nad stosowaniem tej metody estymacji parametrów rozkładów przy określaniu teoretycznych krzywych składu ziarnowego i rozkładów innych wielkości charakteryzujących materiał uziarniony.

Artykuł powstał w ramach pracy statutowej nr 11.11.100.276

Bibliografia

1. Box G.E.P., Tiao G.C.: Bayesian inference in statistical analysis, John Wiley and sons INC, New York/Chichester/Brisbane/Toronto/Singapore, 1992.
2. Domański C., Pruska K.: Nielklasyczne metody statystyczne, Polskie Wydawnictwo Ekonomiczne, Warszawa 2000.
3. Grabski F., Jaźwiński J.: Metody bayesowskie w niezawodności i diagnostyce, Wyd. Komunikacji i Łączności, Warszawa 2001.
4. Krzyśko M.: Statystyka matematyczna II, Wydawnictwo Uniwersytetu im. A. Mickiewicza, Poznań 1997.
5. Niedoba T.: Ocena jakości produktu przy pomocy metod bayesowskich na przykładzie złoża „Rudna”, Inżynieria Mineralna, no. 3, p. 134-141, 2003b.
6. Stanisław A.: Przystępny kurs statystyki w oparciu o program STATISTICA PL na przykładach z medycyny, Statsoft, Kraków 2007.
7. Tumidajski T.: Zastosowanie metod statystycznych w analizie procesów przeróbki surowców mineralnych, Śląskie Wydawnictwo Techniczne, Katowice 1993.
8. Zieliński R.: Tablice statystyczne, PWN, Warszawa 1972.

Recenzent: Prof. dr hab. inż. Stanisław CIERPISZ

Abstract

The particle size distribution is one of the most important features concerning the quality of grained materials. It determines its adequacy to many applications, including road industry, building industry and various sorts of powder applications. Dependably on the sort of material, its characteristics and chemical composition, the proper granulation is often the first need of the producer of aggregates. That is why it is so important to approximate the particle size composition as good as it is possible. Usually, the traditional statistical methods are being applied to that purpose giving the Weibull, GSA, log-norm or other distribution functions,

very well known in many applications. Each such distribution function is featured with 1, 2 or more parameters, characteristic for individual case. The most often, the least squared method is being applied to calculate these parameters and this is done on the way of linearization. However, results given by the linearization method are not always sufficiently accordant to the empirical data. The author propose then the other possibility of estimating the distribution function parameters – the method based on the Bayes theorem.

In the paper the results of porphyry crushing were the basis to apply the new technique (so-called Bayesian estimators). The 10-stageous process of crushing gave various empirical results which are not always sufficiently well estimated by the Weibull distribution function given by classical linearization. In these cases, author decided to try to approximate the particle size distribution function by means of the Bayesian estimators. Both methods, linearization and Bayesian estimators were then statistically compared by means of mean squared method, representing the differences between theoretical and empiric values. The new method is useful, especially when the approximation given by classical least squared method gives not adequate results.